Quantitative Macroeconomics

Winter 2023/24

Week 3

Willi Mutschler
willi@mutschler.eu

Version: 1.0
Latest version available on: GitHub

Contents

1. Properties of AR(1) 1

2. Properties AR(1) With Time Trend 2

3. Law Of Large Numbers 3

4. Central Limit Theorem For Dependent Data 4

A. Solutions 7

1. Properties Of AR(1)

Let $\{\varepsilon_t\}$ be a white noise process with variance σ_ε^2.

1. Consider the univariate first-order autoregressive process AR(1):

$$y_t = c + \phi y_{t-1} + \varepsilon_t$$

   where $|\phi| < 1$ and $c \in \mathbb{R}$. Derive the unconditional first and second moments.

   Hint:

   - If $|\phi| < 1$, then $\sum_{j=0}^{\infty} \phi^j = \frac{1}{(1-\phi)}$

2. Do the following simulation exercise:

   - Simulate AR(1) processes for $t = 1, \ldots, 200$ with $c = 2$, $\sigma_\varepsilon^2 = 1^2$ and different values of $\phi = \{-0.8; 0.4; 0.9; 1.01\}$.
   - Plot the corresponding autocorrelation functions (ACF) for data vector y and maximum number of lags $p^{\max} = 8$ using MATLAB's `autocorr(y,'NumLags',8)` function.
   - Write a function `acfPlots(`y`,`p^{max}`,`α`)` that plots the autocorrelation function (ACF) of the data vector y with maximum number of lags $p^{\max}$. Your plot should also include an approximate (1-α)% confidence interval around zero. Your plots should look similar to MATLAB's `autocorr` command.

   Hints:

   - The empirical autocorrelation function at lag k is defined as $\hat{\rho}_k = \hat{\gamma}_k / \hat{\gamma}_0$ where

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=k+1}^{T} (y_t - \bar{y})(y_{t-k} - \bar{y})$$

     and

$$\hat{\gamma}_0 = \frac{1}{T} \sum_{t=1}^{T} (y_t - \bar{y})(y_t - \bar{y})$$

   - You can either use a for-loop to compute the sum or use vectors: $(y - \bar{y})'(y - \bar{y})$.
   - The sample autocorrelation function is an estimate of the actual autocorrelation only if the process is stationary. If the process is purely random, that is, all members are mutually independent and identically distributed, then y_t and y_{t-k} should be stochastically independent for any $k \neq 0$. According to the Lindeberg-Levy central limit theorem (see corresponding exercise), the normalized estimated autocorrelations are asymptotically standard normally distributed, i.e. $\sqrt{T}\hat{\rho}_k \to U \sim N(0,1)$ and thus $\hat{\rho}_k \to \tilde{U} \sim N(0, 1/T)$.

Readings

- Bjørnland and Thorsrud (2015, Ch.2)
- Lütkepohl (2004)

2. Properties AR(1) With Time Trend

Consider the univariate AR(1) model with a constant and time trend

$$y_t = c + d \cdot t + \phi y_{t-1} + u_t$$

where $u_t \sim WN(0, \sigma_u^2)$, $|\phi| < 1$, $c \in \mathbb{R}$ and $d \in \mathbb{R}$.

1. Compute the unconditional first and second moments, i.e. the unconditional mean, variance, autocovariance and autocorrelation function of y_t.

2. Why is this process not covariance-stationary? How could one proceed to make it covariance-stationary?

Hints:

- If $|\phi| < 1$, then $\sum_{j=0}^{\infty} \phi^j = \frac{1}{(1-\phi)}$
- If $|\phi| < 1$, then $\sum_{j=0}^{\infty} j\phi^j = \frac{\phi}{(1-\phi^2)}$

Readings

- Lütkepohl (2004)

3. Law Of Large Numbers

Let $Y_1, Y_2, \ldots$ be an iid sequence of arbitrarily distributed random variables with finite variance σ_Y^2 and expectation μ. Define the sequence of random variables

$$\overline{Y}_T = \frac{1}{T}\sum_{t=1}^{T} Y_t$$

1. Briefly outline the intuition behind the “law of large numbers”. What are the differences between “almost-sure convergence” and “convergence in probability”?

2. Write a program to illustrate the law of large numbers for uniformly distributed random variables (you may also try different distributions such as normal, gamma, geometric, student’s t with finite or infinite variance). To this end, do the following:
   - Set $T = 10000$ and initialize the $T \times 1$ output vector u.
   - Choose values for the parameters of the uniform distribution. Note that $E[u] = (a + b)/2$, where a is the lower and b the upper bound of the uniform distribution.
   - For $t = 1, \ldots, T$ do the following computations:
     - Draw t random variables from the uniform distribution with lower bound a and upper bound b.
     - Compute and store the mean of the drawn values in your output vector at position t.
   - Plot your output vector and add a line to indicate the theoretical mean of the uniform distribution.

3. Now suppose that the sequence $Y_1, Y_2, \ldots$ is an $AR(1)$ process:

   $$Y_t = \phi Y_{t-1} + \varepsilon_t$$

   where $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$ is not necessarily normally distributed and $|\phi| < 1$. Illustrate numerically that the law of large numbers still holds despite the intertemporal dependence.

Readings

- Lütkepohl (2005, App. C)
- Neusser (2016, App. C)
- Ploberger (2010)
- White (2001, Ch. 3)

4. Central Limit Theorem For Dependent Data

Suppose that the sequence $Y_1, Y_2, \ldots$ is an $AR(1)$ process, i.e.

$$Y_t - \mu = \phi \left(Y_{t-1} - \mu\right) + \varepsilon_t$$

where $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$ is (not necessarily but in our case) normally distributed and $|\phi| < 1$.

1. Briefly state and describe the intuition of the "Lindeberg-Levy Central Limit Theorem" for iid random variables. What does "convergence in distribution" mean? Why can we not use the theorem for the $AR(1)$ process?

2. Show that Y_t has mean equal to μ and finite variance equal to $\sigma_\varepsilon^2/(1-\phi^2)$.

3. To derive the asymptotic distribution of the sample mean, do the following steps:
   a) Derive the asymptotic distribution of $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t$.
   b) Show that

   $$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t = \sqrt{T} \left[(1-\phi)\left(\hat{\mu} - \mu\right) + \phi \left(\frac{Y_T - Y_0}{T} \right) \right]$$

   with $\hat{\mu} = \frac{1}{T} \sum_{t=1}^{T} Y_t$.
   c) Show that

   $$plim \left[\frac{\phi}{1-\phi} \left(\frac{Y_T - Y_0}{\sqrt{T}} \right) \right] = 0$$

   Hint: Use Tchebychev's Inequality, i.e. for a random variable X with expectation μ_x and finite variance σ_x^2:

   $$\Pr(|X - \mu_x| > \delta) \leq \frac{\sigma_x^2}{\delta^2}$$

   for any small real number $\delta > 0$.
   d) Put your results of (a), (b) and (c) together and derive the asymptotic distribution of the sample mean. That is, show that

   $$Z_T = \sqrt{T} \frac{\hat{\mu} - \mu}{\sigma_Z} \xrightarrow{d} U \sim N(0,1)$$

   for $\sigma_Z = \sqrt{\sigma_\varepsilon^2/(1-\phi)^2}$.

4. Write a program to demonstrate the central limit theorem for the AR(1) process. To this end:
   - Simulate $B = 5000$ stationary (e.g. $\phi = 0.8$) AR(1) processes with each $T = 10000$ observations. Store these in a $T \times B$ matrix Y.
   - Compute $\hat{\mu}$ for each column of Y.
   - Plot the histograms of the standardized variables according to the Lindeberg-Levy Central Limit Theorem:

   $$\widetilde{Z}_T = \sqrt{T} \frac{\hat{\mu} - \mu}{\sigma_\varepsilon / \sqrt{1-\phi^2}}$$

   and of the correct standardized variables that we derived in 3(d):

   $$Z_T = \sqrt{T} \frac{\hat{\mu} - \mu}{\sigma_\varepsilon/(1-\phi)}$$

   Compare the histograms to the standard normal distribution.

Readings

- Crack and Ledoit (2010)
- Lütkepohl (2005, App. C)
- Neusser (2016, App. C)
- White (2001, Ch. 5)

References

Bjørnland, Hilde Christiane and Leif Anders Thorsrud (2015). *Applied Time Series for Macroeconomics*. 2. utgave, 1. opplag. Oslo: Gyldendal Akademisk. ISBN: 978-82-05-48089-6.

Crack, Timothy Falcon and Oliver Ledoit (2010). "Central Limit Theorems When Data Are Dependent: Addressing the Pedagogical Gaps". In: *Journal of Financial Education* 36.1/2, pp. 38–60. ISSN: 00933961, 2332421X. JSTOR: 41948634. URL: http://www.jstor.org/stable/41948634.

Lütkepohl, Helmut (2004). "Univariate Time Series Analysis". In: *Applied Time Series Econometrics*. Ed. by Helmut Lütkepohl and Markus Krätzig. 1st ed. Cambridge University Press, pp. 8–85. ISBN: 978-0-521-83919-8 978-0-521-54787-1 978-0-511-60688-5. URL: https://doi.org/10.1017/CBO9780511606885.003.

– (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer. ISBN: 978-3-540-40172-8.

Neusser, Klaus (2016). *Time Series Econometrics*. Springer Texts in Business and Economics. Cham: Springer. ISBN: 978-3-319-81387-5 978-3-319-32861-4. URL: https://doi.org/10.1007/978-3-319-32862-1.

Ploberger, Werner (2010). "Law(s) of Large Numbers". In: *Macroeconometrics and Time Series Analysis*. Ed. by Steven N. Durlauf and Lawrence E. Blume. London: Palgrave Macmillan UK, pp. 158–162. ISBN: 978-0-230-23885-5 978-0-230-28083-0. URL: https://doi.org/10.1057/9780230280830_33.

White, Halbert (2001). *Asymptotic Theory for Econometricians*. Rev. ed. San Diego: Academic Press. ISBN: 978-0-12-746652-1.

A. Solutions

1 Solution to Properties of AR(1)

1. **Geometric sequence:** First, note that we can use the geometric sequence with or without the lag operator, i.e.

$$(1-\phi)^{-1} = \lim_{j\to\infty}(\phi^0 + \phi^1 + \phi^2 + \dots + \phi^j) = \sum_{j=0}^{\infty}\phi^j$$
$$(1-\phi L)^{-1} = \lim_{j\to\infty}((\phi L)^0 + (\phi L)^1 + (\phi L)^2 + \dots + (\phi L)^j) = \sum_{j=0}^{\infty}(\phi L)^j$$

The proof for this is pretty simple. Denote

$$S_k = \sum_{j=0}^{k}\phi^j = 1 + \phi^1 + \phi^2 + \phi^3 + \dots + \phi^k$$

then multiply with ϕ:

$$\phi S_k = \phi^1 + \phi^2 + \phi^3 + \dots + \phi^{k+1}$$

Now look at $S_k - \phi S_k = (1-\phi)S_k$:

$$(1-\phi)S_k = 1 - \phi^{k+1} \Leftrightarrow S_k = \frac{1}{1-\phi} - \frac{\phi^{k+1}}{1-\phi}$$

Looking at the limit of S_k for $k \to \infty$, we get

$$\lim_{k\to\infty} S_k = \frac{1}{1-\phi}$$

Next let's get a representation of the process y_t that is useful to compute the moments.

We can do this in different ways:

- **Recursive substitution** (starting at some infinite time j):

$$\begin{aligned}
y_t &= c + \phi y_{t-1} + \varepsilon_t \\
&= c + \phi\left(c + \phi y_{t-2} + \varepsilon_{t-1}\right) + \varepsilon_t \\
&= c + \phi c + \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2(c + \phi y_{t-3} + \varepsilon_{t-2}) \\
&\vdots \\
&= c + \phi c + \phi^2 c^2 + \dots + \phi^j c^j + \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \dots + \phi^j\varepsilon_{t-j} + \phi^{j+1}y_{t-j+1}
\end{aligned}$$

  y_t is a linear function of an initial value $\phi^{j+1}y_{t-j+1}$, historical values of the white noise process ε_t, and a sum of polynomials in c. If $|\phi| < 1$ and j becomes large, then $\phi^{j+1}y_{t-j+1} \to 0$, thus we get a so-called $MA(\infty)$ process:

$$\begin{aligned}
y_t &= \underbrace{c + \phi c + \phi^2 c^2 + \dots}_{\sum_{j=0}^{\infty}\phi^j c^j} + \underbrace{\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \dots}_{\sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j}} \\
&= c\sum_{j=0}^{\infty}\phi^j + \sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j} = \frac{c}{1-\phi} + \sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j}
\end{aligned}$$

- **With Lag Operators**: works only if $|\phi| < 1$ and $\{y_t\}$ is bounded (that is, there exists a finite number k such that $|y_t| < k$ for all t). Then

$$(1-\phi L)y_t = c + \varepsilon_t$$
$$(1-\phi L)^{-1}(1-\phi L)y_t = y_t = (1-\phi L)^{-1}c + (1-\phi L)^{-1}\varepsilon_t$$

  Using the geometric series, we get:

$$\begin{aligned} y_t &= (1+\phi L + \phi^2 L^2 + \dots + (\phi L)^j)c + (1+\phi L + \phi^2 L^2 + \dots + (\phi L)^j)\varepsilon_t \\ &= \Big(c + \phi c + \phi^2 c + \dots\Big) + \Big(\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \dots\Big) = c\sum_{j=0}^{\infty}\phi^j + \sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j} \\ &= \frac{c}{1-\phi} + \sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j} \end{aligned}$$

If we can express an AR process as a MA process, we call this process invertible. Let's now compute the moments from the MA(∞) representation and using the fact that ε_t is a white noise process:

- **Unconditional Mean:**

$$\begin{aligned} E[y_t] &= E\left[\frac{c}{1-\phi}\right] + E\left[\sum_{j=1}^{\infty}\phi^j\varepsilon_{t-j}\right] = \frac{c}{1-\phi}\sum_{j=1}^{\infty}\phi^j E[\varepsilon_{t-j}] \\ &= \underbrace{\frac{c}{1-\phi}}_{:=\mu} \end{aligned}$$

  As this process is covariance-stationary, the unconditional mean is time invariant. We typically denote this time-independence by using the greek letter μ.

- **Unconditional variance:**

$$\begin{aligned} Var[y_t] &= E\left[(y_t - E[y_t])(y_t - E[y_t])\right] = E\left[\left(\sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j}\right)\left(\sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j}\right)\right] \\ &= E\Big[\phi^0\phi^0\varepsilon_t\varepsilon_t + \phi^0\phi^1\varepsilon_t\varepsilon_{t-1} + \phi^0\phi^2\varepsilon_t\varepsilon_{t-2} + \phi^0\phi^3\varepsilon_t\varepsilon_{t-3} + \dots \\ &\qquad \phi^1\phi^0\varepsilon_{t-1}\varepsilon_t + \phi^1\phi^1\varepsilon_{t-1}\varepsilon_{t-1} + \phi^1\phi^2\varepsilon_{t-1}\varepsilon_{t-2} + \phi^1\phi^3\varepsilon_{t-1}\varepsilon_{t-3} + \dots \\ &\qquad \phi^2\phi^0\varepsilon_{t-2}\varepsilon_t + \phi^2\phi^1\varepsilon_{t-2}\varepsilon_{t-1} + \phi^2\phi^2\varepsilon_{t-2}\varepsilon_{t-2} + \phi^2\phi^3\varepsilon_{t-2}\varepsilon_{t-3} + \dots \\ &\qquad \dots] \\ &= \phi^0\phi^0 E[\varepsilon_t\varepsilon_t] + \phi^0\phi^1 E[\varepsilon_t\varepsilon_{t-1}] + \phi^0\phi^2 E[\varepsilon_t\varepsilon_{t-2}] + \phi^0\phi^3 E[\varepsilon_t\varepsilon_{t-3}] + \dots \\ &\qquad \phi^1\phi^0 E[\varepsilon_{t-1}\varepsilon_t] + \phi^1\phi^1 E[\varepsilon_{t-1}\varepsilon_{t-1}] + \phi^1\phi^2 E[\varepsilon_{t-1}\varepsilon_{t-2}] + \phi^1\phi^3 E[\varepsilon_{t-1}\varepsilon_{t-3}] + \dots \\ &\qquad \phi^2\phi^0 E[\varepsilon_{t-2}\varepsilon_t] + \phi^2\phi^1 E[\varepsilon_{t-2}\varepsilon_{t-1}] + \phi^2\phi^2 E[\varepsilon_{t-2}\varepsilon_{t-2}] + \phi^2\phi^3 E[\varepsilon_{t-2}\varepsilon_{t-3}] + \dots \\ &\qquad \dots \end{aligned}$$

  Note that ε_t is a white-noise process with variance $E[\varepsilon_{t-j}\varepsilon_{t-j}] = \sigma^2_\varepsilon$ for any j, but zero autocovariance, i.e. $E[\varepsilon_{t-j}\varepsilon_{t-k}] = 0$ for any $j \neq k$. Therefore:

$$\begin{aligned} Var[y_t] &= \phi^0\phi^0 E[\varepsilon_t\varepsilon_t] + \phi^1\phi^1 E[\varepsilon_{t-1}\varepsilon_{t-1}] + \phi^2\phi^2 E[\varepsilon_{t-2}\varepsilon_{t-2}] + \phi^3\phi^3 E[\varepsilon_{t-3}\varepsilon_{t-3}] + \dots \\ &= \sum_{j=0}^{\infty}(\phi^2)^j E[\varepsilon_{t-j}\varepsilon_{t-j}] = \sum_{j=0}^{\infty}(\phi^2)^j\sigma^2_\varepsilon = \sigma^2_\varepsilon\sum_{j=0}^{\infty}(\phi^2)^j \\ &= \underbrace{\sigma^2_\varepsilon\frac{1}{1-\phi^2}}_{:=\gamma_0} \end{aligned}$$

As the process is covariance-stationary, the unconditional variance is time invariant. We typically denote this time-independence by using the Greek letter γ_0.

- **Unconditional autocovariance:**

$$
\begin{aligned}
Var[y_t, y_{t-k}] = E\left[(y_t - E[y_t])(y_t - E[y_{t-k}])\right] &= E\left[\left(\sum_{j=0}^{\infty}\phi^j\varepsilon_{t-j}\right)\left(\sum_{j=0}^{\infty}\phi^j\varepsilon_{t-k-j}\right)\right] \\
&= E\Big[\phi^0\phi^0\varepsilon_t\varepsilon_{t-k} + \phi^0\phi^1\varepsilon_t\varepsilon_{t-k-1} + \phi^0\phi^2\varepsilon_t\varepsilon_{t-k-2} + \phi^0\phi^3\varepsilon_t\varepsilon_{t-k-3} + \dots \\
&\quad \phi^1\phi^0\varepsilon_{t-1}\varepsilon_{t-k} + \phi^1\phi^1\varepsilon_{t-1}\varepsilon_{t-k-1} + \phi^1\phi^2\varepsilon_{t-1}\varepsilon_{t-k-2} + \phi^1\phi^3\varepsilon_{t-1}\varepsilon_{t-k-3} + \dots \\
&\quad \phi^2\phi^0\varepsilon_{t-2}\varepsilon_{t-k} + \phi^2\phi^1\varepsilon_{t-2}\varepsilon_{t-k-1} + \phi^2\phi^2\varepsilon_{t-2}\varepsilon_{t-k-2} + \phi^2\phi^3\varepsilon_{t-2}\varepsilon_{t-k-3} + \dots \\
&\quad \dots\Big]
\end{aligned}
$$

This can be simplified due to the white noise property of ε_t to:

$$
\begin{aligned}
Var[y_t, y_{t-k}] &= \phi^k(\phi^0 E[\varepsilon_{t-k}\varepsilon_{t-k}] + \phi^2 E[\varepsilon_{t-k-1}\varepsilon_{t-k-1}] + \phi^4 E[\varepsilon_{t-k-2}\varepsilon_{t-k-2}] + \dots) \\
&= \phi^k \sum_{j=0}^{\infty}(\phi^2)^j\sigma_\varepsilon^2 = \phi^k\frac{\sigma_\varepsilon^2}{1-\phi^2} = \underbrace{\phi^k\gamma_0}_{:=\gamma_k}
\end{aligned}
$$

As the process is covariance-stationary, the unconditional autocovariance is only dependent on the time difference k. We typically denote this by using the Greek letter γ_k.

2. Here is a possible run-script:

progs/matlab/acfPlots_run.m

```
% ---------------------------------------------------------------------------
% Run script for simulating AR(1) processes and plotting their empirical
% autocorrelation function
% ---------------------------------------------------------------------------
% Willi Mutschler, November 07, 2022
% willi@mutschler.eu
% ---------------------------------------------------------------------------

%% Housekeeping
clearvars; clc;close all;

%% Generate and plot autoregressive processes
phi=[-0.8 0.4 0.9 1.01];     % different values for the phi coefficient
c=2;                         % value for constant
sigma=1;                     % value for standard deviation of white noise, experiment with different values
T=200;                       % value for number of observations
Y=nan(T,size(phi,2));        % initialize output vector with nan
Y(1,:)=c./(1-phi); % set first period equal to unconditional mean

for j=1:size(phi,2) % loop over coefficients
    for t=2:T % begin loop to compute AR(1) at t=2, as there is no y(0,j), i.e. you cannot index with 0
        Y(t,j)=phi(j)*Y(t-1,j)+randn()*sigma; % Simulate time seriers, randn simply generates one draw from N(0,1), we scale the standard deviation with sigma
    end
end
```

```matlab

%% Plot autocorrelation functions
% common figure
figure('Name','autocorrelation function')
sgtitle('autocorr (left) vs. ACFPlots (right)')

% use MATLAB's builin function
subplot(4,2,1); autocorr(Y(:,1),'NumLags',8); title('\phi=-0.8');
subplot(4,2,3); autocorr(Y(:,2),'NumLags',8); title('\phi=-0.4');
subplot(4,2,5); autocorr(Y(:,3),'NumLags',8); title('\phi=0.9');
subplot(4,2,7); autocorr(Y(:,4),'NumLags',8); title('\phi=1.01');

% use self-written function
subplot(4,2,2); acfPlots(Y(:,1),8,0.05); title('\phi=-0.8');
subplot(4,2,4); acfPlots(Y(:,2),8,0.05); title('\phi=-0.4');
subplot(4,2,6); acfPlots(Y(:,3),8,0.05); title('\phi=0.9');
subplot(4,2,8); acfPlots(Y(:,4),8,0.05); title('\phi=1.01');
```

and the corresponding `acfPlots.m` function:

progs/matlab/acfPlots.m

```matlab
function RHOHAT = acfPlots(y,pmax,alph)
% RHOHAT = acfPlots(y,pmax,alph)
% -------------------------------------------------------------------------
% Computes and plots the empirical autocorrelation function
% \hat{\gamma}_k = 1/T*\sum_{t=k+1}^T (y_t-\bar{y})(y_{t-k}-\bar{y})
% \hat{rho}_k = \hat{\gamma}_k/\hat{\gamma}_0
% -------------------------------------------------------------------------
% INPUTS
%   - y       [periods x 1] vector of data
%   - pmax    [scalar]      maximum number of lags to plot
%   - alph    [scalar]      significance level for asymptotic bands, e.g. 0.05
% -------------------------------------------------------------------------
% OUTPUTS
%   - RHOHAT [1 x pmax]  Sample autocorrelation coefficient
% -------------------------------------------------------------------------
% Willi Mutschler, November 07, 2022
% willi@mutschler.eu
% -------------------------------------------------------------------------

T=size(y,1);               % get number of periodes
y_demeaned = y-mean(y);    % put y in deviations from mean
RHOHAT = nan(1,pmax);      % initialize output vector

% Compute variance
c0 = 1/T*(y_demeaned' * y_demeaned);
% Compute autocorrelations
for k=1:pmax
    c_k = 1/T * (y_demeaned(1+k:T,:)' * y_demeaned(1:T-k,:));
    RHOHAT(1,k) = c_k/c0;
end

% Asymptotic bands
```

```
critval = norminv(1-alph/2);
ul = repmat(critval/sqrt(T),pmax,1);
ll = -1*ul;

% Barplots
bar(RHOHAT);
hold on;
plot(1:pmax,ul,'color','black','linestyle','--');
plot(1:pmax,ll,'color','black','linestyle','--');
hold off;

% The following is just for pretty plots
acfbarplot = gca; % Get current axes handle
acfbarplot.Title.String = 'Sample autocorrelation coefficients';
acfbarplot.XAxis.Label.String = 'lags';
acfbarplot.XAxis.TickValues = 1:pmax;
acfbarplot.YAxis.Label.String = 'acf value';
acfbarplot.YAxis.Limits = [-1 1];
acfbarplot.XAxis.Limits = [0 pmax];

end % function end
```

2 Solution to Properties AR(1) With Time Trend

1. First, let's get a MA representation using the lag operator or recursive substitution techniques:

$$\begin{aligned} & y_t = c + d \cdot t + \phi y_{t-1} + u_t \\ \Leftrightarrow\ & y_t = \frac{c + d \cdot t + u_t}{1 - \phi L} \\ \Leftrightarrow\ & y_t = \sum_{j=0}^{\infty} \phi^j (c + d(t-j) + u_{t-j}) = \frac{c}{1-\phi} + \frac{dt}{1-\phi} - d \sum_{j=0}^{\infty} \phi^j j + \sum_{j=0}^{\infty} \phi^j u_{t-j} \end{aligned}$$

As $|\phi| < 1$ the geometric series holds. For $\sum_{j=0}^{k} \phi^j j$ we also have a closed-form formula. Therefore, y_t is given by

$$y_t = \frac{c}{1-\phi} + \frac{dt}{1-\phi} - d\frac{\phi}{(1-\phi)^2} + \sum_{j=0}^{\infty} \phi^j u_{t-j}.$$

Unconditional mean:

$$\begin{aligned} E[y_t] &= \frac{c}{1-\phi} + \frac{dt}{1-\phi} - d\frac{\phi}{(1-\phi)^2} + \underbrace{\sum_{j=0}^{\infty} \phi^j E[u_{t-j}]}_{=0,\text{ as } u_t \sim iid(0,\sigma^2)} \\ &= \frac{c}{1-\phi} + \frac{dt}{1-\phi} - d\frac{\phi}{(1-\phi)^2} \end{aligned}$$

Unconditional variance:

$$\begin{aligned} \gamma_0 = Var[y_t] &= E[(y_t - E[y_t])^2] \\ &= E[(\sum_{j=0}^{\infty} \phi^j E[u_{t-j}]) \cdot (\sum_{j=0}^{\infty} \phi^j E[u_{t-j}])] \\ &= E[(u_t + \phi^1 u_{t-1} + \phi^2 u_{t-2} + \dots)(u_t + \phi^1 u_{t-1} + \phi^2 u_{t-2} + \dots)] \\ &= E[u_t^2 + 2\phi^1 u_t u_{t-1} + 2\phi^2 u_t u_{t-2} + \dots + \phi^2 u_{t-1}^2 + 2\phi^3 u_{t-1} u_{t-2} + 2\phi^4 u_{t-1} u_{t-3} + \dots \\ &+ \phi^4 u_{t-2}^2 + 2\phi^5 u_{t-2} u_{t-3} + 2\phi^5 u_{t-2} u_{t-4} + \dots] \\ &\overset{iid}{=} E[u_t^2 + \phi^2 u_{t-1}^2 + \phi^4 u_{t-2}^2 + \dots] \end{aligned}$$

with $Var[u_t] = E[u_t^2] - E[u_t]^2 = E[u_t^2] = \sigma^2$ we get

$$Var[y_t] = \sigma^2(\phi^0 + \phi^2 + \phi^4 + \dots) = \frac{\sigma^2}{1-\phi^2}.$$

Autocovariance:

$$
\begin{aligned}
\gamma(k) &= E[(y_t - E[y_t])(y_{t-k} - E[y_{t-k}])] \\
&= E[(\sum_{j=0}^{\infty} \phi^j u_{t-j})(\sum_{j=0}^{\infty} \phi^j u_{t-j-k})] \\
&= E[(u_t + \phi^1 u_{t-1} + \phi^2 u_{t-2} + \dots + \phi^k u_{t-k} + \phi^{k+1} u_{t-k-1} + \phi^{k+2} u_{t-k-2} + \dots) \\
&(u_{t-k} + \phi^1 u_{t-k-1} + \phi^2 u_{t-k-2} + \dots)] \\
&= E[u_t u_{t-k} + \phi^1 u_t u_{t-k-1} + \phi^2 u_t u_{t-k-3} + \dots \\
&\phi^1 u_{t-1} u_{t-k} + \phi^2 u_{t-1} u_{t-k-1} + \phi^3 u_{t-1} u_{t-k-2} + \dots \\
&\phi^2 u_{t-2} u_{t-k} + \phi^3 u_{t-2} u_{t-k-1} + \phi^4 u_{t-2} u_{t-k-2} + \dots \\
&\vdots \\
&\phi^k u_{t-k}^2 + 2\phi^{k+1} u_{t-k} u_{t-k-1} + 2\phi^{k+2} u_{t-k} u_{t-k-2} + \dots \\
&\phi^{k+2} u_{t-k-1}^2 + 2\phi^{k+3} u_{t-k-1} u_{t-k-2} + 2\phi^{k+4} u_{t-k-1} u_{t-k-3} + \dots \\
&\phi^{k+4} u_{t-k-2}^2 + 2\phi^{k+5} u_{t-k-2} u_{t-k-3} + 2\phi^{k+6} u_{t-k-2} u_{t-k-4} + \dots] \\
&\overset{iid}{=} E[\phi^k u_{t-k}^2 + \phi^{k+2} u_{t-k-1}^2 + \phi^{k+4} u_{t-k-2}^2 + \dots] \\
&= \phi^k \sigma^2 (\phi^0 + \phi^2 + \phi^4 + \dots) \\
&= \frac{\phi^k \sigma^2}{1 - \phi^2}
\end{aligned}
$$

Autocorrelation: $\rho_k = \frac{\gamma_k}{\gamma_0}$. Therefore:

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{\frac{\phi^k \sigma^2}{1-\phi^2}}{\frac{\sigma^2}{1-\phi^2}} = \phi^k.$$

2. The expectation is time-dependent, hence it is not stationary. One could subtract the expectation, e.g. look at $y_t^{demeaned} = y_t - E[y_t]$, where $y_t^{demeaned}$ is now covariance stationary. The unknown coefficient d must be estimated first though.

3 Solution to Law Of Large Numbers

1. In probability theory, the law of large numbers (LLN) is a theorem that describes the result of performing the same experiment a large number of times (repetitions, trials, experiments, iterations, sample size). According to the LLN, the average of the results obtained from a large number of times will be close to the theoretical expected value, and will tend to become closer as more iterations are performed. There are different laws of large numbers that differ in the underlying assumptions on the stochastic process. These laws are the cornerstones of asymptotic theory in statistics and econometrics.

   In this exercise, the LLN is about determining what happens to $\overline{Y}_T$ as $T \to \infty$ (note that $\overline{Y}_T$ is a random variable). The LLN states that this series converges to the (unknown) expected value $E[Y_t] = \mu$. More precisely, the strong LLN implies that at the limit, we can exactly determine μ, whereas the weak LLN implies that we can only approximately determine μ, even though we can make the approximation very close to the unknown number μ.

   Econometrically speaking:

   - Strong LLN means **almost-sure convergence**:
     At some point adding more observation does not matter at all for the average, $\overline{Y}_T$ will be exactly equal to the expected value μ. That is, the sequence $\overline{Y}_1, \overline{Y}_2, \ldots$ of random variables converges **almost surely** to the variable μ, if

     $$\Pr\left(\left\{\lim_{T\to\infty} \overline{Y}_T = \mu\right\}\right) = 1$$

     or simply:

     $$\overline{Y}_T \overset{a.s.}{\to} \mu$$

     This definition of convergence is not very important in Quantitative Macroeconomics.

   - Weak LLN means that the sample mean $\overline{Y}_T$ converges in probability to the population mean μ. That is, the sequence $\overline{Y}_1, \overline{Y}_2, \ldots$ of random variables converges **in probability** to the variable μ, if

     $$\lim_{T\to\infty} \Pr\left(|Y_T - \mu| < \delta\right) = 1$$

     As $T \to \infty$, the probability is approaching 1 very closely, but typically it will not be exactly equal to 1. In other words, the probability that the average is "far away" from the expectation μ is zero, where we measure closeness by an arbitrary small number $\delta > 0$. More compact notation:

     $$\begin{aligned} \overline{Y}_T &\overset{p}{\to} \mu \\ plim\, \overline{Y}_T &= \mu \end{aligned}$$

     This definition of convergence is very important in Quantitative Macroeconomics.

     In Quantitative Macroeconomics, we are mainly concerned with identically distributed processes that are either independent of each other (like the white noise process) or that are homogenously dependent (like the VAR(1) process). Given assumptions on existence and boundedness of the unconditional moments of these processes, the weak LLN typically applies.

2. Here is an extended illustration for several distributions:

progs/matlab/lawOfLargeNumbers.m

```
% ------------------------------------------------------------------------
% Illustration of the weak law of large numbers for several distributions:
```

```
% normal,uniform,geometric,student's t (finite and infinite variance),gamma.
% Note that the draws are i.i.d.
% -------------------------------------------------------------------------
% Willi Mutschler, October 28, 2022
% willi@mutschler.eu
% -------------------------------------------------------------------------

%% Housekeeping
clearvars; clc;close all;

%% Initializations
T = 10000;                                  % maximum horizon of periods
z = nan(1,T); sig_z = 0.2; mu_z = 10;       % normal distribution
u = nan(1,T); a = 2; b = 4; mu_u = (a+b)/2; % uniform distribution
ge = nan(1,T); p = 0.2; mu_ge = (1-p)/p;    % geometric distribution
ga = nan(1,T); k=2; thet=2; mu_ga=k*thet;   % gamma distribution
st1 = nan(1,T); nu1 = 8; mu_st1 = 0;        % student t with finite variance
st2 = nan(1,T); nu2 = 2; mu_st2 = 0;        % student t with infinite variance

%% Draw random variables
ZZ = mu_z + sig_z.*randn(1,T); % normal distribution
UU = a + (b-a).*rand(1,T);     % uniform distribution
GeGe = geornd(p,1,T);          % geometric distribution
GaGa = gamrnd(k,thet,1,T);     % gamma distribution
ST1 = trnd(nu1,1,T);           % student t with finite variance
ST2 = trnd(nu2,1,T);           % student t with infinite variance

%% Compute and store mean for growing sample sizes
wait = waitbar(0,'Please wait...'); % open waitbar
for t = 1:T
    % get random numbers with growing sample size
    Zt = ZZ(1:t);              % normal distribution
    Ut = UU(1:t);              % uniform distribution
    Get = GeGe(1:t);           % geometric distribution
    Gat = GaGa(1:t);           % gamma distribution
    ST1t = ST1(1:t);           % student t with finite variance
    ST2t = ST2(1:t);           % student t with infinite variance

    % Compute and store averages
    z(t) = mean(Zt);
    u(t) = mean(Ut);
    ge(t) = mean(Get);
    ga(t) = mean(Gat);
    st1(t) = mean(ST1t);
    st2(t) = mean(ST2t);

    waitbar(t/T); % update waitbar
end
close(wait); % close waitbar

%% Create plot for different distributions
figure('name','Law of Large Numbers for different distributions');
```

```
subplot(2,3,1);
    plot(z,'linewidth',2);
    line(0:T,repmat(mu_z,1,T+1),'linestyle','--','color','black');
    title('Normal');
subplot(2,3,2);
    plot(u,'linewidth',2);
    line(0:T,repmat(mu_u,1,T+1),'linestyle','--','color','black');
    title('Uniform');
subplot(2,3,3);
    plot(ge,'linewidth',2);
    line(0:T,repmat(mu_ge,1,T+1),'linestyle','--','color','black');
    title('Geometric');
subplot(2,3,4);
    plot(ga,'linewidth',2);
    line(0:T,repmat(mu_ga,1,T+1),'linestyle','--','color','black');
    title('Gamma');
subplot(2,3,5);
    plot(st1,'linewidth',2);
    line(0:T,repmat(mu_st1,1,T+1),'linestyle','--','color','black');
    title('Student''s t finite variance');
subplot(2,3,6);
    plot(st2,'linewidth',2);
    line(0:T,repmat(mu_st2,1,T+1),'linestyle','--','color','black');
    title('Student''s t infinite variance');
```

Note that for the t-distribution with infinite variance the weak LLN actually does not apply.

3. Here is an extended illustration for different error term distributions:

progs/matlab/lawOfLargeNumbersAR1.m

```
% ---------------------------------------------------------------------------
% Illustration of the weak law of large numbers for the AR(1) process based
% on different error term distributions. Distributions considered:
% normal,uniform,geometric,student's t (finite and infinite variance),gamma.
% Note that the AR(1) process is not i.i.d. (it is NOT independently distributed)
% ---------------------------------------------------------------------------
% Willi Mutschler, October 28, 2022
% willi@mutschler.eu
% ---------------------------------------------------------------------------

%% Housekeeping
clearvars; clc;close all;

%% Initializations
T = 10000;                      % maximum horizon of periods
sig_z = 0.2; mu_z = 10;         % parameters for normal distribution
a = 2; b = 4; mu_u = (a+b)/2;   % parameters for uniform distribution
p = 0.2; mu_ge = (1-p)/p;       % parameters for geometric distribution
k=2; thet=2; mu_ga=k*thet;      % parameters for gamma distribution
nu1 = 8; mu_st1 = 0;            % parameters for student t with finite variance
nu2 = 2; mu_st2 = 0;            % parameters for student t with infinite variance
phi=0.8; mu_y = 0;              % parameters for stable AR(1) process

```

```matlab
Y = nan(T,6);                   % initialize output vector: sample size in rows, distributions in columns
%% Draw random variables
ZZ = mu_z + sig_z.*randn(1,T); % normal distribution
UU = a + (b-a).*rand(1,T);     % uniform distribution
GeGe = geornd(p,1,T);          % geometric distribution
GaGa = gamrnd(k,thet,1,T);     % gamma distribution
ST1 = trnd(nu1,1,T);           % student t with finite variance
ST2 = trnd(nu2,1,T);           % student t with infinite variance

%% Compute and store mean for growing sample sizes
wait = waitbar(0,'Please wait...'); % open waitbar
for t = 1:T
    Yt = nan(t,6);  % initialize matrix for simulated data of sample size t in the rows
                    % different processes due to the 6 different error term distributions are in the columns
    Yt(1,:) = mu_y; % initialize the first observation with the unconditional mean

    if t>1
        for tt=2:t
            % Note that we demean the errors
            Yt(tt,1) = phi*Yt(tt-1,1) + (ZZ(tt)-mu_z);    % normal distribution with mean zero
            Yt(tt,2) = phi*Yt(tt-1,2) + (UU(tt)-mu_u);    % uniform
            Yt(tt,3) = phi*Yt(tt-1,3) + (GeGe(tt)-mu_ge); % geometric
            Yt(tt,4) = phi*Yt(tt-1,4) + (GaGa(tt)-mu_ga); % gamma
            Yt(tt,5) = phi*Yt(tt-1,5) + (ST1(tt)-mu_st1); % finite variance student t
            Yt(tt,6) = phi*Yt(tt-1,6) + (ST2(tt)-mu_st2); % infinite variance student t
        end
    end
    % Compute and store averages
    Y(t,:) = mean(Yt,1);
    waitbar(t/T); % update waitbar
end
close(wait); % close waitbar

%% Create plot for AR(1) with different distributions
figure('name','Law of Large Numbers for stable AR(1)');
subplot(2,3,1);
    plot(Y(:,1),'linewidth',2);
    line(0:T,repmat(mu_y,1,T+1),'linestyle','--','color','black');
    title('Normal errors');
subplot(2,3,2);
    plot(Y(:,2),'linewidth',2);
    line(0:T,repmat(mu_y,1,T+1),'linestyle','--','color','black');
    title('Uniform errors');
subplot(2,3,3);
    plot(Y(:,3),'linewidth',2);
```

```
    line(0:T,repmat(mu_y,1,T+1),'linestyle','--','color','black');
    title('Geometric errors');
subplot(2,3,4);
    plot(Y(:,4),'linewidth',2);
    line(0:T,repmat(mu_y,1,T+1),'linestyle','--','color','black');
    title('Gamma errors');
subplot(2,3,5);
    plot(Y(:,5),'linewidth',2);
    line(0:T,repmat(mu_y,1,T+1),'linestyle','--','color','black');
    title('Student''s t finite variance errors');
subplot(2,3,6);
    plot(Y(:,6),'linewidth',2);
    line(0:T,repmat(mu_y,1,T+1),'linestyle','--','color','black');
    title('Student''s t infinite variance errors');
```

Note that we need to make sure that $E[\varepsilon_t] = 0$ when we simulate data. We see that the weak law of large numbers holds under weaker conditions than iid. For instance, one can show that for the stationary AR(1), necessary and sufficient conditions are: $Var[y_t] < \infty$ and $|\gamma(k)| \to 0$ as $k \to \infty$. This does not hold for all considered distributions in the code.

4 Solution to Central Limit Theorem For Dependent Data

1. We usually consider the *Lindeberg-Levy Central Limit Theorem* for identically and independently (iid) random variables with finite mean μ and finite variance σ_Y^2. Then the *Lindeberg-Levy Central Limit Theorem* establishes the distribution of the sample mean $\overline{Y}_T$ for a growing sample size:

$$\sqrt{T}(\hat{\mu} - \mu) \xrightarrow{d} N(0, \sigma_Y^2)$$

with $\hat{\mu} = \bar{Y}_T = \frac{1}{T}\sum_{t=1}^{T} y_t$. Or more compactly using standardized variables:

$$z = \sqrt{T}\frac{\hat{\mu} - \mu}{\sigma_Y} \xrightarrow{d} U \sim N(0, 1)$$

The central limit theorem is closely related to the LLN, but while the LLN is a statement about **converging to a constant**, central limit theorems look at **convergence in distribution**, i.e. the distribution of the sample mean. More formally, a sequence $\bar{Y}_1, \bar{Y}_2, \ldots$ of random variables with distribution functions $F_1, F_2, \ldots$ converges **in distribution (weakly; in law)** to a variable μ with distribution function F, if

$$\lim_{T\to\infty} F_T(x) = F(x)$$

for all $x \in \mathbb{R}$ where $F(x)$ is continuous. Typically, we use the following notation for this:

$$\bar{Y}_T \xrightarrow{d} \mu$$

Unfortunately, the *Lindeberg-Levy Central Limit Theorem* does not apply for the AR(1) process as we have dependent and not iid data. For stationary AR(1) processes, we can however use similar central limit theorems either for Martingale-Difference-processes or mixing processes.

2. First let's derive the expectation and variance of the AR(1) process with $|\phi| < 1$. For this, we use recursive substitution techniques given a starting value Y_0:

$$Y_t = (1 - \phi)(1 + \phi + \phi^2 + \cdots + \phi^T)\mu + \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \cdots + \phi^T\varepsilon_{t-T} + \phi^{T+1}Y_0$$

Note that $\lim_{T\to\infty} \phi^{T+1} = 0$ and $\lim_{T\to\infty}\sum_{j=0}^{\infty}\phi^j = \frac{1}{1-\phi}$, since $|\phi| < 1$. The AR(1) process with $|\phi| < 1$ can therefore be equally represented by

$$Y_t = \mu + \sum_{j=1}^{\infty}\phi^j\varepsilon_{t-j}$$

Its expectation and variance are then equal to

$$E[Y_t] = \mu + \sum_{j=1}^{\infty}\phi^j E[\varepsilon_{t-j}] = \mu$$
$$Var[Y_t] = \sum_{j=1}^{\infty}(\phi^j)^2 Var[\varepsilon_{t-j}] = \sum_{j=1}^{\infty}(\phi^2)^j\sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{1 - \phi^2}$$

where we use the white noise property of ε_t.

3. Let's derive the asymptotic distribution of the sample mean:

   a) Due to the white noise assumption on ε_t, we can use the Lindeberg-Levy central limit theorem such that

$$\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t\right) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t \xrightarrow{d} U_\varepsilon \sim N(0, \sigma_\varepsilon^2)$$

b) Let's have a look at $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t$:

$$\begin{aligned}
\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}[(Y_t-\mu)-\phi(Y_{t-1}-\mu)] \\
&= \frac{1}{\sqrt{T}}\left[\sum_{t=1}^{T}(Y_t-\mu)-\phi\sum_{t=1}^{T}(Y_{t-1}-\mu)\right] \\
&= \frac{1}{\sqrt{T}}\left[\sum_{t=1}^{T}(Y_t-\mu)-\phi\left[\sum_{t=1}^{T}(Y_t-\mu)-(Y_T-Y_0)\right]\right] \\
&= \sqrt{T}\left[\frac{1}{T}\sum_{t=1}^{T}(Y_t-\mu)-\phi\left[\frac{1}{T}\sum_{t=1}^{T}(Y_t-\mu)-\left(\frac{Y_T-Y_0}{T}\right)\right]\right] \\
&= \sqrt{T}\left[\hat{\mu}-\mu-\phi\left[\hat{\mu}-\mu-\left(\frac{Y_T-Y_0}{T}\right)\right]\right] \\
&= \sqrt{T}\left[(1-\phi)\left(\hat{\mu}-\mu\right)+\phi\left(\frac{Y_T-Y_0}{T}\right)\right]
\end{aligned}$$

c) Using the definition of the probability limit, i.e.

$$plim\left[\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right]=0$$

is shorthand notation for

$$\lim_{T\to\infty} Pr\left(\left|\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right|>\delta\right)=0$$

We are going to show this by using Tchebychev's Inequality. That is, according to the inequality we have:

$$Pr\left(\left|\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right|>\delta\right)\leq\frac{1}{\delta^2}var\left[\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right]$$

for any $\delta>0$. Let's have a look at $var\left[\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right]$:

$$\begin{aligned}
var\left[\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right] &= \frac{1}{T}\left(\frac{\phi}{1-\phi}\right)^2 Var[Y_T-Y_0] \\
&= \frac{1}{T}\left(\frac{\phi}{1-\phi}\right)^2 (Var[Y_T]+Var[Y_0]-2Cov[Y_T,Y_0]] \\
&= \frac{1}{T}\left(\frac{\phi}{1-\phi}\right)^2\left[\frac{\sigma_\varepsilon^2}{1-\phi^2}+\frac{\sigma_\varepsilon^2}{1-\phi^2}-2Corr[Y_T,Y_0]\sqrt{\frac{\sigma_\varepsilon^2}{1-\phi^2}}\sqrt{\frac{\sigma_\varepsilon^2}{1-\phi^2}}\right] \\
&\leq \frac{1}{T}\left(\frac{\phi}{1-\phi}\right)^2 4\left(\frac{\sigma_\varepsilon^2}{1-\phi^2}\right)
\end{aligned}$$

since $corr(Y_T,Y_0)\geq -1$.

Thus for any $\delta>0$, we have

$$Pr\left(\left|\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right|>\delta\right)\leq\frac{1}{\delta^2}\frac{1}{T}\left(\frac{\phi}{1-\phi}\right)^2 4\left(\frac{\sigma_\varepsilon^2}{1-\phi^2}\right)$$

In the limit for $T\to\infty$ the right hand side converges to 0; hence:

$$\lim_{T\to\infty} Pr\left(\left|\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right|>\delta\right)=0.$$

or using our shorthand notation:

$$plim\left[\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right]=0$$

d) Now, let's go back to

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t = \sqrt{T}\left[(1-\phi)(\hat{\mu}-\mu) + \phi\left(\frac{Y_T - Y_0}{T}\right)\right]$$

Let's divide by $(1-\phi)$

$$\frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t}{1-\phi} = \sqrt{T}(\hat{\mu}-\mu) + \frac{\phi}{1-\phi}\left(\frac{Y_T - Y_0}{\sqrt{T}}\right)$$

For the left-hand-side we have

$$\frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t}{1-\phi} \xrightarrow{d} \tilde{U_\varepsilon} \sim N\left(0, \frac{\sigma_\varepsilon^2}{(1-\phi)^2}\right)$$

This is also the distribution of the right-hand side. However, in the limit, the right-hand side actually simplifies as we just derived that $plim\left[\frac{\phi}{1-\phi}\left(\frac{Y_T-Y_0}{\sqrt{T}}\right)\right] = 0$. Therefore:

$$\sqrt{T}(\hat{\mu}-\mu) \xrightarrow{d} \tilde{U} \sim N\left(0, \frac{\sigma_\varepsilon^2}{(1-\phi)^2}\right)$$

and we're done. We have just derived the distribution to which the sample mean $\hat{\mu} = \overline{Y}_T$ of an AR(1) process converges to asymptotically. Note that the required standardization is different than the Lindeberg-Levy central limit theorem would suggest. The correct standardized variable to consider is:

$$Z_T = \sqrt{T}\frac{\hat{\mu}-\mu}{\sigma_Z} \xrightarrow{d} U \sim N(0,1)$$

where we need to set $\sigma_Z^2 = \frac{\sigma_\varepsilon^2}{(1-\phi)^2}$ to get convergence to the standard normal distribution.

```
% ----------------------------------------------------------------------
% Illustration of Central Limit Theorem For Dependent Data (Gaussian AR(1))
% ----------------------------------------------------------------------
% Willi Mutschler, October 28, 2022
% willi@mutschler.eu
% ----------------------------------------------------------------------

% Housekeeping
clearvars; clc;close all;

% Initializations
B       = 5000;      % repetitions
T       = 10000;     % time periods, t=1,...,T
c       = 3;         % constant for AR(1)
phi     = 0.8;       % AR(1) parameter
mu      = c/(1-phi); % theoretical expectation of AR(1)
sig_eps = 0.5;       % standard deviation of error in AR(1) process
Y = nan(T,B);        % output matrix

% Compute distributions
Y(1,:) = repmat(mu,1,B); % Initialize first period to expectation of AR(1)
for b = 1:B
    epsi = sig_eps*randn(T,1);
```

```
    for t=2:T
        Y(t,b) = c + phi*Y(t-1,b) + epsi(t);
    end
end
muhat = mean(Y);              % average
var_Y = sig_eps^2/(1-phi^2); % analytical variance of an AR(1)-process
var_Z = sig_eps^2/(1-phi)^2; % variance of standardized variable

% Standardizations
ZT = sqrt(T).*(muhat - mu)./sqrt(var_Z);      % correct standardization
ZTnaive = sqrt(T).*(muhat - mu)./sqrt(var_Y); % naive standardization

% Plot histograms
f=figure('name','Central Limit Theorems');
x = -5:0.1:5; % values to plot normal distribution
subplot(1,2,1);
    histogram(ZTnaive,'Normalization','pdf');
    hold on;
    plot(x,normpdf(x),'linewidth',2);
    title('Lindeberg-Levy (wrong)');
    ylim([0 0.45]);
    hold off;
subplot(1,2,2);
    histogram(ZT,'Normalization','pdf');
    hold on;
    plot(x,normpdf(x),'linewidth',2);
    title('Dependent Data (correct)');
    ylim([0 0.45]);
    hold off;
sgtitle('Central Limit Theorems');
```