

17.5: Convergence

Basic Theory

Basic Assumptions

As in the Introduction, we start with a stochastic process $\mathbf{X} = \{X_t : t \in T\}$ on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, having state space \mathbb{R} , and where the index set T (representing time) is either \mathbb{N} (discrete time) or $[0, \infty)$ (continuous time). Next, we have a filtration $\mathfrak{F} = \{\mathcal{F}_t : t \in T\}$, and we assume that \mathbf{X} is *adapted* to \mathfrak{F} . So \mathfrak{F} is an increasing family of sub σ -algebras of \mathcal{F} and X_t is measurable with respect to \mathcal{F}_t for $t \in T$. We think of \mathcal{F}_t as the collection of events up to time $t \in T$. We assume that $\mathbb{E}(|X_t|) < \infty$, so that the mean of X_t exists as a real number, for each $t \in T$. Finally, in continuous time where $T = [0, \infty)$, we need the additional assumptions that $t \mapsto X_t$ is right continuous and has left limits, and that the filtration \mathfrak{F} is standard (that is, right continuous and complete). Recall also that $\mathcal{F}_\infty = \sigma(\bigcup_{t \in T} \mathcal{F}_t)$, and this is the σ -algebra that encodes our information over all time.

The Martingale Convergence Theorems

If \mathbf{X} is a sub-martingale relative to \mathfrak{F} then \mathbf{X} has an *increasing property* of sorts: $E(X_t | \mathcal{F}_s) \geq X_s$ for $s, t \in T$ with $s \leq t$. Similarly, if \mathbf{X} is a super-martingale relative to \mathfrak{F} then \mathbf{X} has a *decreasing property* of sorts, since the last inequality is reversed. Thus, there is hope that if this increasing or decreasing property is coupled with an appropriate boundedness property, then the sub-martingale or super-martingale might converge, in some sense, as $t \rightarrow \infty$. This is indeed the case, and is the subject of this section. The martingale convergence theorems, first formulated by Joseph Doob, are among the most important results in the theory of martingales. The *first martingale convergence theorem* states that if the expected absolute value is bounded in the time, then the martingale process converges with probability 1.

Suppose that $\mathbf{X} = \{X_t : t \in T\}$ is a sub-martingale or a super-martingale with respect to $\mathfrak{F} = \{\mathcal{F}_t : t \in T\}$ and that $\mathbb{E}(|X_t|)$ is bounded in $t \in T$. Then there exists a random variable X_∞ that is measurable with respect to \mathcal{F}_∞ such that $\mathbb{E}(|X_\infty|) < \infty$ and $X_t \rightarrow X_\infty$ as $t \rightarrow \infty$ with probability 1.

Proof

The proof is simple using the up-crossing inequality. Let $T_t = \{s \in T : s \leq t\}$ for $t \in T$. For $a, b \in \mathbb{R}$ with $a < b$, let $U_t(a, b)$ denote the number of up-crossings of the interval $[a, b]$ by the process \mathbf{X} on T_t , and let $U_\infty(a, b)$ denote the number of up-crossings of $[a, b]$ by \mathbf{X} on T . Recall that $U_t \uparrow U_\infty$ as $t \rightarrow \infty$. Suppose that $\mathbb{E}(|X_t|) < c$ for $t \in T$, where $c \in (0, \infty)$. By the up-crossing inequality,

$$\mathbb{E}[U_t(a, b)] \leq \frac{1}{b-a} [|a| + \mathbb{E}(|X_t|)] \leq \frac{|a| + c}{b-a}, \quad n \in \mathbb{N} \quad (17.5.1)$$

By the monotone convergence theorem, it follows that

$$\mathbb{E}[U_\infty(a, b)] < \frac{|a| + c}{b-a} < \infty \quad (17.5.2)$$

Hence $\mathbb{P}[U_\infty(a, b) < \infty] = 1$. Therefore with probability 1, $U_\infty(a, b) < \infty$ for every $a, b \in \mathbb{Q}$ with $a < b$. By our characterization of convergence in terms of up-crossings, it follows that there exists a random variable X_∞ with values in $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, \infty\}$ such that with probability 1, $X_t \rightarrow X_\infty$ as $t \rightarrow \infty$. Note that X is measurable with respect to \mathcal{F}_∞ . By Fatou's lemma,

$$\mathbb{E}(|X_\infty|) \leq \liminf_{t \rightarrow \infty} \mathbb{E}(|X_t|) < \infty \quad (17.5.3)$$

Hence $\mathbb{P}(X_\infty \in \mathbb{R}) = 1$.

The boundedness condition means that \mathbf{X} is bounded (in norm) as a subset of the vector space \mathcal{L}_1 . Here is a very simple, but useful corollary:

If $\mathbf{X} = \{X_t : t \in T\}$ is a nonnegative super-martingale with respect to $\mathfrak{F} = \{\mathcal{F}_t : t \in T\}$ then there exists a random variable X_∞ , measurable with respect to \mathcal{F}_∞ , such that $X_t \rightarrow X_\infty$ with probability 1.

Proof

Since \mathbf{X} is a nonnegative super-martingale, $\mathbb{E}(|X_t|) = \mathbb{E}(X_t) \leq \mathbb{E}(X_0)$ for $t \in T$. Hence the previous martingale convergence theorem applies.

Of course, the corollary applies to a nonnegative martingale as a special case. For the second *martingale convergence theorem* you will need to review uniformly integrable variables. Recall also that for $k \in [1, \infty)$, the k -norm of a random variable X is

$$\|X\|_k = \left[\mathbb{E}(|X|^k) \right]^{1/k} \quad (17.5.4)$$

and \mathcal{L}_k is the normed vector space of all real-valued random variables for which this norm is finite. *Convergence in mean* refers to convergence in \mathcal{L}_1 and more generally, *convergence in k th mean* refers to convergence in \mathcal{L}_k .

Suppose that \mathbf{X} is a uniformly integrable and is a sub-martingale or super-martingale with respect to \mathfrak{F} . Then there exists a random variable X_∞ , measurable with respect to \mathcal{F}_∞ such that $X_t \rightarrow X_\infty$ as $t \rightarrow \infty$ with probability 1 and in mean. Moreover, if \mathbf{X} is a martingale with respect to \mathfrak{F} then $X_t = \mathbb{E}(X_\infty | \mathcal{F}_t)$ for $t \in T$.

Proof

Since $\mathbf{X} = \{X_t : t \in T\}$ is uniformly integrable, $\mathbb{E}(|X_t|)$ is bounded in $t \in T$. Hence the by the first martingale convergence theorem, there exists X_∞ that is measurable with respect to \mathcal{F}_∞ such that $\mathbb{E}(|X_\infty|) < \infty$ and $X_t \rightarrow X_\infty$ as $t \rightarrow \infty$ with probability 1. By the uniform integrability theorem, the convergence is also in mean, so that $\mathbb{E}(|X_t - X|) \rightarrow 0$ as $t \rightarrow \infty$. Suppose now that \mathbf{X} is a martingale with respect to \mathfrak{F} . For fixed $s \in T$ we know that $\mathbb{E}(X_t | \mathcal{F}_s) \rightarrow \mathbb{E}(X_\infty | \mathcal{F}_s)$ as $t \rightarrow \infty$ (with probability 1). But $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$ for $t \geq s$ so it follows that $X_s = \mathbb{E}(X_\infty | \mathcal{F}_s)$.

As a simple corollary, recall that if $\|X_t\|_k$ is bounded in $t \in T$ for some $k \in (1, \infty)$ then \mathbf{X} is uniformly integrable, and hence the second martingale convergence theorem applies. But we can do better.

Suppose again that $\mathbf{X} = \{X_t : t \in T\}$ is a sub-martingale or super-martingale with respect to $\mathfrak{F} = \{\mathcal{F}_t : t \in T\}$ and that $\|X_t\|_k$ is bounded in $t \in T$ for some $k \in (1, \infty)$. Then there exists a random variable $X_\infty \in \mathcal{L}_k$ such that $X_t \rightarrow X_\infty$ as $t \rightarrow \infty$ in \mathcal{L}_k .

Proof

Suppose that $\|X_t\|_k \leq c$ for $t \in T$ where $c \in (0, \infty)$. Since $\|X\|_1 \leq \|X\|_k$, we have $\mathbb{E}(|X_t|)$ bounded in $t \in T$ so the first martingale convergence theorem applies. Hence there exists X_∞ , measurable with respect to \mathcal{F}_∞ , such that $X_t \rightarrow X_\infty$ as $t \rightarrow \infty$ with probability 1. Equivalently, with probability 1,

$$|X_t - X_\infty|^k \rightarrow 0 \text{ as } t \rightarrow \infty \quad (17.5.5)$$

Next, for $t \in T$, let $T_t = \{s \in T : s \leq t\}$ define $W_t = \sup\{|X_s| : s \in T_t\}$. by the norm version of the maximal inequality,

$$\|W_t\|_k \leq \frac{k}{k-1} \|X_t\| \leq \frac{kc}{k-1}, \quad t \in T \quad (17.5.6)$$

If we let $W_\infty = \sup\{|X_s| : s \in T\}$, then by the montone convergence theorem

$$\|W_\infty\|_k = \lim_{t \rightarrow \infty} \|W_t\|_k \leq \frac{ck}{k-1} \quad (17.5.7)$$

So $W_\infty \in \mathcal{L}_k$. But $|X_\infty| \leq W_\infty$ so $X_\infty \in \mathcal{L}_k$ also. Moreover, $|X_t - X_\infty|^k \leq 2^k W_\infty^k$, so applying the dominated convergence theorem to the first displayed equation above, we have $\mathbb{E}(|X_t - X_\infty|^k) \rightarrow 0$ as $t \rightarrow \infty$.

Example and Applications

In this subsection, we consider a number of applications of the martingale convergence theorems. One indication of the importance of martingale theory is the fact that many of the classical theorems of probability have simple and elegant proofs when formulated in terms of martingales.

Simple Random Walk

Suppose now that $\mathbf{V} = \{V_n : n \in \mathbb{N}\}$ is a sequence of independent random variables with $\mathbb{P}(V_i = 1) = p$ and $\mathbb{P}(V_i = -1) = 1 - p$ for $i \in \mathbb{N}_+$, where $p \in (0, 1)$. Let $\mathbf{X} = \{X_n : n \in \mathbb{N}\}$ be the partial sum process associated with \mathbf{V} so that

$$X_n = \sum_{i=0}^n V_i, \quad n \in \mathbb{N} \quad (17.5.8)$$

Recall that \mathbf{X} is the simple random walk with parameter p . From our study of Markov chains, we know that $p > \frac{1}{2}$ then $X_n \rightarrow \infty$ as $n \rightarrow \infty$ and if $p < \frac{1}{2}$ then $X_n \rightarrow -\infty$ as $n \rightarrow \infty$. The chain is transient in these two cases. If $p = \frac{1}{2}$, the chain is (null) recurrent and so visits every state in \mathbb{N} infinitely often. In this case X_n does not converge as $n \rightarrow \infty$. But of course $\mathbb{E}(X_n) = n(2p - 1)$ for $n \in \mathbb{N}$, so the martingale convergence theorems do not apply.

Doob's Martingale

Recall that if X is a random variable with $\mathbb{E}(|X|) < \infty$ and we define $X_t = \mathbb{E}(X | \mathcal{F}_t)$ for $t \in T$, then $\mathbf{X} = \{X_t : t \in T\}$ is a martingale relative to \mathfrak{F} and is known as a Doob martingale, named for you know whom. So the second martingale convergence theorem states that every uniformly integrable martingale is a Doob martingale. Moreover, we know that the Doob martingale \mathbf{X} constructed from X and \mathfrak{F} is uniformly integrable, so the second martingale convergence theorem applies. The last remaining question is the relationship between X and the limiting random variable X_∞ . The answer may come as no surprise.

Let $\mathbf{X} = \{X_t : t \in T\}$ be the Doob martingale constructed from X and \mathfrak{F} . Then $X_t \rightarrow X_\infty$ as $t \rightarrow \infty$ with probability 1 and in mean, where

$$X_\infty = \mathbb{E}(X | \mathcal{F}_\infty) \quad (17.5.9)$$

Of course if $\mathcal{F}_\infty = \mathcal{F}$, which is quite possible, then $X_\infty = X$. At the other extreme, if $\mathcal{F}_t = \{\emptyset, \Omega\}$, the trivial σ -algebra for all $t \in T$, then $X_\infty = \mathbb{E}(X)$, a constant.

Kolmogorov Zero-One Law

Suppose that $\mathbf{X} = (X_n : n \in \mathbb{N}_+)$ is a sequence of random variables with values in a general state space (S, \mathcal{S}) . Let $\mathcal{G}_n = \sigma\{X_k : k \geq n\}$ for $n \in \mathbb{N}_+$, and let $\mathcal{G}_\infty = \bigcap_{n=1}^\infty \mathcal{G}_n$. So \mathcal{G}_∞ is the *tail σ -algebra* of \mathbf{X} , the collection of events that depend only on the terms of the sequence with arbitrarily large indices. For example, if the sequence is real-valued (or more generally takes values in a metric space), then the event that X_n has a limit as $n \rightarrow \infty$ is a tail event. If $B \in \mathcal{S}$, then the event that $X_n \in B$ for infinitely many $n \in \mathbb{N}_+$ is another tail event. The *Kolmogorov zero-one law*, named for Andrei Kolmogorov, states that if \mathbf{X} is an independent sequence, then the tail events are essentially deterministic.

Suppose that \mathbf{X} is a sequence of independent random variables. If $A \in \mathcal{G}_\infty$ then $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.

Proof

Let $\mathcal{F}_n = \sigma\{X_k : k \leq n\}$ for $n \in \mathbb{N}_+$ so that $\mathfrak{F} = \{\mathcal{F}_n : n \in \mathbb{N}_+\}$ is the natural filtration associated with \mathbf{X} . As with our notation above, let $\mathcal{F}_\infty = \sigma\left(\bigcup_{n \in \mathbb{N}_+} \mathcal{F}_n\right)$. Now let $A \in \mathcal{G}_\infty$ be a tail event. Then $\{\mathbb{E}(\mathbf{1}_A | \mathcal{F}_n) : n \in \mathbb{N}_+\}$ is the Doob martingale associated with the indicator variable $\mathbf{1}_A$ and \mathfrak{F} . By our results above, $\mathbb{E}(\mathbf{1}_A | \mathcal{F}_n) \rightarrow \mathbb{E}(\mathbf{1}_A | \mathcal{F}_\infty)$ as $n \rightarrow \infty$ with probability 1. But $A \in \mathcal{F}_\infty$ so $\mathbb{E}(\mathbf{1}_A | \mathcal{F}_\infty) = \mathbf{1}_A$. On the other hand, $A \in \mathcal{G}_{n+1}$ and the σ -algebras \mathcal{G}_{n+1} and \mathcal{F}_n are independent. Therefore $\mathbb{E}(\mathbf{1}_A | \mathcal{F}_n) = \mathbb{P}(A)$ for each $n \in \mathbb{N}_+$. Thus $\mathbb{P}(A) = \mathbf{1}_A$.

Tail events and the Kolmogorov zero-one law were studied earlier in the section on measure in the chapter on probability spaces. A random variable that is measurable with respect to \mathcal{G}_∞ is a *tail random variable*. From the Kolmogorov zero-one law, a real-valued tail random variable for an independent sequence must be a constant (with probability 1).

Branching Processes

Recall the discussion of the simple *branching process* from the Introduction. The fundamental assumption is that the particles act independently, each with the same offspring distribution on \mathbb{N} . As before, we will let f denote the (discrete) probability density function of the number of offspring of a particle, m the mean of the distribution, and q the probability of extinction starting with a single particle. We assume that $f(0) > 0$ and $f(0) + f(1) < 1$ so that a particle has a positive probability of dying without children and a positive probability of producing more than 1 child.

The stochastic process of interest is $\mathbf{X} = \{X_n : n \in \mathbb{N}\}$ where X_n is the number of particles in the n th generation for $n \in \mathbb{N}$. Recall that \mathbf{X} is a discrete-time Markov chain on \mathbb{N} . Since 0 is an absorbing state, and all positive states lead to 0, we know that the positive states are transient and so are visited only finitely often with probability 1. It follows that either $X_n \rightarrow 0$ as $n \rightarrow \infty$ (extinction) or $X_n \rightarrow \infty$ as $n \rightarrow \infty$ (explosion). We have quite a bit of information about which of these events will occur from our study of Markov chains, but the martingale convergence theorems give more information.

Extinction and explosion

1. If $m \leq 1$ then $q = 1$ and extinction is certain.
2. If $m > 1$ then $q \in (0, 1)$. Either $X_n \rightarrow 0$ as $n \rightarrow \infty$ or $X_n \rightarrow \infty$ as $n \rightarrow \infty$ at an exponential rate.

Proof

The new information is the rate of divergence to ∞ in (b). The other statements are from our study of discrete-time branching Markov chains. We showed in the Introduction that $\{X_n/m^n : n \in \mathbb{N}\}$ is a martingale. Since this martingale is nonnegative, it has a limit as $n \rightarrow \infty$, and the limiting random variable takes values in \mathbb{R} . So if $m > 1$ and $X_n \rightarrow \infty$ as $n \rightarrow \infty$, then the divergence to ∞ must be at essentially the same rate as m^n .

The Beta-Bernoulli Process

Recall that the beta-Bernoulli process is constructed by randomizing the success parameter in a Bernoulli trials process with a beta distribution. Specifically, we start with a random variable P having the beta distribution with parameters $a, b \in (0, \infty)$. Next we have a sequence $\mathbf{X} = (X_1, X_2, \dots)$ of indicator variables with the property that \mathbf{X} is conditionally independent given $P = p \in (0, 1)$ with $\mathbb{P}(X_i = 1 \mid P = p) = p$ for $i \in \mathbb{N}_+$. Let $\mathbf{Y} = \{Y_n : n \in \mathbb{N}\}$ denote the partial sum process associated with \mathbf{X} , so that once again, $Y_n = \sum_{i=1}^n X_i$ for $n \in \mathbb{N}$. Next let $M_n = Y_n/n$ for $n \in \mathbb{N}_+$ so that M_n is the *sample mean* of (X_1, X_2, \dots, X_n) . Finally let

$$Z_n = \frac{a + Y_n}{a + b + n}, \quad n \in \mathbb{N} \quad (17.5.10)$$

We showed in the Introduction that $\mathbf{Z} = \{Z_n : n \in \mathbb{N}\}$ is a martingale with respect to \mathbf{X} .

$M_n \rightarrow P$ and $Z_n \rightarrow P$ as $n \rightarrow \infty$ with probability 1 and in mean.

Proof

We showed in the section on the beta-Bernoulli process that $Z_n \rightarrow P$ as $n \rightarrow \infty$ with probability 1. Note that $0 \leq Z_n \leq 1$ for $n \in \mathbb{N}$, so the martingale \mathbf{Z} is uniformly integrable. Hence the [second martingale convergence theorem](#) applies, and the convergence is in mean also.

This is a very nice result and is reminiscent of the fact that for the ordinary Bernoulli trials sequence with success parameter $p \in (0, 1)$ we have the law of large numbers that $M_n \rightarrow p$ as $n \rightarrow \infty$ with probability 1 and in mean.

Pólya's Urn Process

Recall that in the simplest version of Pólya's urn process, we start with an urn containing a red and b green balls. At each discrete time step, we select a ball at random from the urn and then replace the ball and add c new balls of the same color to the urn. For the parameters, we need $a, b \in \mathbb{N}_+$ and $c \in \mathbb{N}$. For $i \in \mathbb{N}_+$, let X_i denote the color of the ball selected on the i th draw, where 1 means red and 0 means green. For $n \in \mathbb{N}$, let $Y_n = \sum_{i=1}^n X_i$, so that $\mathbf{Y} = \{Y_n : n \in \mathbb{N}\}$ is the partial sum process associated with $\mathbf{X} = \{X_i : i \in \mathbb{N}_+\}$. Since Y_n is the number of red balls in the urn at time $n \in \mathbb{N}_+$, the *average* number of balls at time n is $M_n = Y_n/n$. On the other hand, the total number of balls in the urn at time $n \in \mathbb{N}$ is $a + b + cn$ so the *proportion* of red balls in the urn at time n is

$$Z_n = \frac{a + cY_n}{a + b + cn} \quad (17.5.11)$$

We showed in the Introduction, that $\mathbf{Z} = \{Z_n : n \in \mathbb{N}\}$ is a martingale. Now we are interested in the limiting behavior of M_n and Z_n as $n \rightarrow \infty$. When $c = 0$, the answer is easy. In this case, Y_n has the binomial distribution with trial parameter n and success parameter $a/(a+b)$, so by the law of large numbers, $M_n \rightarrow a/(a+b)$ as $n \rightarrow \infty$ with probability 1 and in mean. On the other hand, $Z_n = a/(a+b)$ when $c = 0$. So the interesting case is when $c > 0$.

Suppose that $c \in \mathbb{N}_+$. Then there exists a random variable P such that $M_n \rightarrow P$ and $Z_n \rightarrow P$ as $n \rightarrow \infty$ with probability 1 and in mean. Moreover, P has the beta distribution with left parameter a/c and right parameter b/c .

Proof

In our study of Pólya's urn process we showed that when $c \in \mathbb{N}_+$ the process \mathbf{X} is a beta-Bernoulli process with parameters a/c and b/c . So the result follows from our previous theorem.

Likelihood Ratio Tests

Recall the discussion of likelihood ratio tests in the Introduction. To review, suppose that (S, \mathcal{S}, μ) is a general measure space, and that $\mathbf{X} = \{X_n : n \in \mathbb{N}\}$ is a sequence of independent, identically distributed random variables, taking values in S , and having a common probability density function with respect to μ . The likelihood ratio test is a hypothesis test, where the null and alternative hypotheses are

- H_0 : the probability density function is g_0 .
- H_1 : the probability density function is g_1 .

We assume that g_0 and g_1 are positive on S . Also, it makes no sense for g_0 and g_1 to be the same, so we assume that $g_0 \neq g_1$ on a set of positive measure. The test is based on the *likelihood ratio test statistic*

$$L_n = \prod_{i=1}^n \frac{g_0(X_i)}{g_1(X_i)}, \quad n \in \mathbb{N} \quad (17.5.12)$$

We showed that under the alternative hypothesis H_1 , $\mathbf{L} = \{L_n : n \in \mathbb{N}\}$ is a martingale with respect to \mathbf{X} , known as the *likelihood ratio martingale*.

Under H_1 , $L_n \rightarrow 0$ as $n \rightarrow \infty$ with probability 1.

Proof

Assume that H_1 is true. \mathbf{L} is a nonnegative martingale, so the first martingale convergence theorem applies, and hence there exists a random variable L_∞ with values in $[0, \infty)$ such that $L_n \rightarrow L_\infty$ as $n \rightarrow \infty$ with probability 1. Next note that

$$\ln(L_n) = \sum_{i=1}^n \ln \left[\frac{g_0(X_i)}{g_1(X_i)} \right] \quad (17.5.13)$$

The variables $\ln[g_0(X_i)/g_1(X_i)]$ for $i \in \mathbb{N}_+$ are also independent and identically distributed, so let m denote the common mean. The natural logarithm is concave and the martingale \mathbf{L} has mean 1, so by Jensen's inequality,

$$m = \mathbb{E} \left(\ln \left[\frac{g_0(X)}{g_1(X)} \right] \right) < \ln \left(\mathbb{E} \left[\frac{g_0(X)}{g_1(X)} \right] \right) = \ln(1) = 0 \quad (17.5.14)$$

Hence $m \in [-\infty, 0)$. By the strong law of large numbers, $\frac{1}{n} \ln(L_n) \rightarrow m$ as $n \rightarrow \infty$ with probability 1. Hence we must have $\ln(L_n) \rightarrow -\infty$ as $n \rightarrow \infty$ with probability 1. But by continuity, $\ln(L_n) \rightarrow \ln(L_\infty)$ as $n \rightarrow \infty$ with probability 1, so $L_\infty = 0$ with probability 1.

This result is good news, statistically speaking. Small values of L_n are evidence in favor of H_1 , so the decision rule is to reject H_0 in favor of H_1 if $L_n \leq l$ for a chosen critical value $l \in (0, \infty)$. If H_1 is true and the sample size n is sufficiently large, we will reject H_0 . In the proof, note that $\ln(L_n)$ must diverge to $-\infty$ at least as fast as n diverges to ∞ . Hence $L_n \rightarrow 0$ as $n \rightarrow \infty$ exponentially fast, at least. It also worth noting that \mathbf{L} is a mean 1 martingale (under H_1) so trivially $\mathbb{E}(L_n) \rightarrow 1$ as $n \rightarrow \infty$ even though $L_n \rightarrow 0$ as $n \rightarrow \infty$ with probability 1. So the likelihood ratio martingale is a good example of a sequence where the interchange of limit and expected value is not valid.

Partial Products

Suppose that $\mathbf{X} = \{X_n : n \in \mathbb{N}_+\}$ is an independent sequence of nonnegative random variables with $\mathbb{E}(X_n) = 1$ for $n \in \mathbb{N}_+$. Let

$$Y_n = \prod_{i=1}^n X_i, \quad n \in \mathbb{N} \quad (17.5.15)$$

so that $\mathbf{Y} = \{Y_n : n \in \mathbb{N}\}$ is the partial product process associated with \mathbf{X} . From our discussion of this process in the Introduction, we know that \mathbf{Y} is a martingale with respect to \mathbf{X} . Since \mathbf{Y} is nonnegative, the [second martingale convergence theorem](#) applies, so there exists a random variable Y_∞ such that $Y_n \rightarrow Y_\infty$ as $n \rightarrow \infty$ with probability 1. What more can we say? The following result, known as the *Kakutani product martingale theorem*, is due to Shizuo Kakutani.

Let $a_n = \mathbb{E}(\sqrt{X_n})$ for $n \in \mathbb{N}_+$ and let $A = \prod_{i=1}^{\infty} a_i$.

1. If $A > 0$ then $Y_n \rightarrow Y_\infty$ as $n \rightarrow \infty$ in mean and $\mathbb{E}(Y_\infty) = 1$.
2. If $A = 0$ then $Y_\infty = 0$ with probability 1.

Proof

Note that $a_n > 0$ for $n \in \mathbb{N}_+$ since X_n is nonnegative and $\mathbb{P}(X_n > 0) > 0$. Also, since $x \mapsto \sqrt{x}$ is concave on $(0, \infty)$ it follows from Jensen's inequality that

$$a_n = \mathbb{E}(\sqrt{X_n}) \leq \sqrt{\mathbb{E}(X_n)} = 1 \quad (17.5.16)$$

Let $A_n = \prod_{i=1}^n a_i$ for $n \in \mathbb{N}$. Since $a_n \in (0, 1]$ for $n \in \mathbb{N}_+$, it follows that $A_n \in (0, 1]$ for $n \in \mathbb{N}$ and that A_n is decreasing in $n \in \mathbb{N}$ with limit $A = \prod_{i=1}^{\infty} a_i \in [0, 1]$. Next let $Z_n = \prod_{i=1}^n \sqrt{X_i}/a_i$ for $n \in \mathbb{N}$, so that $\mathbf{Z} = \{Z_n : n \in \mathbb{N}\}$ is the partial product process associated with $\{\sqrt{X_n}/a_n : n \in \mathbb{N}\}$. Since $\mathbb{E}(\sqrt{X_n}/a_n) = 1$ for $n \in \mathbb{N}_+$, the process \mathbf{Z} is also a nonnegative martingale, so there exists a random variable Z_∞ such that $Z_n \rightarrow Z_\infty$ as $n \rightarrow \infty$ with probability 1. Note that $Z_n^2 = Y_n/A_n^2$, $Y_n = A_n^2 Z_n^2$, and $Y_n \leq Z_n^2$ for $n \in \mathbb{N}$.

1. Suppose that $A > 0$. Since the martingale \mathbf{Y} has mean 1,

$$\mathbb{E}(Z_n^2) = \mathbb{E}(Y_n/A_n^2) = 1/A_n^2 \leq 1/A^2 < \infty, \quad n \in \mathbb{N} \quad (17.5.17)$$

Let $W_n = \max\{Z_k : k \in \{0, 1, \dots, n\}\}$ for $n \in \mathbb{N}$ so that $\mathbf{W} = \{W_n : n \in \mathbb{N}\}$ is the maximal process associated with \mathbf{Z} . Also, let $W_\infty = \sup\{Z_k : k \in \mathbb{N}\}$ and note that $W_n \uparrow W_\infty$ as $n \rightarrow \infty$. By the \mathcal{L}_2 maximal inequality,

$$\mathbb{E}(W_n^2) \leq 4\mathbb{E}(Z_n^2) \leq 4/A^2, \quad n \in \mathbb{N} \quad (17.5.18)$$

By the monotone convergence theorem, $\mathbb{E}(W_\infty^2) = \lim_{n \rightarrow \infty} \mathbb{E}(W_n^2) \leq 4/A^2$. Since $x \rightarrow x^2$ is strictly increasing on $[0, \infty)$, $W_\infty^2 = \sup\{Z_n^2 : n \in \mathbb{N}\}$ and so $Y_n \leq W_\infty^2$ for $n \in \mathbb{N}$. Since $\mathbb{E}(W_\infty^2) < \infty$, it follows that the martingale \mathbf{Y} is uniformly integrable. Hence by the [third martingale convergence theorem](#) above, $Y_n \rightarrow Y_\infty$ is mean. Since convergence in mean implies that the means converge, $\mathbb{E}(Y_\infty) = \lim_{n \rightarrow \infty} \mathbb{E}(Y_n) = 1$.

2. Suppose that $A = 0$. Then $Y_n = A_n^2 Z_n^2 \rightarrow 0 \cdot Z_\infty^2 = 0$ as $n \rightarrow \infty$ with probability 1. Note that in this case, the convergence is not in mean, and trivially $\mathbb{E}(Y_\infty) = 0$.

Density Functions

This discussion continues the one on density functions in the Introduction. To review, we start with our probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\mathfrak{F} = \{\mathcal{F}_n : n \in \mathbb{N}\}$ in discrete time. Recall again that $\mathcal{F}_\infty = \sigma(\bigcup_{n=0}^{\infty} \mathcal{F}_n)$. Suppose now that μ is a finite measure on the sample space (Ω, \mathcal{F}) . For each $n \in \mathbb{N} \cup \{\infty\}$, the restriction of μ to \mathcal{F}_n is a measure on (Ω, \mathcal{F}_n) and similarly the restriction of \mathbb{P} to \mathcal{F}_n is a probability measure on (Ω, \mathcal{F}_n) . To save notation and terminology, we will refer to these as μ and \mathbb{P} on \mathcal{F}_n , respectively. Suppose now that μ is absolutely continuous with respect to \mathbb{P} on \mathcal{F}_n for each $n \in \mathbb{N}$. By the Radon-Nikodym theorem, μ has a density function (or Radon-Nikodym derivative) $X_n : \Omega \rightarrow \mathbb{R}$ with respect to \mathbb{P} on \mathcal{F}_n for each $n \in \mathbb{N}$. The theorem and the derivative are named for Johann Radon and Otto Nikodym. In the Introduction we showed that $\mathbf{X} = \{X_n : n \in \mathbb{N}\}$ is a martingale with respect to \mathfrak{F} . Here is the convergence result:

There exists a random variable X_∞ such that $X_n \rightarrow X_\infty$ as $n \rightarrow \infty$ with probability 1.

1. If μ is absolutely continuous with respect to \mathbb{P} on \mathcal{F}_∞ then X_∞ is a density function of μ with respect to \mathbb{P} on \mathcal{F}_∞ .
2. If μ and \mathbb{P} are mutually singular on \mathcal{F}_∞ then $X_\infty = 0$ with probability 1.

Proof

Again, as shown in the Introduction, \mathbf{X} is a martingale with respect to \mathfrak{F} . Moreover, $\mathbb{E}(|X_n|) = \|\mu\|$ (the total variation of μ) for each $n \in \mathbb{N}$. Since μ is a finite measure, $\|\mu\| < \infty$ so the first [martingale convergence theorem](#) applies. Hence there exists a random variable X_∞ , measurable with respect to \mathcal{F}_∞ , such that $X_n \rightarrow X_\infty$ as $n \rightarrow \infty$.

1. If μ is absolutely continuous with respect to \mathbb{P} on \mathcal{F}_∞ , then μ has a density function Y_∞ with respect to \mathbb{P} on \mathcal{F}_∞ . Our goal is to show that $X_\infty = Y_\infty$ with probability 1. By definition, Y_∞ is measurable with respect to \mathcal{F}_∞ and

$$\int_A Y_\infty d\mathbb{P} = \mathbb{E}(Y_\infty; A) = \mu(A), \quad A \in \mathcal{F}_\infty \quad (17.5.19)$$

Suppose now that $n \in \mathbb{N}$ and $A \in \mathcal{F}_n$. Then again by definition, $\mathbb{E}(X_n; A) = \mu(A)$. But $A \in \mathcal{F}_\infty$ also, so $\mathbb{E}(Y_\infty; A) = \mu(A)$. So to summarize, X_n is \mathcal{F}_n -measurable and $\mathbb{E}(X_n; A) = \mathbb{E}(Y_\infty; A)$ for each $A \in \mathcal{F}_n$. By definition, this means that $X_n = \mathbb{E}(Y_\infty | \mathcal{F}_n)$, so X is the [Doob martingale](#) associated with Y_∞ . Letting $n \rightarrow \infty$ and using the result [above](#) gives $X_\infty = \mathbb{E}(Y_\infty | \mathcal{F}_\infty) = Y_\infty$ (with probability 1, of course).

2. Suppose that μ and \mathbb{P} are mutually singular on \mathcal{F}_∞ . Assume first that μ is a positive measure, so that X_n is nonnegative for $n \in \mathbb{N} \cup \{\infty\}$. By the definition of mutual singularity, there exists $B \in \mathcal{F}_\infty$ such that $\mu_\infty(B) = 0$ and $\mathbb{P}_\infty(B^c) = 0$, so that $\mathbb{P}(B) = 1$. Our goal is to show that $\mathbb{E}(X_\infty; A) \leq \mu(A)$ for every $A \in \mathcal{F}_\infty$. Towards that end, let

$$\mathcal{M} = \{A \in \mathcal{F}_\infty : \mathbb{E}(X_\infty; A) \leq \mu(A)\} \quad (17.5.20)$$

Suppose that $A \in \bigcup_{k=0}^\infty \mathcal{F}_k$, so that $A \in \mathcal{F}_k$ for some $k \in \mathbb{N}$. Then $A \in \mathcal{F}_n$ for all $n \geq k$ and therefore $\mathbb{E}(X_n; A) = \mu(A)$ for all $n \geq k$. By Fatou's lemmas,

$$\mathbb{E}(X_\infty; A) \leq \liminf_{n \rightarrow \infty} \mathbb{E}(X_n; A) \leq \mu(A) \quad (17.5.21)$$

so $A \in \mathcal{M}$. Next, suppose that $\{A_n : n \in \mathbb{N}\}$ is an increasing or decreasing sequence in \mathcal{M} , and let $A_\infty = \lim_{n \rightarrow \infty} A_n$ (the union in the first case and the intersection in the second case). Then $\mathbb{E}(X_\infty; A_n) \leq \mu(A_n)$ for each $n \in \mathbb{N}$. By the continuity theorems, $\mathbb{E}(X_\infty; A_n) \rightarrow \mathbb{E}(X_\infty; A_\infty)$ and $\mu(A_n) \rightarrow \mu(A_\infty)$ as $n \rightarrow \infty$. Therefore $\mathbb{E}(X_\infty; A_\infty) \leq \mu(A_\infty)$ and so $A_\infty \in \mathcal{M}$. It follows that \mathcal{M} is a monotone class. Since \mathcal{M} contains the algebra $\bigcup_{n=0}^\infty \mathcal{F}_n$, it then follows from the monotone class theorem that $\mathcal{F}_\infty \subseteq \mathcal{M}$. In particular $B \in \mathcal{M}$, so $\mathbb{E}(X_\infty) = \mathbb{E}(X_\infty; B) \leq \mu(B) = 0$ and therefore $X_\infty = 0$ with probability 1. If μ is a general finite measure, then by the Jordan decomposition theorem, μ can be written uniquely in the form $\mu = \mu^+ - \mu^-$ where μ^+ and μ^- are finite positive measures. Moreover, X_n^+ is the density function of μ^+ on \mathcal{F}_n and X_n^- is the density function of μ^- on \mathcal{F}_n . By the first part of the proof, $X^+ = 0$, $X^- = 0$, and also $X = 0$, all with probability 1.

The martingale approach can be used to give a probabilistic proof of the Radon-Nikodym theorem, at least in certain cases. We start with a sample set Ω . Suppose that $\mathcal{A}_n = \{A_i^n : i \in I_n\}$ is a countable partition of Ω for each $n \in \mathbb{N}$. Thus I_n is countable, $A_i^n \cap A_j^n = \emptyset$ for distinct $i, j \in I_n$, and $\bigcup_{i \in I_n} A_i^n = \Omega$. Suppose also that \mathcal{A}_{n+1} refines \mathcal{A}_n for each $n \in \mathbb{N}$ in the sense that A_i^n is a union of sets in \mathcal{A}_{n+1} for each $i \in I_n$. Let $\mathcal{F}_n = \sigma(\mathcal{A}_n)$. Thus \mathcal{F}_n is generated by a countable partition, and so the sets in \mathcal{F}_n are of the form $\bigcup_{j \in J} A_j^n$ where $J \subseteq I_n$. Moreover, by the refinement property $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for $n \in \mathbb{N}$, so that $\mathcal{F} = \{\mathcal{F}_n : n \in \mathbb{N}\}$ is a filtration. Let $\mathcal{F} = \mathcal{F}_\infty = \sigma(\bigcup_{n=0}^\infty \mathcal{F}_n) = \sigma(\bigcup_{n=0}^\infty \mathcal{A}_n)$, so that our sample space is (Ω, \mathcal{F}) . Finally, suppose that \mathbb{P} is a probability measure on (Ω, \mathcal{F}) with the property that $\mathbb{P}(A_i^n) > 0$ for $n \in \mathbb{N}$ and $i \in I_n$. We now have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Interesting probability spaces that occur in applications are of this form, so the setting is not as specialized as you might think.

Suppose now that μ a finite measure on (Ω, \mathcal{F}) . From our assumptions, the only null set for \mathbb{P} on \mathcal{F}_n is \emptyset , so μ is automatically absolutely continuous with respect to \mathbb{P} on \mathcal{F}_n . Moreover, for $n \in \mathbb{N}$, we can give the density function of μ with respect to \mathbb{P} on \mathcal{F}_n explicitly:

The density function of μ with respect to \mathbb{P} on \mathcal{F}_n is the random variable X_n whose value on A_i^n is $\mu(A_i^n)/\mathbb{P}(A_i^n)$ for each $i \in I_n$. Equivalently,

$$X_n = \sum_{i \in I_n} \frac{\mu(A_i^n)}{\mathbb{P}(A_i^n)} \mathbf{1}(A_i^n) \quad (17.5.22)$$

Proof

We need to show that $\mu(A) = \mathbb{E}(X_n; A)$ for each $A \in \mathcal{F}_n$. So suppose $A = \bigcup_{j \in J} A_j^n$ where $J \subseteq I_n$. Then

$$\mathbb{E}(X_n; A) = \sum_{j \in J} \mathbb{E}(X_n; A_j^n) = \sum_{j \in J} \frac{\mu(A_j^n)}{\mathbb{P}(A_j^n)} \mathbb{P}(A_j^n) = \sum_{j \in J} \mu(A_j^n) = \mu(A) \quad (17.5.23)$$

By our theorem above, there exists a random variable X such that $X_n \rightarrow X$ as $n \rightarrow \infty$ with probability 1. If μ is absolutely continuous with respect to \mathbb{P} on \mathcal{F} , then X is a density function of μ with respect to \mathbb{P} on \mathcal{F} . The point is that we have given a more or less explicit construction of the density.

For a concrete example, consider $\Omega = [0, 1)$. For $n \in \mathbb{N}$, let

$$\mathcal{A}_n = \left\{ \left[\frac{j}{2^n}, \frac{j+1}{2^n} \right) : j \in \{0, 1, \dots, 2^n - 1\} \right\} \quad (17.5.24)$$

This is the partition of $[0, 1)$ into 2^n subintervals of equal length $1/2^n$, based on the *dyadic rationals* (or *binary rationals*) of rank n or less. Note that every interval in \mathcal{A}_n is the union of two adjacent intervals in \mathcal{A}_{n+1} , so the refinement property holds. Let \mathbb{P} be ordinary Lebesgue measure on $[0, 1)$ so that $\mathbb{P}(A_i^n) = 1/2^n$ for $n \in \mathbb{N}$ and $i \in \{0, 1, \dots, 2^n - 1\}$. As above, let $\mathcal{F}_n = \sigma(\mathcal{A}_n)$ and $\mathcal{F} = \sigma(\bigcup_{n=0}^{\infty} \mathcal{F}_n) = \sigma(\bigcup_{n=0}^{\infty} \mathcal{A}_n)$. The dyadic rationals are dense in $[0, 1)$, so \mathcal{F} is the ordinary Borel σ -algebra on $[0, 1)$. Thus our probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is simply $[0, 1)$ with the usual Euclidean structures. If μ is a finite measure on $([0, 1), \mathcal{F})$ then the density function of μ on \mathcal{F}_n is the random variable X_n whose value on the interval $[j/2^n, (j+1)/2^n)$ is $2^n \mu[j/2^n, (j+1)/2^n)$. If μ is absolutely continuous with respect to \mathbb{P} on \mathcal{F} (so absolutely continuous in the usual sense), then a density function of μ is $X = \lim_{n \rightarrow \infty} X_n$.

This page titled [17.5: Convergence](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist](#) ([Random Services](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.