Economics Letters 151 (2017) 75–78

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Ergodic for the mean

Uwe Hassler

Goethe University Frankfurt, Campus Westend, RuW Building, 60629 Frankfurt, Germany

HIGHLIGHTS

- Stationary time series.
- Consistent in mean square.
- Sucient and necessary condition.
- Asymptotic average uncorrelatedness.

ARTICLE INFO

Article history:
Received 2 November 2016
Received in revised form 6 December 2016
Accepted 8 December 2016
Available online 16 December 2016

JEL classification:
C22

Keywords:
Stationary time series
consistency in mean square

ABSTRACT

We discuss ergodicity for the mean in the sense that the sample average converges in mean square to the population mean of a stationary stochastic process. This differs from ergodicity in a measure theoretic sense. It is widely known that asymptotic uncorrelatedness is sufficient for ergodicity for the mean. We weaken this assumption to "asymptotic average uncorrelatedness" and show that it cannot be further weakened: Our condition is necessary and sufficient.

1. Introduction and summary

A time series of length T is considered as one realization of a stochastic process. We assume that the process is stationary with constant expectation μ at all time. This population mean μ is estimated by the time average over T variables $x_1, \dots, x_T$:

$$\bar{x} = \frac{1}{T} \sum_{t=1}^{T} x_t .$$

Under what circumstances will $\bar{x}$ converge to μ as $T \to \infty$? The answer depends on the mode of convergence. If convergence is almost surely, then a sufficient condition is strict stationarity and ergodicity in a measure theoretic sense (see Definition 1 and Proposition 1 below). If convergence is in mean square we speak of ergodicity for the mean. Almost sure convergence is neither implied by nor does it imply convergence in mean square, such that ergodicity in a measure theoretic sense and ergodicity for the mean are somehow separate concepts. Many authors prefer the rather simple and intuitive concept of ergodicity for the mean, since the ergodicity in a measure theoretic sense is mathematically much more involved, and strict stationarity is an assumption that is hard to test empirically. Brockwell and Davis (1991, Thm. 7.1.1) and Fuller (1996, Cor. 6.1.1.1) e.g. establish a sufficient condition for ergodicity for the mean under covariance stationarity. In this letter we provide a weaker condition and show that it is necessary, too, see Proposition 2 below. We hence find that ergodicity for the mean is equivalent to "asymptotic average uncorrelatedness" in the sense of Proposition 2. Two examples illustrate what ergodicity for the mean is and what it is not.

The rest of this letter is organized as follows. The next section becomes precise on the definitions and notation. Section 3 gives our equivalent characterization of ergodicity for the mean and discusses two illustrative examples. The proof of our result is presented in the final section.

2. Preliminaries

Let the index set $\mathbb{T}$ be a subset of the integers, $\mathbb{T} \subseteq \mathbb{Z}$, and $\{x_t\}_{t\in\mathbb{T}}$ denotes a univariate stochastic, real-valued process. The process is called covariance stationary if the first and second moments are constant over time (where we assume these moments to exist), and if the covariance between x_t and x_{t+h} depends on the distance h only:

E-mail address: hassler@wiwi.uni-frankfurt.de.

http://dx.doi.org/10.1016/j.econlet.2016.12.013

1. $\mathrm{E}(x_t) = \mu$ for $t \in \mathbb{T}$,
2. $\mathrm{Cov}(x_t, x_{t+h}) = \mathrm{E}[(x_t - \mu)(x_{t+h} - \mu)] = \gamma(h)$ for all $t, t + h \in \mathbb{T}$.

The process is said to be strictly stationary if the distribution of any n-vector $\left(x_{t_1}, \ldots, x_{t_n}\right)'$ is invariant for any $t_1 < t_2 < \cdots < t_n$, such that a shift from $t_1 < t_2 < \cdots < t_n$ to $t_1 + h < t_2 + h < \cdots < t_n + h$ does not change the joint distribution. Examples of covariance and strict stationarity are, respectively, white noise processes and pure random processes defined as follows. Assume a sequence $\{\varepsilon_t\}$ free of serial correlation:

$$\mathrm{Cov}(\varepsilon_t, \varepsilon_{t+h}) = 0, \quad h \neq 0 \quad \text{with } \mathrm{E}(\varepsilon_t) = 0. \tag{1}$$

The process is called white noise in case of constant variance, $\mathrm{E}(\varepsilon_t^2) = \gamma(0) = \sigma^2$, often abbreviated as $\{\varepsilon_t\} \sim \mathrm{WN}(0, \sigma^2)$. If a white noise process is independent over time and has an identical distribution (iid) at each point of time, then it is called a pure random (or iid) process, in short $\{\varepsilon_t\} \sim \mathrm{iid}(0, \sigma^2)$.

We are concerned with consistent estimation of the population mean from a sample of size T in that $\bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t$ converges to μ. Such results have been called "law of large numbers" (LLN). Here, we distinguish between almost sure convergence $\overset{a.s.}{\to}$ and convergence in mean square $\overset{2}{\to}$ as $T \to \infty$, see for instance Pötscher and Prucha (2001) for a concise review of modes of convergence. The latter convergence is defined by

$$\mathrm{E}\left((\bar{x} - \mu)^2\right) \to 0 \quad \text{as } T \to \infty. \tag{2}$$

Consistent estimation of μ is related to the concept of "ergodicity", which is mathematically involved and requires measure theoretic foundations. We refrain from the precise definition given e.g. in Stout (1974), Breiman (1992) or Davidson (1994); note that the earlier exposition by Doob (1953) calls an ergodic process "metrically transitive".

Definition 1 (*Ergodicity*)**.** A strictly stationary process is called ergodic when satisfying metric transitivity as defined by Doob (1953, p. 457); see also Stout (1974, Def. 3.5.8) or Breiman (1992, Def. 6.30).

The relevance of ergodicity is rooted in the following well known result.

Proposition 1 (*Ergodicity*)**.** *Let $\{x_t\}$ be a strictly stationary, ergodic process with $E(|x_t|) < \infty$. Then $\bar{x} \overset{a.s.}{\to} \mu$ as $T \to \infty$.*

Proof. Doob (1953, Thm. 2.1, p. 465), Stout (1974, Thm. 3.5.8) or Breiman (1992, Prop. 6.31). ■

Doob (1953, Thm. 1.2, p. 460) or Breiman (1992, Cor. 6.33) show that every iid process is ergodic. Obviously, if $\{\varepsilon_t\}$ is iid, then so is $\{\varepsilon_t^2\}$. From this we conclude the following implication:

$$\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \overset{a.s.}{\to} 0 \text{ and } \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t^2 \overset{a.s.}{\to} \sigma^2 \quad \text{if } \varepsilon_t \sim \mathrm{iid}(0, \sigma^2). \tag{3}$$

A drawback of Proposition 1 is the assumption of strict stationarity, which is hard to verify in practice. Further, it has been mentioned that the concept behind Definition 1 is mathematically somewhat involved. For these reasons many economists refer to mean square convergence when speaking of ergodicity, which implies so-called convergence in probability. According to Hamilton (1994, p. 47) or Fuller (1996, p. 308) such a property is called "ergodic for the mean".

Definition 2 (*Ergodic for The Mean*)**.** A covariance stationary process is called ergodic for the mean when the sample average converges to $\mu = \mathrm{E}(x_t)$ in mean square, see (2).

In the next section we establish a necessary and sufficient condition for ergodic for the mean.

3. Result and discussion

Hamilton (1994, Prop. 7.5) establishes ergodicity for the mean under the assumption of an absolutely summable autocovariance sequence. This is much more restrictive than necessary. Brockwell and Davis (1991) and Fuller (1996) work under the weaker sufficient condition that the sequence of autocovariances $\{\gamma(h)\}$ converges to zero as $h \to \infty$:

$$\gamma(h) \to 0 \quad \text{as } h \to \infty. \tag{4}$$

Brockwell and Davis (1991, Thm. 7.1.1) and Fuller (1996, Cor. 6.1.1.1) prove that (4) implies (2). An even weaker sufficient condition is given in Proposition 2. It only assumes that the average over the autocovariances converges to zero. This assumption cannot be weakened, since we show that it is also necessary.

Proposition 2 (*Ergodic for the Mean*)**.** *Let $\{x_t\}$ be a covariance stationary process with finite μ and $\{\gamma(h)\}$. The sample average converges in mean square to μ if and only if*

$$\frac{1}{H}\sum_{h=1}^{H} \gamma(h) \to 0 \tag{5}$$

as $H \to \infty$.

Proof. See next section.

Hence, a necessary and sufficient condition for ergodicity for the mean is asymptotic average uncorrelatedness in the sense of (5). A sequence $\{\gamma(h)\}$ meeting (5) is also called Cesàro summable with Cesàro sum equal to zero.

We discuss three examples to illustrate what ergodicity for the mean is, and what it is not. First, we show that ergodicity for the mean does not imply ergodicity in the sense of Definition 1.

Example 1. Let B be a Bernoulli variable with equal probability, i.e. $\mathrm{P}(B = 1) = \mathrm{P}(B = 0) = 1/2$. Further, assume that $\{\varepsilon_t\}$ is iid with mean zero and variance 1, and independent of B. Next, define $\{x_t\} = \{\varepsilon_t B\}$, which is a strictly stationary process. The claim is that $\{x_t\}$ is ergodic for the mean but not ergodic in the sense of Definition 1. To see this we first obtain the second moments:

$$\mathrm{E}(x_t) = 0, \ \mathrm{E}(x_t^2) = \mathrm{E}(B^2) = \frac{1}{2} \quad \text{and} \quad \mathrm{E}(x_t x_s) = 0 \text{ for } t \neq s.$$

Note that Proposition 2 and (3) imply that, respectively,

$$\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \overset{2}{\to} 0 \quad \text{and} \quad \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t^2 \overset{a.s.}{\to} 1.$$

Consequently, one has

$$\frac{1}{T}\sum_{t=1}^{T} x_t = B\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \overset{2}{\to} 0,$$

$$\frac{1}{T}\sum_{t=1}^{T} x_t^2 = B^2\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t^2 \overset{a.s.}{\to} B^2 \neq \frac{1}{2}.$$

While $\frac{1}{T}\sum_{t=1}^{T} x_t$ is consistent in mean square for $\mathrm{E}(x_t)$ (ergodic for the mean), the process is not ergodic. This follows from the fact that $\frac{1}{T}\sum_{t=1}^{T} x_t^2$ would converges almost surely to $\mathrm{E}(x_t^2)$ if $\{x_t\}$ was ergodic. The reason for that is: if $\{x_t\}$ was ergodic, then we know from Doob (1953, p. 458) and Breiman (1992, Prop. 6.31) that $\{x_t^2\}$ would be strictly stationary and ergodic too. Hence, if $\{x_t\}$ was ergodic, then Proposition 1 would imply $\frac{1}{T}\sum_{t=1}^{T} x_t^2 \overset{a.s.}{\to} \frac{1}{2}$. This proves the claim.