# Durrett's Exercises

Thanh Le

May 2024

This file covers the following problems of Durrett's *Probability: Theory of Examples*, $5^{\text{th}}$ edition: 1.2.4, 1.3.1, 1.5.9, 1.6.3, 2.2.1, 2.3.14, 2.4.2, 3.2.4, 3.3.1, 3.4.4, 4.1.2, 4.2.1, 4.3.3

**1.2.4.**

Suppose that $F(x) = P(X \le x)$ is continuous. We will show that $Y = F(X)$ has a uniform distribution on $(0,1)$, that is, if $y \in [0,1]$, $P(Y \le y) = y$.

Let $y \in [0,1]$ be arbitrary. If $y = 1$, then $P(F \circ X \le 1) = 1$, since $F$ is always $\le 1$. Consider $y \in [0,1)$. Let $x = \sup\{r \in \mathbb{R} : F(r) \le y\}$ (as $f$ is continuous, $f(x) = y$). Now, we have

$$\begin{aligned} P(Y \le y) = P(F \circ X \le y) &= P(\{w \in \Omega : F \circ X(w) \le y\}) \\ &= P(\{w \in \Omega : X(w) \le x\}) = F(x) = y, \end{aligned}$$

where the third equality uses the fact that $f$ is increasing.

Thus, $Y$ has a uniform distribution on $(0,1)$. $\square$

**1.3.1**

Suppose that $\mathcal{A}$ generates $\mathcal{S}$. Then, we will show that $X^{-1}(\mathcal{A}) = \{\{X \in A\} : A \in \mathcal{A}\}$ generates $\sigma(X) = \{\{X \in B\} : B \in \mathcal{S}\}$. (Here, as usual, $X$ denotes a measure from $(\Omega, \mathcal{F})$ to $(S, \mathcal{S})$). Here, in other words, we need to show that $\sigma(X^{-1}(\mathcal{A})) = \sigma(X)$. We will do that by mutual containment.

$\boxed{\subseteq}$ We have that $X^{-1}(\mathcal{A}) \subseteq X^{-1}(\mathcal{S})$ (Since $\mathcal{A} \subseteq \mathcal{S}$). Therefore, $\sigma(X^{-1}(\mathcal{A})) \subseteq \sigma(X^{-1}(\mathcal{S}))$. Now, note that $\sigma(X^{-1}(\mathcal{S})) = \sigma(X)$ (by the definition of $\sigma(X)$). Hence, $\sigma(X^{-1}(\mathcal{A})) \subseteq \sigma(X)$.

$\boxed{\supseteq}$ Our idea is to show that $X$ is $\sigma(X^{-1}(\mathcal{A}))$-measurable. (This implies $\sigma(X) \subseteq \sigma(X^{-1}(\mathcal{A}))$).

We will use the following **Lemma**: for any $\sigma$-field $\mathcal{D}$ on $\Omega$, $\mathcal{T} = \{B \in \mathcal{S} : \{X \in B\} \in \mathcal{D}\}$ is a $\sigma$-field in $\mathcal{S}$. (We will prove this at the end of this proof).

Now, construct $\mathcal{T}$ with $\mathcal{D} := \sigma(X^{-1}(\mathcal{A}))$, then $\mathcal{T}$ is a $\sigma$-field containing $\mathcal{A}$. Since $\mathcal{A}$ generates $\mathcal{S}$, $\mathcal{S} \subseteq \mathcal{T}$. Therefore, $\sigma(X) \subseteq \sigma$. But, by definition of $\mathcal{T}$, $\mathcal{T} \subseteq \mathcal{S}$. Hence, $\mathcal{T} = \mathcal{S}$. This means $\{B \in \mathcal{S} : \{X \in B\} \subseteq \mathcal{D}\} = \mathcal{S}$. Hence, $X$ is

$\mathcal{D}$-measurable, i.e., $\sigma(X^{-1}(\mathcal{A}))$-measurable.

So, we have shown that $\sigma(X^{-1}(\mathcal{A})) = \sigma(X)$. □

*Proof of Lemma.* We will prove the Lemma stated above.

Let $\mathcal{D}$ be arbitrary, and $\mathcal{T}$ is defined as stated. We want to show that $\mathcal{T}$ is a $\sigma$-field in $\mathcal{S}$.

By definition of $\mathcal{T}$, $\mathcal{T} \subseteq \mathcal{S}$. Now, to show that $\mathcal{T}$ is a $\sigma$-field, we will show that $\mathcal{T}$ is closed under countable union and complement.

Let $\{B_i\} \in \mathcal{T}$. We have that $\{X \in B_i\} \in \mathcal{D}$. Now, let $B = \cup B_i$, we have

$$\{X \in B\} = \{X \in \cup B_i\} = \cup\{X \in B_i\} \in \mathcal{D},$$

since $\mathcal{D}$ is a $\sigma$-field (hence closed under countable union). Hence, $B \in \mathcal{T}$.

Now, let $B \in \mathcal{T}$ be arbitrary. We have $\{X \in B\} \in \mathcal{D}$. This means

$$\{X \in B^c\} = \Omega \setminus \{X \in B\} = \{X \in B\}^c \in \mathcal{D},$$

since $\mathcal{D}$ is closed under complement. Thus, $B^c \in \mathcal{T}$.

We have shown that $\mathcal{T}$ is closed under countable union and complement, which means $\mathcal{T}$ is a $\sigma$-field. □

**1.5.9**

Suppose that $f$ has $||f||_p = (\int |f|^p d\mu)^{1/p} < \infty$. We will show that there are simple function $\varphi_n$ such that $||f_n - f||_p \to 0$.

From Theorem 2.10 (or around that) in Folland, we know that there exists a sequence $\varphi_n$ of simple functions such that $|\varphi_n| \leq |f|$ and $\varphi_n \to f$ pointwisely.

Note that

$$|\varphi_n - f|^p \leq (|\varphi_n| + |f|)^p \leq (|f| + |f|)^p = 2^p |f|^p.$$

Note that $|f|^p$ is integrable, so $2^p|f|^p$ is also integrable. Also, note that $|\varphi_n - f|^p$ converges pointwisely to 0. Now, by Dominated Convergence Theorem, we yield $\int |\varphi_n - f|^p d\mu \to \int 0 d\mu = 0$. Therefore, $||\varphi_n - f||_p \to 0$, as desired. □

**1.6.3**

Recall that we know (a special case of) Chebyshev's inequality: for a random variable $X$ and any $a \geq 0$, $P(|X| \geq a) \leq \frac{EX^2}{a^2}$.

(i) Let $0 < b \leq a$. We will show that there exists a random variable $X$ with $EX^2 = b^2$ and $P(|X| \geq a) = b^2/a^2$. (This means that the equality in Chebyshev's inequality can be achieved, i.e. the inequality is sharp).

Our goal is just to construct $X$ such that $P(|X| \geq a) = b^2/a^2$ and $EX^2 = b^2$. It's natural to come up with $X : [0,1] \to \mathbb{R}$ (so the probability space is $[0,1]$) such that

$$X(w) = \begin{cases} a & \text{if } 1 \geq w \geq \frac{a^2 - b^2}{a^2} = 1 - \frac{b^2}{a^2} \\ 0 & \text{if } 0 \leq w < 1 - \frac{b^2}{a^2}. \end{cases}$$

From this definition, we have that $P(|X| \geq a) = b^2/a^2$. Moreover,

$$\begin{aligned} EX^2 &= \int X^2 1_{\left[0,1-b^2/a^2\right)} dP + \int X^2 1_{\left[1-b^2/a^2,1\right]} dP \\ &= 0 + a^2 P([1-b^2/a^2, 1]) = a^2 \frac{b^2}{a^2} = b^2, \end{aligned}$$

as desired.

(ii) Let $X$ be a random variable such that $0 < EX^2 < \infty$. We will show that $\lim_{a\to\infty} \frac{a^2 P(|X|\geq a)}{EX^2} = 0$. (Chebyshev inequality only tells $\frac{a^2 P(|X|\geq a)}{b^2} \leq 1$, which is not as sharp as this limit).

We have

$$\begin{aligned} EX^2 &= \int X^2 1_{|X|\geq a} dP + \int X^2 1_{|X|<a} dP \\ &\geq a^2 P(|X| \geq a) + \int X^2 1_{|X|<a} dP. \end{aligned}$$

This is equivalent to

$$a^2 P(|X| \geq a) \leq EX^2 - \int X^2 1_{|X|<a} dP. \qquad (1)$$

Now, note that $X^2 1_{|X|<n}$ ($n \in \mathbb{N}$) is a sequence of non-negative random variable, and that $X^2 1_{|X|<n} \uparrow X^2$. Therefore, by Monotone Convergence Theorem, $\int X^2 1_{|X|<n} dP \uparrow EX^2$. Now, take the limit of both sides of (1), we yield

$$\lim_{a\to\infty} a^2 P(|X| \geq a) \leq EX^2 - EX^2 = 0.$$

Thus, LHS $= 0$ (as it must be $\geq 0$). Now, dividing this equality by the constant $EX^2$, we yield

$$\lim_{a\to\infty} \frac{a^2 P(|X| \geq a)}{EX^2} = 0,$$

as desired. (In the proof for this part, as $0 < EX^2 < \infty$, every step is well-defined). $\square$

**2.2.1**

Let $X_1, X_2, \ldots$ be uncorrelated random variables with $\frac{\mathrm{var}(X_i)}{i} \to 0$ as $i \to \infty$. Let $S_n = X_1 + \cdots + X_n$ and $\nu_n = \frac{ES_n}{n}$ as $n \to \infty$. We will show that $\frac{S_n}{n} - \nu_n \to 0$ in $L^2$ and in probability.

Our idea is to use Theorem 2.2.6. Our triangular array will be $X_{n,i} := X_i$. Let $\mu_n := ES_n$ and $\sigma_n^2 := \mathrm{var}(S_n)$. The theorem say that if $\frac{\sigma_n^2}{b_n^2} \to 0$ (for some sequence $b_n$) then $\frac{S_n - \mu_n}{b_n}$ in $L^2$ and in probability (the theorem itself does not conclude about convergence $L^2$, but its proof does). (Here, our $b_n$ will be $n$).

We will show that $\frac{\sigma_n^2}{n^2} \to 0$. Let $\epsilon > 0$ be arbitrary. As $X_i$'s are uncorrelated, we have

$$\sigma_n^2 = \mathrm{var}(S_n) = \sum_{i=1}^{n} \mathrm{var}(X_i).$$

Thus,

$$\frac{\sigma_n^2}{n^2} = \sum_{i=1}^{n} \frac{\mathrm{var}(X_i)}{n^2} \leq \frac{1}{n} \sum_{i=1}^{n} \frac{\mathrm{var}(X_i)}{i}. \qquad (1)$$

Now, as $\frac{\mathrm{var}(X_i)}{i} \to 0$, there exists an $N$ such that for $n \geq N$, $\frac{\mathrm{var}(X_n)}{n} < \epsilon/2$. Let $M := \sum_{i=1}^{N} \frac{\mathrm{var}(X_i)}{i} < \infty$.

Let $N' \in \mathbb{N}$ $(N' \geq N)$ such that $\frac{1}{N'} < \frac{\epsilon}{2M}$. Then, for any $n \geq N'$, continuing on the inequality in (1), we yield

$$\begin{aligned} \frac{\sigma_n^2}{n^2} &\leq \frac{1}{n}\big(M + \sum_{i=N+1}^{n} \frac{\epsilon}{2}\big) \\ &\leq \frac{1}{N'} M + \frac{\epsilon}{2} < \epsilon. \end{aligned}$$

Thus, $\frac{\sigma_n^2}{n^2} \to 0$. Now, by Theorem 2.2.6, $\frac{S_n - \mu_n}{n} \to 0$ in $L^2$ and in probability. In other words,

$$\frac{S_n - \mu_n}{n} = \frac{S_n - \nu_n n}{n} = \frac{S_n}{n} - \nu_n \to 0$$

in $L^2$ and in probability, as desired. □

**2.3.14**

Let $X_1, X_2, \ldots$ be independent. We will show $\sup X_n < \infty$ a.s. iff $\sum_n P(X_n > A) < \infty$ for some $A$.

$\boxed{\Leftarrow}$ Suppose that there exists $A \in \mathbb{R}$ such that $\sum_n P(X_n > A) < \infty$. Then, by Borel-Cantelli Lemma, $P(X_n > A \text{ i.o.}) = 0$. Now, suppose that $w \in \Omega$ satisfies $(\sup X_n)(w) = \infty$, then we have $w \in (X_n > A \text{ i.o.})$. Therefore, $\{\sup X_n = \infty\} \subseteq (X_n > A \text{ i.o.})$. But, we know $P(X_n > A \text{ i.o.}) = 0$, so $P(\sup X_n = \infty) = 0$. Therefore, we yield $\sup X_n < \infty$ a.s..

$\boxed{\Rightarrow}$ Suppose that $\sup X_n < \infty$ a.s.. Suppose, FSOC, that $\sum_n P(X_n > A) = \infty$ for any $A \in \mathbb{R}$. Fix an $A \in \mathbb{R}$. We know that $X_i$'s are indepedent, so $(X_i > A)$'s are also independent. Now, by the second Borel-Cantelli Lemma, we yield $P(X_n > A \text{ i.o.}) = 1$.

We have the following lemma: let $(\Omega, \mathcal{F}, P)$ be a probability space; let $C, D \in \mathcal{F}$; then if $P(C) = P(D) = 1$, then $P(C \cap D) = 1$. The proof of this lemma goes as follows. As $P(C) = P(D) = 1$, we yield $P(C \cap D) = 1$, $P(C \setminus D) = 0$, and $P(D \setminus C) = 0$. Therefore, $P(C \cap D) = P(C \cup D) - P(C \setminus D) - P(D \setminus C) = 1 - 0 - 0 = 1$.

Now, from the lemma, we yield $P(\cap_{A \in \mathbb{R}}(X_n > A \text{ i.o.})) = 1$. This implies $P(\sup X_n = \infty) = 1$, i.e. $\sup X_n = \infty$ a.s., contradiction.

Thus, there exists $A \in \mathbb{R}$ such that $\sum_n P(X_n > A) < \infty$. □

**2.4.2**

Let $X_0 = (0, 1)$ and define $X_n \in \mathbb{R}^2$ inductively by declaring that $X_{n+1}$ is chosen at random from the ball of radius $|X_n|$ centered at the origin. That is,

$X_{n+1}/|X_n|$ is uniformly distributed on the ball of radius 1 and independent of $X_1, \dots, X_n$. We will show that $n^{-1}\log|X_n| \to -1/2$ a.s. (Here, log means the natural log).

From the definition of $X_n$'s, we know that $\frac{|X_n|}{|X_{n-1}|}$'s are iid random variables (note that $X_n$ is a random variable and the absolute value function is continuous imply that $|X_n|$ is also a random variable (Exercise 1.3.3)). Consider the function log on the domain $[0,1]$, with $\log(0) := \infty$. We have that log is continuous, so $\log\frac{|X_n|}{|X_{n-1}|}$'s are random variables. Furthermore, as $\frac{|X_n|}{|X_{n-1}|}$'s are iid, so are $\log\frac{|X_n|}{|X_{n-1}|}$'s.

We want to apply SSLN for $\log\frac{|X_n|}{|X_{n-1}|}$, so let's compute its expected value. Note that the distribution function of $\frac{|X_n|}{|X_{n-1}|}$ is

$$F(a) = P(\frac{|X_1|}{|X_0|} \le a) = P(|X_1| \le a) = \frac{\pi a^2}{\pi 1^2} = a^2,$$

where the third equal sign is because the point $X_n$ is chosen uniformly. Thus, applying Theorem 1.6.9 for the function $|\log|$ (which is $\ge 0$), we have

$$E\big|\log\frac{|X_1|}{|X_0|}\big| = \int_0^1 |\log(y)|2y\,dy = -\int_0^1 \log(y)2y\,dy = \frac{1}{2} < \infty,$$

by integration by parts. Therefore, we can apply Theorem 1.6.9 for $E\big(\log\frac{|X_1|}{|X_0|}\big)$. We yield

$$E\big(\log\frac{|X_1|}{|X_0|}\big) = \int_0^1 \log(y)2ydy = -\frac{1}{2}.$$

Now, by SLLN,

$$\frac{\sum_{i=1}^n \log\frac{|X_i|}{|X_{i-1}|}}{n} \to -\frac{1}{2}\text{ a.s.}.$$

Note that the numerator on LHS is just $\log\frac{|X_n|}{|X_0|} = \log|X_n|$, so $\frac{\log|X_n|}{n} \to -1/2$ a.s.. □

**3.2.4**

Let $g \ge 0$ be continuous. Suppoe that $X_n \Rightarrow X_\infty$. We will show

$$\liminf_{n\to\infty} Eg(X_n) \ge Eg(X_\infty).$$

As $X_n \Rightarrow X_\infty$, by Theorem 3.2.8, there exist $Y_n$ ($1 \le n \le \infty$) with the same distribution as $X_n$ such that $Y_n \to Y_\infty$ a.s.. As $g$ is continuous, by Exercise 1.3.3, $g(Y_n) \to g(Y_\infty)$ a.s.. Also, we know that $g \ge 0$, so $g(Y_n) \ge 0$ for all $n$. So, by Fatou's Lemma,

$$\liminf_{n\to\infty} Eg(Y_n) \ge E\big(\liminf_{n\to\infty} g(Y_n)\big) = Eg(Y_\infty).$$

But, as $X_n$ and $Y_n$ have same distrubtion, we know $Eg(X_n) = Eg(Y_n)$ (for any $1 \leq n \leq \infty$). Therefore,

$$\liminf_{n\to\infty} Eg(X_n) \geq Eg(X_\infty).$$

□

**3.3.1**

Suppose $\varphi$ is a characteristic function (ch.f.). We will show that $\text{Re}(\varphi)$ and $|\varphi|^2$ are also ch.f.'s.

We will prove for $|\varphi|^2$ first. Say $\phi$ is the ch.f. for a r.v. $X$, i.e. $\varphi(t) = Ee^{itX}$. Let $X_1$ be a r.v. that have the same distribution as $X$. Let $X_2$ be a r.v. that have the same distribution as $-X$ and that is independent to $X_1$. Then, the ch.f. of $X_2$ is

$$\varphi_2(t) = Ee^{it(-X)} = E\cos(it(-X)) + iE\sin(it(-X)) = E\cos(itX) - iE\sin(itX) = \overline{\varphi(t)}.$$

Now, as $X_1$ and $X_2$ are independent, by Theorem 3.3.2, $X_1 + X_2$ has ch.f.

$$\varphi(t)\varphi_2(t) = \varphi(t)\overline{\varphi(t)} = |\varphi(t)|^2.$$

Now, we will prove that $\text{Re}(\varphi)$ is also a ch.f.. Consider $X_1$ and $X_2$ as above, i.e. $X_1$ has the same distribution as $X$; $X_2$ the same distribution as $-X$; $X_1$ and $X_2$ are independent. Let $F_1, F_2$ be the distribution functions of $X_1, X_2$, respectively. Then, by Lemma 3.3.9, $\frac{1}{2}F_1 + \frac{1}{2}F_2$ is a distribution function, and has ch.f.

$$\frac{1}{2}\varphi(t) + \frac{1}{2}\varphi_2(t) = \text{Re}(\varphi(t)).$$

*(Explanation for why $F := \frac{1}{2}F_1 + \frac{1}{2}F_2$ is a distribution function: By Theorem 1.2.2, we just need to check 3 conditions: $F$ is non-decreasing; $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$; $F$ is right continuous. Then, the 3 conditions hold because $F_1$ and $F_2$ both satisfy all such conditions, by Theorem 1.2.1).* □

**3.4.4**

Let $X_1, X_2, \ldots$ be i.i.d. with $X_i \geq 0$, $EX_i = 1$, and $\text{var}(X_i) = \sigma^2 \in (0, \infty)$. We will show that $2(\sqrt{S_n} - \sqrt{n}) \Rightarrow \sigma\chi$.

The idea is to manipulate $2(\sqrt{S_n} - \sqrt{n})$ so that we can use the Central Limit Theorem (CLT). Indeed, we have

$$2(\sqrt{S_n} - \sqrt{n}) = \frac{S_n - n}{\sigma n^{1/2}} \cdot \frac{2\sigma\sqrt{n}}{\sqrt{S_n} + \sqrt{n}}. \qquad (1)$$

Now, by CLT, we have that $\frac{S_n - n}{\sigma\sqrt{n}} \Rightarrow \chi$. Furthermore, by SLLN, we have that $\frac{S_n}{n} \to EX_1 = 1$ a.s.. This implies

$$\frac{2\sigma\sqrt{n}}{\sqrt{S_n} + \sqrt{n}} = \frac{2\sigma}{\sqrt{\frac{S_n}{n}} + 1} \to \frac{2\sigma}{1+1} = \sigma \text{ a.s.}.$$

Now, from (1), and by Exercise 3.2.14 (which says $Y_n \Rightarrow Y$ and $Z_n \Rightarrow c$ then $Y_n Z_n \Rightarrow cY$), we yield

$$2(\sqrt{S_n} - \sqrt{n}) \Rightarrow \sigma\chi.$$

□

**4.1.2**

Suppose that $a > 0$. We will show $P(|X| \geq a \big| \mathcal{F}) \leq \frac{1}{a^2} E(X^2|\mathcal{F})$. (Note that this is a version of Chebyshev's inequality, for conditional expectation!)

By definition, we have that $P(|X| \geq a \big| \mathcal{F}) = E(1_{(|X|\geq a)}|\mathcal{F})$. Note that $1_{(|X|\geq a)} \leq \frac{X^2}{a^2}$ (because if $|X| < a$, $1_{(|X|\geq a)} = 0$; otherwise, $1_{(|X|\geq a)} = 1 \leq \frac{X^2}{a^2}$). Therefore, let $A \in \mathcal{F}$ be arbitrary, we have

$$\int_A 1_{(|X|\geq a)} dP \leq \int_A \frac{1}{a^2} X^2 dP = \frac{1}{a^2} \int_A X^2 dP.$$

By definition of conditional expectation, the above is equivalent to

$$\int_A E(1_{(|X|\geq A)}|\mathcal{F}) dP \leq \frac{1}{a^2} \int_A E(X^2|\mathcal{F}) dP, \qquad (1)$$

and this is true for any $A \in \mathcal{F}$.

Now, we will use the following *Lemma*: if $X_1, X_2$ are $\mathcal{F}$-measurable r.v.'s, and for any $A \in \mathcal{F}$, we have $\int_A X_1 dP \leq \int_A X_2 dP$, then $X_1 \leq X_2$ a.s. on $\Omega$. (We will show the proof for this Lemma at the end of this problem).

Applying the lemma, we immediately yield that $E(1_{(|X|\geq A)}|\mathcal{F}) \leq E(X^2|\mathcal{F})$ a.s., as desired.

*Proof of Lemma.* Consider $\epsilon < 0$ be arbitrary. Let $Y := X_2 - X_1$, we know that $\int_A Y dP \geq 0$ for any $A \in \mathcal{F}$. Let $A_\epsilon = (Y \in (-\infty, \epsilon))$. As $Y$ is $\mathcal{F}$-measurable, $A_\epsilon \in \mathcal{F}$. Thus, we have that $\int_{A_\epsilon} Y dP \geq 0$. This implies

$$0 \leq \int_{A_\epsilon} Y dP \leq \int_{A_\epsilon} \epsilon dP = \epsilon P(A_\epsilon).$$

As $\epsilon < 0$, we have that $P(A_\epsilon) \leq 0$. This means $P(A_\epsilon) = 0$. As this fact is true for any $\epsilon < 0$, we yield that $P(Y \in (-\infty, 0)) = 0$. Therefore, $Y \geq 0$ a.s.. □

**4.2.1**

Suppose that $X_n$ is a martingale wrt $\mathcal{G}_n$. Let $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. We will show that $\mathcal{G}_n \supset \mathcal{F}_n$ and $X_n$ is a martingale wrt $\mathcal{F}_n$.

As $X_n$ is a martingale wrt $\mathcal{G}_n$, $X_n$ is adapted to $\mathcal{G}_n$. That means $X_1, \ldots, X_n$ are $\mathcal{G}_n$-measurable. But, $\mathcal{F}_n$ is the smallest $\sigma$-algebra that makes $X_1, \ldots, X_n$ measurable. Therefore, $\mathcal{F}_n \subset G_n$.

By definition, we have that $\mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n$ (that is, $\mathcal{F}_n$ is a filtration). Now, for any $n$, we have

$$E(X_{n+1}|\mathcal{F}_n) = E\big(E(X_{n+1}|\mathcal{G}_n)\big|\mathcal{F}_n\big) = E(X_n|\mathcal{F}_n) = X_n,$$

where the first equality is by Theorem 4.1.13, the second because $X_n$ is a martingale wrt $\mathcal{F}_n$, and the third due to $X_n \in \mathcal{F}_n$. Therefore, $X_n$ is a martingale wrt $\mathcal{F}_n$. □

**4.3.3**

Let $X_n, Y_n$ be positive, integrable, and adapted to $\mathcal{F}_n$. Suppose that $E(X_{n+1}|\mathcal{F}_n) \leq X_n + Y_n$, with $\sum Y_n < \infty$ a.s.. We will show that $X_n$ converges a.s. to a finite limit.

We have that

$$\begin{aligned} E(X_{n+1} - \sum_{k=1}^{n} Y_k|\mathcal{F}_n) &= E(X_{n+1}|\mathcal{F}_n) - E(\sum_{k=1}^{n} Y_k|\mathcal{F}_n) \\ &= E(X_{n+1}|\mathcal{F}_n) - \sum_{k=1}^{n} Y_k \\ &\leq X_n - \sum_{k=1}^{n-1} Y_k, \end{aligned}$$

where the second equality is because $Y_n$ is adapted to $\mathcal{F}_n$, and the inequality is due to our hypothesis. Furthermore, note that as $X_n$ and $Y_n$ are adapted to $\mathcal{F}_n$, $X_n - \sum_{k=1}^{n-1} Y_k$ is adapted to $\mathcal{F}_n$. Therefore, $X_n - \sum_{k=1}^{n-1} Y_k$ is a supermartingale.

Now, let $N = \inf_k \left(\sum_{m=1}^{k} Y_m > M\right)$ for some $M > 0$. $N$ is a random variable, and we have that $\{N = n\} \in \mathcal{F}_n$ for any $n$, so $N$ is a stopping time. Then, by Theorem 4.2.9, $Z_n := X_{n \wedge N} - \sum_{k=1}^{(n \wedge N)-1} Y_k$ is a supermartingale. Furthermore, by the definition of $N$, we have $Z_n + M$ is a positive supermartingale. Therefore, by Theorem 4.2.12, $Z_n + M$ converges a.s. to a finte limit. This means then that $Z_n$ converges a.s. to a finite limit.

Now, let $M \to \infty$. Then, $N \to \infty$. Therefore, $X_{n \wedge N} = X_n$, and $\sum_{k=1}^{(n \wedge N)-1} Y_k = \sum_{k=1}^{n-1} Y_k$. Now, we have $n \to \infty$, $Z_n$ converges a.s. to a finite limit, and $\sum_{k=1}^{\infty} Y_k$ is finite, so $X_n = Z_n + \sum_{k=1}^{n-1}$ converges a.s. to a finite limit. □