## Exercise: Show that

$$
y_{n} \xrightarrow{L} c \Rightarrow y_{n} \xrightarrow{p} c .
$$

Solution:

$$
\begin{aligned}
& y_{n} \xrightarrow{L} c \Rightarrow F_{y_{n}}(\lambda) \rightarrow F_{c}(\lambda)=I_{[c, \infty)}(\lambda) \forall \lambda \neq c \\
& \Rightarrow P\left(\left|y_{n}-c\right|<\varepsilon\right)=P\left(c-\varepsilon<y_{n}<c+\varepsilon\right) \\
& \geq P\left(c-\frac{\varepsilon}{2}<y_{n} \leq c+\frac{\varepsilon}{2}\right) \\
&=F_{y_{n}}\left(\mathrm{c}+\frac{\varepsilon}{2}\right)-F_{y_{n}}\left(c-\frac{\varepsilon}{2}\right) \\
& \rightarrow F_{c}\left(c+\frac{\varepsilon}{2}\right)-F_{c}\left(c-\frac{\varepsilon}{2}\right)=1-0
\end{aligned}
$$

## Remarks:

(i) $F_{c}(\lambda)$ is not continuous at $\lambda=c$.
(ii) In general, we have

$$
y_{n} \xrightarrow{p} y \Rightarrow y_{n} \xrightarrow{L} y .
$$

The last exercise therefore establishes the equivalence of convergence in probability and convergence in law (distribution) in the case of convergence to a constant.

The mean of a sample $x_{l}, \ldots, x_{n}$ from an $\mathrm{AR}(1)$ process

$$
\left(x_{t}-\mu\right)=\phi\left(x_{t-1}-\mu\right)+u_{t}
$$

with mean $\mu$, AR-parameter $\phi$ satisfying $|\phi|<1$, innovations $u_{t}$ with variance $\sigma^{2}>0$, and autocovariances

$$
\gamma(k)=E\left(x_{\mathrm{t}}-\mu\right)\left(x_{t+k}-\mu\right)=\frac{\phi^{|k|} \sigma^{2}}{1-\phi^{2}}
$$

is a consistent estimator of $\mu$.
We establish (weak) consistency ${ }^{1}$ by using the fact that mean square convergence implies convergence in probability.

We have

$$
E(\bar{x})=E \frac{1}{n} \sum_{t=1}^{n} x_{t}=\frac{1}{n} \sum_{t=1}^{n} E x_{t}=\mu
$$

[^0]and
\[

$$
\begin{aligned}
& 0 \leq \operatorname{Var}(\bar{x})= E\left(\frac{1}{n} \sum_{t=1}^{n} x_{t}-\mu\right)^{2}=E\left(\frac{1}{n} \sum_{t=1}^{n}\left(x_{t}-\mu\right)\right)^{2} \\
&= \frac{1}{n^{2}}\left(E\left(x_{1}-\mu\right)^{2}+\ldots+E\left(x_{n}-\mu\right)^{2}\right. \\
&+E\left(x_{1}-\mu\right)\left(x_{2}-\mu\right)+\ldots+E\left(x_{n-1}-\mu\right)\left(x_{n}-\mu\right) \\
& \vdots \\
&+E\left(x_{1}-\mu\right)\left(x_{n}-\mu\right) \\
&+E\left(x_{2}-\mu\right)\left(x_{1}-\mu\right)+\ldots+E\left(x_{n}-\mu\right)\left(x_{n-1}-\mu\right) \\
& \vdots \\
&\left.+E\left(x_{n}-\mu\right)\left(x_{1}-\mu\right)\right) \\
&= \frac{1}{n^{2}}[n \gamma(0)+2(n-1) \gamma(1)+\ldots+2 \gamma(n-1)] \\
&= \frac{2}{n^{2}} \sum_{k=0}^{n-1}(n-k) \gamma(k)-\frac{1}{n} \gamma(0) \\
& \leq \frac{2}{n} \sum_{k=0}^{n-1}|\gamma(k)|= \frac{2}{n} \sum_{k=0}^{n-1} \frac{\sigma^{2} \mid \phi \phi^{k}}{1-\phi^{2}}=\frac{2}{n} \frac{\sigma^{2}}{1-\phi^{2}} \frac{1-|\phi|^{n}}{1-\phi \mid} \rightarrow 0 .
\end{aligned}
$$
\]

Now assume that $\mu=0$ and the innovations $u_{t}$ are an i.i.d. sequence of $N\left(0, \sigma^{2}\right)$ random variables.

Exercise: Show that $E u_{j} u_{k} u_{l} u_{m} \leq E u_{j}^{4}=3 \sigma^{4}$.

$$
\text { Solution: } \begin{gathered}
\quad j \neq k, l, m: \quad E u_{j} u_{k} u_{l} u_{m}=E u_{j} E u_{k} u_{l} u_{m}=0 \\
\\
j=k \neq l=m: E u_{j} u_{k} u_{l} u_{m}=E u_{j}^{2} E u_{l}^{2}=\sigma^{2} \sigma^{2}=\sigma^{4} \\
j=k=l=m: \quad E u_{j} u_{k} u_{l} u_{m}=E u_{j}^{4}=3 \sigma^{4}
\end{gathered}
$$

Exercise: Show that $E x_{t}^{4}<\infty$.
Hint: Use the $\mathrm{MA}(\infty)$ representation

$$
x_{t}=\sum_{j=0}^{\infty} \phi^{j} u_{t-j}
$$

of the $\mathrm{AR}(1)$ process

$$
x_{t}=\phi x_{t-1}+u_{t}
$$

Solution: $E x_{t}^{4}=E\left(\sum_{j=0}^{\infty} \phi^{j} u_{t-j}\right)^{4}$

$$
\begin{aligned}
& =\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} \phi^{j} \phi^{k} \phi^{l} \phi^{m} E u_{t-j} u_{t-k} u_{t-l} u_{t-m} \\
& \leq 3 \sigma^{4} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty}|\phi|^{j}|\phi|^{k}|\phi|^{l}|\phi|^{m} \\
& =3 \sigma^{4}\left(\sum_{j=0}^{\infty}|\phi|^{j}\right)^{4} \\
& =3 \sigma^{4}\left(\frac{1}{1-|\phi|}\right)^{4}<\infty
\end{aligned}
$$

Rewriting the $\mathrm{MA}(\infty)$ representation of the $\mathrm{AR}(1)$ process

$$
x_{t}=\phi x_{t-1}+u_{t}
$$

as

$$
x_{t}=\sum_{j=-\infty}^{\infty} \psi_{j} u_{t-j},
$$

where $\psi_{j}=\phi^{j}$ for $j \geq 0$ and $\psi_{j}=0$ for $j<0$, we obtain

$$
\begin{aligned}
\gamma(k) & =\operatorname{Cov}\left(x_{t}, x_{t-k}\right)=E x_{t} x_{t-k}=E \sum_{j=-\infty}^{\infty} \psi_{j} u_{t-j} \sum_{j=-\infty}^{\infty} \psi_{j} u_{(t-k)-j} \\
& =\sum_{r=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \psi_{r} \psi_{j} E u_{t-r} u_{t-(j+k)}=\sum_{j=-\infty}^{\infty} \psi_{j+k} \psi_{j} \underbrace{E u_{t-(j+k)} u_{t-(j+k)}}_{\sigma^{2}} .
\end{aligned}
$$

## Lemma:

$E x_{s} x_{s+h} x_{t} x_{t+h}=\gamma(h)^{2}+\gamma(t-s)^{2}+\gamma(t-s+h) \gamma(t-s-h)$

## Proof:

$E x_{s} x_{s+h} x_{t} x_{t+h}$
$=E \sum \psi_{j} u_{s-j} \sum \psi_{j} u_{s+h-j} \sum \psi_{j} u_{t-j} \sum \psi_{j} u_{t+h-j}$
$=E \sum \psi_{j} u_{s-j} \sum \psi_{j} u_{(s-j)+h} \sum \psi_{j} u_{(s-j)+(t-s)} \sum \psi_{j} u_{(s-j)+(t-s+h)}$
$=E \sum \psi_{j} u_{s-j} \sum \psi_{j-h} u_{s-j} \sum \psi_{j-(t-s)} u_{s-j} \sum \psi_{j-(t-s+h)} u_{s-j}$
$=\sum_{j} \sum_{k} \sum_{l} \sum_{m} \psi_{j} \psi_{k-h} \psi_{l-(t-s)} \psi_{m-(t-s+h)} E u_{s-j} u_{s-k} u_{s-l} u_{s-m}$
$=\sum_{j=k=l=m} \cdots+\sum \sum_{j=k \neq l=m} \cdots+\sum \sum_{j=l \neq k=m} \cdots+\sum \sum_{j=m \neq k=l} \cdots$
$=\underbrace{\sum \psi_{j} \psi_{j-h} \psi_{j-(t-s)} \psi_{j-(t-s+h)}}_{\Psi} \underbrace{E u_{s-j}^{4}}_{3 \sigma^{4}}$
$+(\underbrace{\sum \psi_{j} \psi_{j-h} \sigma^{2}}_{\gamma(h)} \sum \psi_{k-(t-s)} \psi_{k-(t-s+h)} \sigma^{2}-\sigma^{4} \Psi)$
$+\left(\sum \psi_{j} \psi_{j-(t-s)} \sigma^{2} \sum \psi_{k-h} \psi_{k-(t-s+h)} \sigma^{2}-\sigma^{4} \Psi\right)$
$+\left(\sum \psi_{j} \psi_{j-(t-s+h)} \sigma^{2} \sum \psi_{k-h} \psi_{k-(t-s)} \sigma^{2}-\sigma^{4} \Psi\right)$


$$
\sum_{s=1}^{n} \sum_{t=1}^{n} w_{t-s}=\sum_{k=-(n-1)}^{n-1}(n-|k|) w_{k} .
$$

If it is known that $\mu=0$, the simplified sample covariance

$$
\hat{\gamma}(h)=\frac{1}{n-h} \sum_{t=1}^{n-h} x_{t} x_{t+h}=\frac{1}{n-h} \sum_{t=1}^{n-h}\left(x_{t}-\mu\right)\left(x_{t+h}-\mu\right), h>0,
$$

is a consistent estimator of $\gamma(h)$, because

$$
E(\hat{\gamma}(h))=\frac{1}{n-h} \sum_{t=1}^{n-h} E\left(x_{t}-\mu\right)\left(x_{t+h}-\mu\right)=\frac{1}{n-h}(n-h) \gamma(h)=\gamma(h)
$$ and

$$
0 \leq \operatorname{Var}(\hat{\gamma}(h))=E\left(\frac{1}{n-h} \sum_{t=1}^{n-h} x_{t} x_{t+h}\right)^{2}-\gamma(h)^{2}
$$

$$
=\frac{1}{(n-h)^{2}} \sum_{s=1}^{n-h} \sum_{t=1}^{n-h} E x_{s} x_{s+h} x_{t} x_{t+h}-\gamma(h)^{2}
$$

$$
=\frac{1}{(n-h)^{2}} \sum_{s=1}^{n-h} \sum_{t=1}^{n-h}\left[\gamma(h)^{2}+\gamma(t-s)^{2}+\gamma((t-s)+h) \gamma((t-s)-h)\right]-\gamma(h)^{2}
$$

$$
\begin{aligned}
& =\frac{1}{(n-h)^{2}} \sum_{s=1}^{n-h} \sum_{t=1}^{n-h}\left[\gamma(t-s)^{2}+\gamma((t-s)+h) \gamma((t-s)-h)\right] \\
& =\frac{1}{(n-h)^{2}} \sum_{k=-(n-h-1)}^{n-h-1}((n-h)-|k|)\left(\gamma(k)^{2}+\gamma(k+h) \gamma(k-h)\right) \\
& \leq \frac{2}{n-h} \sum_{\mathrm{k}=0}^{\mathrm{n}-\mathrm{h}-1}\left(\left(\frac{\phi^{|k|} \sigma^{2}}{1-\phi^{2}}\right)^{2}+\frac{|\varphi|^{|k+h| \sigma^{2}}}{1-\varphi^{2}} \frac{|\varphi|^{|k-h| \sigma^{2}}}{1-\varphi^{2}}\right) \\
& \leq \frac{2}{n-h} \frac{\sigma^{4}}{\left(1-\phi^{2}\right)^{2}} \sum_{k=0}^{n-h-1}\left(\left(\phi^{2}\right)^{k}+\phi^{2 k}\right)=\frac{2}{n-h} \frac{\sigma^{4}}{\left(1-\phi^{2}\right)^{2}} 2 \frac{1-\left(\phi^{2}\right)^{n-h}}{1-\phi^{2}} \rightarrow 0 .
\end{aligned}
$$

Note that $|\phi|^{|k+h|+|k-h|} \leq|\phi|^{2 k}=\phi^{2 k}$, because

$$
|k+h|+|k-h| \geq|(k+h)+(k-h)|=|2 k|=2 k .
$$

Remark: We have shown above that the first and second sample moments of a Gaussian AR(1) process converge in probability to their theoretical counterparts. Such a process is said to be ergodic for the first and second moments.

The OLS-estimator of the parameter $\phi$ is given by

$$
\hat{\phi}=\frac{\sum_{t=2}^{n} x_{t-1} x_{t}}{\sum_{t=2}^{n} x_{t-1}^{2}}=\frac{\sum_{t=2}^{n} x_{t-1}\left(\phi x_{t-1}+u_{t}\right)}{\sum_{t=2}^{n} x_{t-1}^{2}}=\phi+\frac{\sum_{t=2}^{n} x_{t-1} u_{t}}{\sum_{t=2}^{n} x_{t-1}^{2}} .
$$

The random variables

$$
z_{t}=u_{t} x_{t-1}
$$

satisfy

$$
E z_{t}=E u_{t} x_{t-1}=E u_{t} E x_{t-1}=0
$$

and

$$
\begin{aligned}
E\left(z_{t} \mid z_{t-1}, z_{t-2}, \ldots\right) & =E\left(u_{t} x_{t-1} \mid u_{t-1} x_{t-2}, u_{t-2} x_{t-3}, \ldots\right) \\
& =E u_{t} E\left(x_{t-1} \mid u_{t-1} x_{t-2}, u_{t-2} x_{t-3}, \ldots\right)=0
\end{aligned}
$$

because $u_{t}$ is independent of $x_{t-1}$ and $u_{t-1} x_{t-2}, u_{t-2} x_{t-3}, \ldots$
Thus, the sequence of random variables $z_{t}$ is a martingale difference sequence ${ }^{2}$.

[^1]Exercise: Show that the sequence of random variables $z_{t}$ is white noise.

Exercise: Show that the sequence of random variables

$$
w_{t}=\left(u_{t}^{2}-\sigma^{2}\right) x_{t-1}^{2}
$$

is a martingale difference sequence.
Exercise: Show that the sequence of random variables $w_{t}$ is white noise.

Remark: Since the random variables $u_{t}$ are Gaussian, the sequences of random variables $z_{t}$ and $w_{t}$, respectively, are not only weakly stationary but also strictly stationary, which means that the joint distribution of any finite subsequence of random variables with indices $t_{1}, \ldots, t_{k}$ is the same as that of the random variables with indices $t_{1+\tau}, \ldots, t_{k+\tau}$.

Weak laws of large numbers state conditions under which the sample mean converges in probability towards the population mean. The weak law of large numbers for strictly stationary martingale difference sequences $\varepsilon_{\mathrm{t}}$ requires only absolute integrability.

Thus,

$$
\begin{array}{r}
E\left|z_{t}\right|=E\left|u_{t}\right|\left|x_{t-1}\right|=E\left|u_{t}\right| E\left|x_{t-1}\right| \leq E\left(1+u_{t}^{2}\right) E\left(1+x_{t-1}^{2}\right)<\infty \\
\Rightarrow \overline{\mathrm{z}} \xrightarrow{p} 0
\end{array}
$$

and

$$
\begin{aligned}
& E\left|w_{t}\right|=E\left|u_{t}^{2}-\sigma^{2}\right| x_{t-1}^{2} \leq E u_{t}^{2} E x_{t-1}^{2}+\sigma^{2} E x_{t-1}^{2}<\infty \\
& \Rightarrow \bar{w} \xrightarrow{p} 0 .
\end{aligned}
$$

The central limit theorem for martingale difference sequences states that

$$
\sqrt{n} \bar{\varepsilon} \xrightarrow{L} N\left(0, E \varepsilon_{t}^{2}\right)
$$

if $E \varepsilon_{t}^{2}>0, E \varepsilon_{t}^{4}<\infty, \overline{\varepsilon^{2}} \xrightarrow{p} E \varepsilon_{t}^{2}$.

Thus,

$$
\sqrt{n-1} \bar{z}=\sqrt{n-1}\left(\frac{1}{n-1} \sum_{t=2}^{n} z_{t}\right) \xrightarrow{L} N\left(0, \frac{\sigma^{4}}{1-\phi^{2}}\right),
$$

because

$$
\begin{aligned}
& E z_{t}^{2}=E\left(u_{t} x_{t-1}\right)^{2}=E u_{t}^{2} E x_{t-1}^{2}=\sigma^{2} \frac{\left.\phi^{0}\right|_{\sigma^{2}}}{1-\phi^{2}}=\frac{\sigma^{4}}{1-\phi^{2}}>0, \\
& E z_{t}^{4}=E\left(u_{t} x_{t-1}\right)^{4}=E u_{t}^{4} E x_{t-1}^{4}=3 \sigma^{4} E x_{t}^{4}<\infty,
\end{aligned}
$$

$$
\text { and } \quad \frac{1}{n-1} \sum_{t=2}^{n} z_{t}^{2}=\frac{1}{n-1} \sum_{t=2}^{n} u_{t}^{2} x_{t-1}^{2}
$$

$$
\begin{aligned}
& =\underbrace{\frac{1}{n-1} \sum_{t=2}^{n}\left(u_{t}^{2}-\sigma^{2}\right) x_{t-1}^{2}}_{=\bar{w}}+\frac{\sigma^{2}}{n-1} \sum_{t=2}^{n} x_{t-1}^{2} \\
& \xrightarrow{p} 0+\sigma^{2} \frac{\sigma^{2}}{1-\varphi^{2}}=E z_{t}^{2} .
\end{aligned}
$$

Now it is straightforward to derive the asymptotic distribution of the OLS-estimator $\hat{\phi}$.

The numerator of the statistic

$$
\sqrt{n-1}(\hat{\phi}-\phi)=\frac{\frac{\sqrt{n-1}}{n-1} \sum_{t=2}^{n} x_{t-1} u_{t}}{\frac{1}{n-1} \sum_{t=2}^{n} x_{t-1}^{2}}
$$

converges in law to

$$
N\left(0, \frac{\sigma^{4}}{1-\phi^{2}}\right)
$$

and its denominator converges in probability to

$$
\frac{\sigma^{2}}{1-\varphi^{2}}
$$

hence

$$
\begin{aligned}
\sqrt{n-1}(\hat{\phi}-\phi) \xrightarrow{L} & \operatorname{plim} \frac{1}{\frac{1}{n-1} \sum_{t=2}^{n} x_{t-1}^{2}} N\left(0, \frac{\sigma^{4}}{1-\phi^{2}}\right) \\
& =\frac{1}{p \lim \frac{1}{n-1} \sum_{t=2}^{n} x_{t-1}^{2}} N\left(0, \frac{\sigma^{4}}{1-\phi^{2}}\right) \\
& =\frac{1}{\frac{\sigma^{2}}{1-\phi^{2}}} N\left(0, \frac{\sigma^{4}}{1-\phi^{2}}\right) \\
& =N\left(0, \frac{\frac{\sigma^{4}}{1-\phi^{2}}}{\left(\frac{\sigma^{2}}{1-\phi^{2}}\right)^{2}}\right) \\
& =N\left(0,1-\phi^{2}\right) .
\end{aligned}
$$


[^0]:    ${ }^{1}$ For strong consistency, almost sure convergence is required rather than convergence in probability.

[^1]:    ${ }^{2}$ In contrast, for a martingale sequence it is required that $E\left|z_{t}\right|<\infty, E\left(z_{t} \mid z_{t-1}, z_{t-2}, \ldots\right)=z_{t-1}$.

