02417: Time Series Analysis
Lasse Engbo Christiansen
Fall 2019

DTU Compute
Technical University of Denmark

Reference solution to assignment 2: ARMA Processes and Seasonal Processes

NOTE that the model parametrizations in the assignment and in your favorite software package may not be identical. You will study how the choice of coefficients (of the operator polynomials in ARMA processes) affects the structure of the process, through simulated data and empirical autocorrelation functions. The assignment is to some extent specified in terms of R commands. Users of other software packages should find similar functions to answer the questions.

Question 2.1: Stability *Let the process X_t be given by*

$$X_t - 0.8X_{t-1} = \epsilon_t + 0.8\epsilon_{t-1} - 0.5\epsilon_{t-2}$$

where ϵ_t is a white noise process with $\sigma = 0.4$.

1. *Is the process stationary and invertible?*
2. *Investigate analytically the second order moment representation of the process.*
3. *Simulate 10 realisations with 200 observations from the process and plot them (Preferably in the same plot).*
4. *Estimate the ACF for each realisation and plot those (Again, preferably in the same plot). Comment on the results.*
5. *Repeat for the PACF of the same realisations.*
6. *Calculate the variance of each of the realisations.*
7. *Discuss and compare the analytical and numerical results.*

Q2.1.1: Stationary and invertible

The process is stationary since:

$$\begin{aligned} 1 - 0.8z^{-1} &= 0 \\ \Leftrightarrow z - 0.8 &= 0 \\ \Rightarrow z &= 0.8 < 1 \end{aligned}$$

Checking invertibility:

$$\begin{aligned} 1 + 0.8z^{-1} - 0.5z^{-2} &= 0 \\ \Leftrightarrow z^2 + 0.8z - 0.5 &= 0 \\ \Rightarrow z = -0.4 \pm 1/2\sqrt{0.8^2 + 2} &\approx \{0.41; -1.21\} \end{aligned}$$

The process is not invertible as one of the roots of the MA-part is outside the unit circle.

Q2.1.2: Second order moment representation

The mean is zero.

For the second order moment, consider the following linear equation system (p. 126-127):

$$\begin{aligned}
&\gamma(0) - 0.8\gamma(1) = \gamma_{\varepsilon X}(0) + 0.8\gamma_{\varepsilon X}(1) - 0.5\gamma_{\varepsilon X}(2)\\
&\gamma(1) - 0.8\gamma(0) = 0.8\gamma_{\varepsilon X}(0) - 0.5\gamma_{\varepsilon X}(1)\\
&\gamma(2) - 0.8\gamma(1) = -0.5\gamma_{\varepsilon X}(0)\\
&\gamma(3) - 0.8\gamma(2) = 0
\end{aligned}$$

Using equations (5.96) and (5.97) to calculate $\gamma_{\varepsilon X}(k)$:

$\gamma_{\varepsilon X}(0) = \sigma_\varepsilon^2$,
$\gamma_{\varepsilon X}(1) = (\theta_1 - \phi_1\theta_0)\cdot\sigma_\varepsilon^2 = 1.6\cdot\sigma_\varepsilon^2$ and
$\gamma_{\varepsilon X}(2) = (\theta_2 - \phi_1(\theta_1 - \phi_1\theta_0))\cdot\sigma_\varepsilon^2 = 0.78\cdot\sigma_\varepsilon^2$.

From the equations above we get

$$\begin{pmatrix} 1 & -0.8 \\ -0.8 & 1 \end{pmatrix}\begin{pmatrix}\gamma(0)\\ \gamma(1)\end{pmatrix} = \begin{pmatrix}\gamma_{\varepsilon X}(0) + 0.8\gamma_{\varepsilon X}(1) - 0.5\gamma_{\varepsilon X}(2)\\ 0.8\gamma_{\varepsilon X}(0) - 0.5\gamma_{\varepsilon X}(1)\end{pmatrix} = \begin{pmatrix}(1 + 0.8\cdot 1.6 - 0.5\cdot 0.78)\cdot 0.4^2\\ (0.8\cdot 1 - 0.5\cdot 1.6)\cdot 0.4^2\end{pmatrix} = \begin{pmatrix}0.3024\\ 0\end{pmatrix}$$

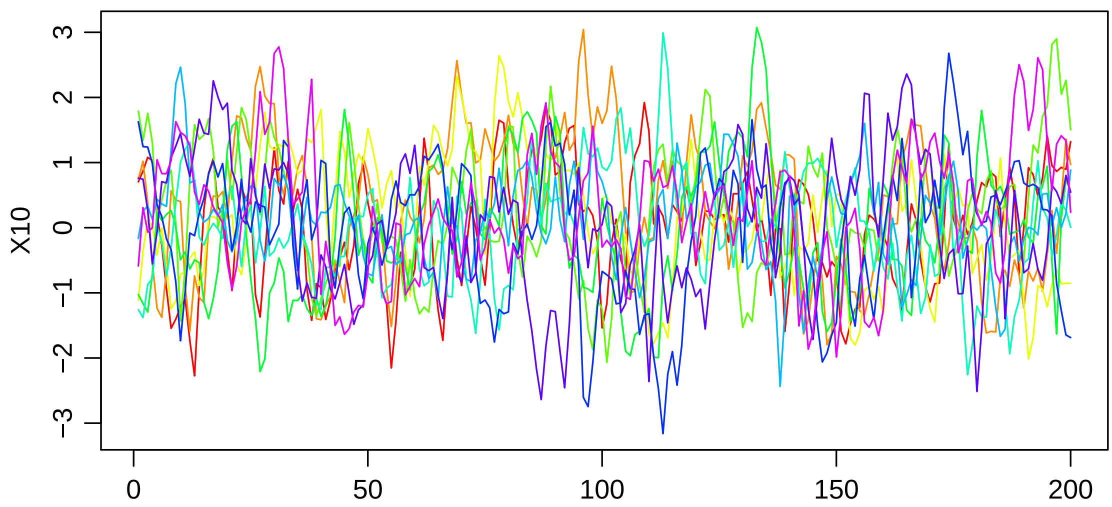
The solution is

$$\begin{pmatrix}\gamma(0)\\ \gamma(1)\end{pmatrix} = \begin{pmatrix}0.840\\ 0.672\end{pmatrix}$$

The auto-covariance is given by:

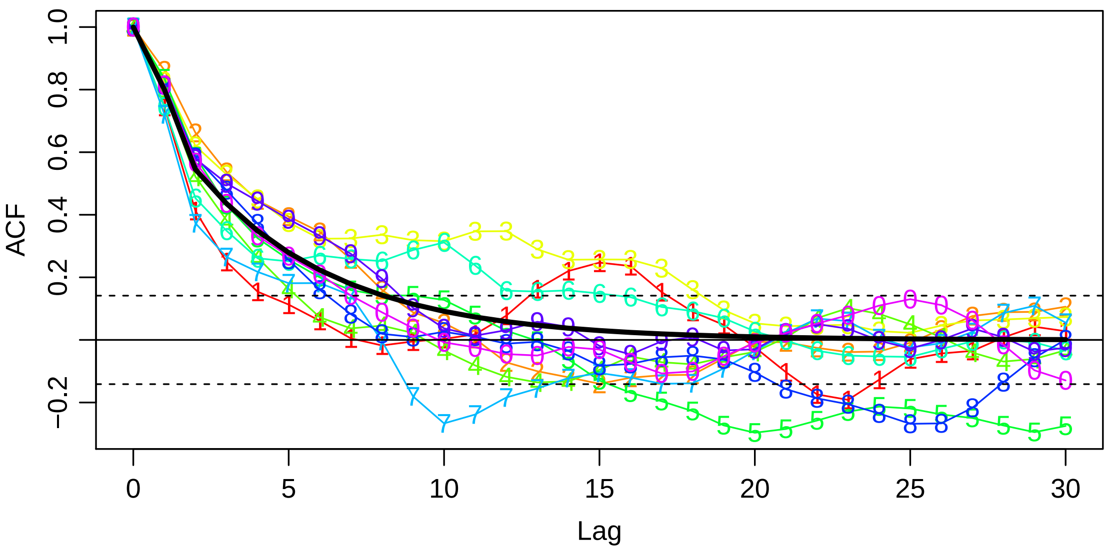
$$\begin{aligned}
\gamma(0) &= V[X_t] = 0.840 && (1)\\
\gamma(1) &= 0.672 && (2)\\
\gamma(2) &= -0.5\cdot 0.4^2 + 0.8\cdot 0.672 = 0.4576 && (3)\\
\gamma(k) &= 0.8\gamma(k-1) \quad for \;\; k > 2 && (4)
\end{aligned}$$

Q2.1.3: 10 realisations



The plot shows that the ten realisations are moving around on top of oneanother. This is in good agreement with the process being stationary.

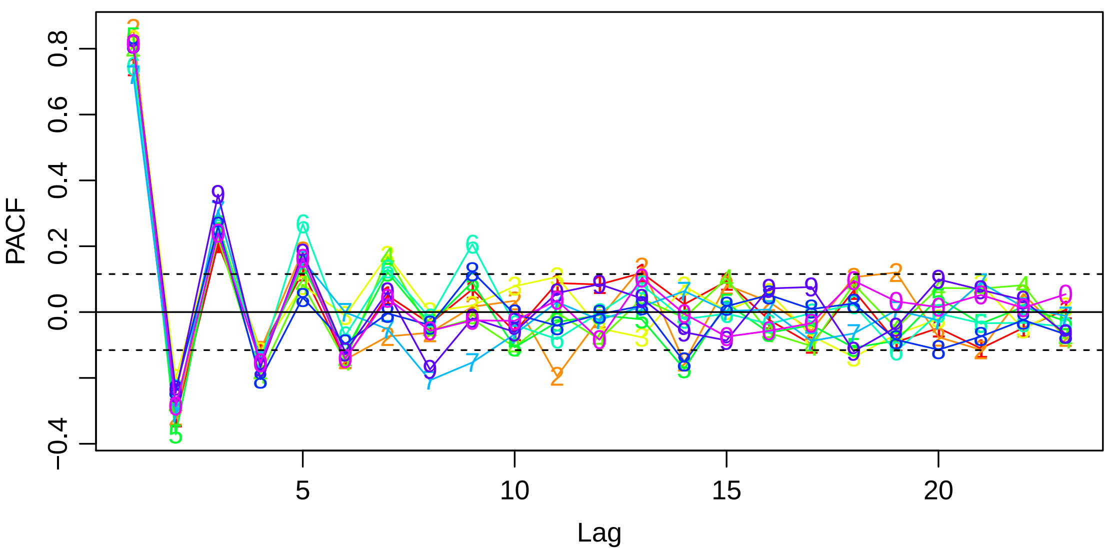
Q2.1.4: 10 realisations - ACF



First of all it is clearly seen that all the ACF drops slowly to zero so it is clear that there is an autoregressive part. The fact that it takes about 15 lags before it insignificant indicates that there is a root relatively close to one.

The thick black line is the theoretical acf.

Q2.1.5: 10 realisations - PACF



Again the behaviour is the same for all ten realisations. Here we see that the PACF changes sign for each lag and that the amplitude is decreasing. It is not possible to say from this if it is a damped harmonic or just alternating sign.

Q2.1.6: 10 realisations - variance

1	2	3	4	5	6	7	8	9	10
0.75	1.10	0.93	0.90	1.03	0.66	0.48	1.06	1.05	0.96

From the table is is clear that variance of the simulated realisations vary quite a bit. The mean of the ten variances is 0.8932563 which is quite close to the theoretical variance.

Q2.1.7: Discuss and compare

The theoretical ACF is plotted on top of the estimated ACF for each of the ten realisations. It is seen that the estimated ones follow the theoretical. *Above expected:* It is also seen that the estimated ACF and are themselves autocorrelated - which is not uncommon.

The variances of the simulations are close to the theoretical value.

Question 2.2: Predicting the the number of sales of apartments *The quarterly number of sales of apartments in the capital region of Denmark has been modelled.*

Based on historical data the following model has been identified:

$$(1 - 1.04B + 0.2B^2)(1 - 0.86B^4)(Y_t - \mu) = (1 - 0.42B^4)\epsilon_t$$

where ϵ_t is a white-noise process with variance σ_ϵ^2. Based on six days of data, it is found that $\sigma_\epsilon^2 = 36963$. Furthermore, μ was estimated to 2070. The file `A2_sales.txt` *contains the sales for the last five years of Y_t.*

Predict the values of Y_t corresponding to t = 2019Q1 and 2019Q2 (Two steps ahead), together with 95% prediction intervals for the predictions.

Plot the observations along with the two predictions and their 95% prediction intervals.

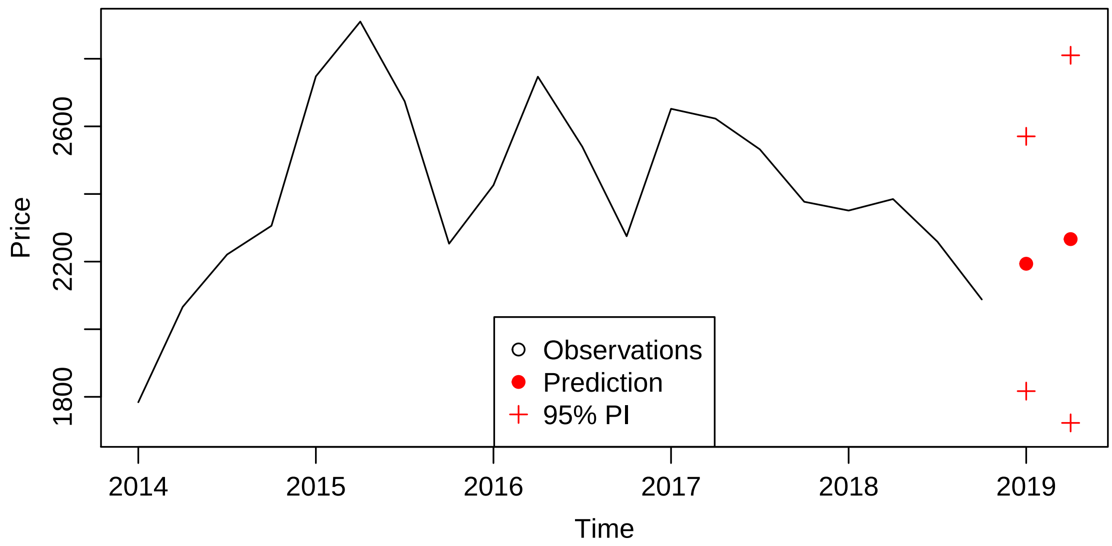
Q2.2: Predicting

You are not directly asked to check for stationarity and invertability but it is constructive to do so. *Count it as a constructive extra when included.*

The predictions for the next two observations including 95% prediction intervals (Using a normal distribution) are:

	Time	Prediction	sd	lower	upper
1	2019Q1	2193.66	192.26	1816.85	2570.48
2	2019Q2	2266.56	277.38	1722.90	2810.22

Plotting the data is always a good idea.



The prediction interval is quite wide and grows rather fast. The reason for the width is the high value of σ in this model. *Additional background: For the present data the value of* σ *would have been smaller if a shorter period had been used for fitting the data. Here data from 1992 and onwards are used.*

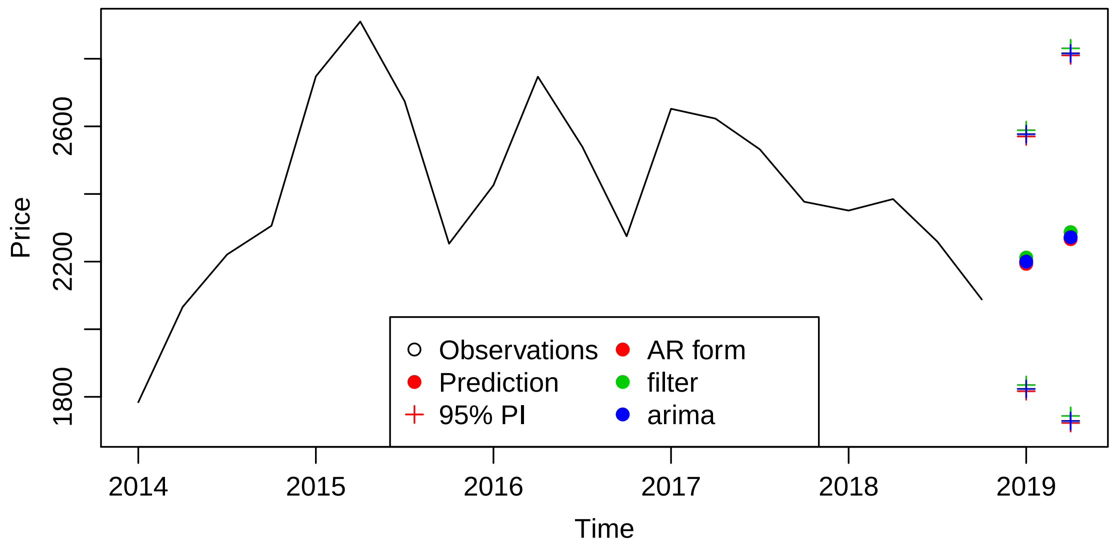
An alternative is filtering the provided samples. In theory the expected values can be found by filtering with the arima function with fixed parameters, however the variance will be different. It can also be done with your own filter assuming the first six epsilons to be zero.

	Time	OwnFilter	lower	upper
1	2019Q1	2211.98	1835.16	2588.80
2	2019Q2	2287.26	1743.59	2830.92

	Time	PredARIMA	lower	upper
1	2019Q1	2200.28	1823.46	2577.10
2	2019Q2	2272.45	1728.78	2816.11

The predictions in the above table are similar to the first values and the tendency is the same: the difference between one and two steps ahead is about 73. And there are small differences in the level - most probably due to the assumptions about the first six errors and older observations. The variances for the predictions are not included in the above formula as they are the same in all cases.

And plotting them all together it is seen that they generally agree well:



Question 2.3: *Consider the following ARMA(2,0) process*

$$\phi(B)X_t = X_t - 1.5X_{t-1} + \phi_2 X_{t-2} = \epsilon_t$$

Consider the four variations of the process with $\phi_2 \in \{0.52, 0.98\}$ and $\sigma^2 \in \{0.1^2, 5^2\}$. Simulate 300 observations of each of the four processes 100 times. Subsequently, estimate the model parameters based on the simulated sequences.

1. *For both values of ϕ_2 you should calculate the roots of $\phi(z^{-1}) = 0$*
2. *For each process, make a histogram plot of the estimates of parameter ϕ_2 and indicate the 95% quantiles.*
3. *How do different values of ϕ_2 affect the variance/distribution of the estimated ϕ_2?*
4. *How do different values of σ affect the variance/distribution of the estimated ϕ_2?*
5. *Plot all the estimated pairs of parameters (ϕ_1, ϕ_2) for the four variations. Comment on what you see and compare with the true values.*

Q2.3.1: Roots

```
(r1 <- polyroot(c(0.52, -1.5, 1)))

## [1] 0.5438447-0i 0.9561553+0i

(r2 <- polyroot(c(0.98, -1.5, 1)))

## [1] 0.75+0.6461424i 0.75-0.6461424i

Mod(r2)

## [1] 0.9899495 0.9899495
```

When $\phi_2 = 0.52$ the system has two real roots and the largest one is close to unity.

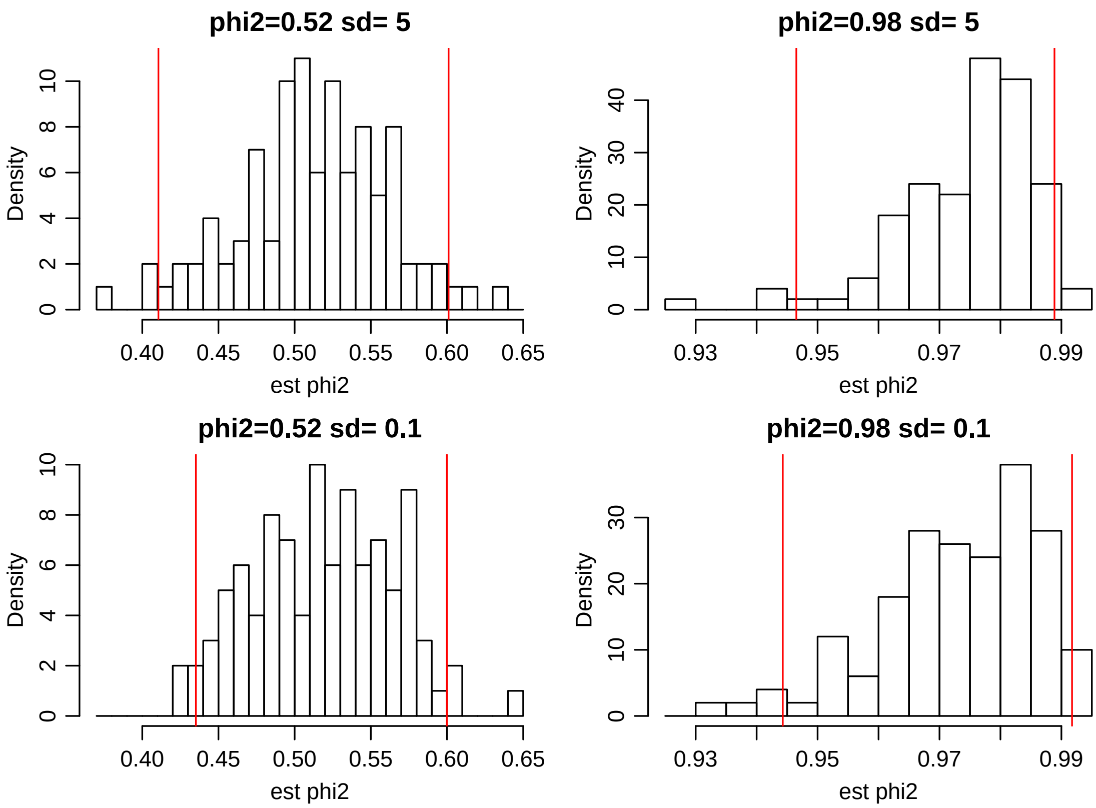
When $\phi_2 = 0.98$ the system has two complex conjugate roots. The modulus is quite close to unity.

The system is stationary for both values.

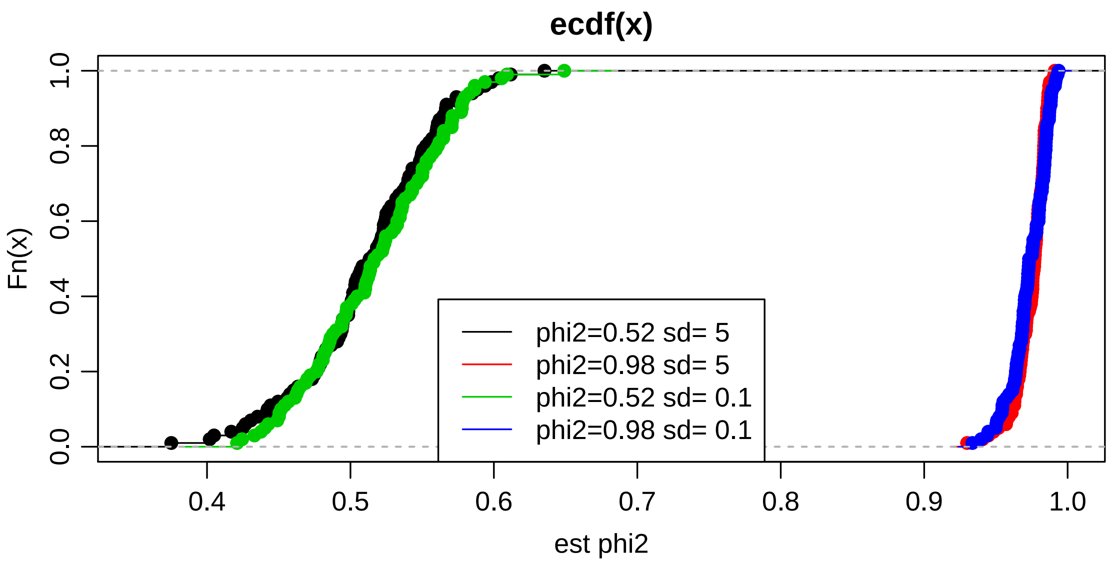
Q2.3.2-4: Histograms

It may be chosen to reset the seed before simulating each of the four combinations. This would result in identical ϵ's. It is OK if this is done.

First, plotting the requested four histograms and indicating the 2.5% and 97.5% quantiles in the distribution:

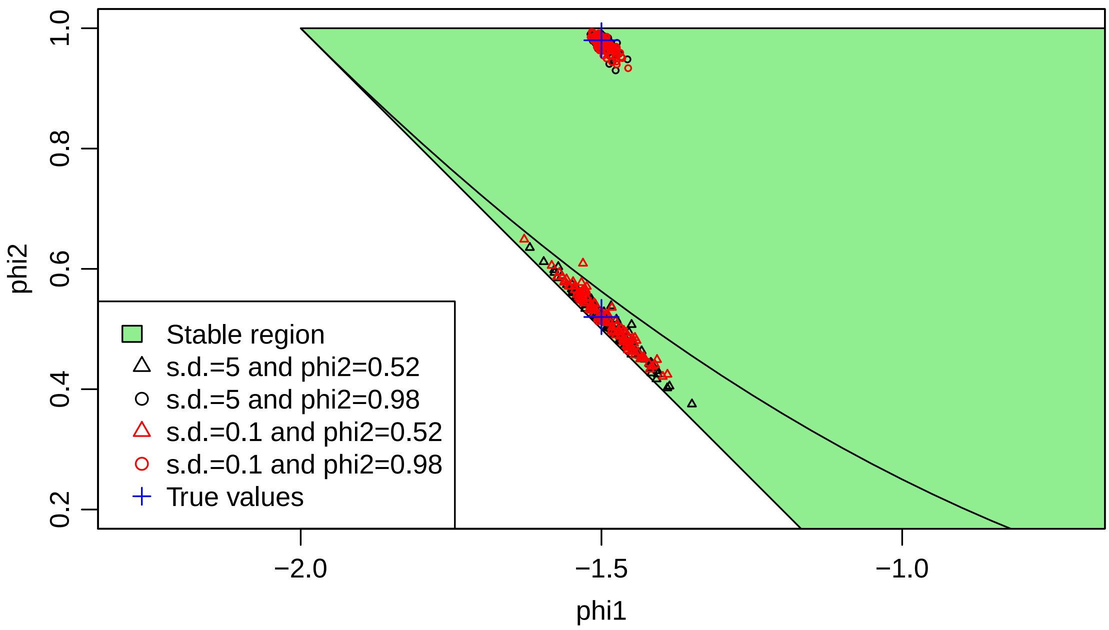


It is seen that there is almost no influence of the variance of the noise. (There will be no variation if the seed is reset). The reason for this is that this just scales the entire simulation. Another way (*Not expected*) to visuallize that the variance doesn't matter is to plot the emperical cumulated distribution function:



As expected there are groups for the different values of ϕ_2 but the value of σ doesn't make a difference.

Q2.3.5: Estimated pairs of parameters



The black line inside the stable region splits between real roots (below) and imaginary roots (above) Again it is seen that the variance in the simulation doesn't matter (red and black colors) and that the estimated values are near the true values. There are also differences: The variance is much smaller for $\phi_2 = 0.98$. It is noticed that all estimates are inside the stable region - this is a feature of the arima function in R (And OK if other packages does it different). The stability boundary is the reason why the estimates for $\phi_2 = 0.52$ are placed along a line as this process is stationary but doesn't oscillate.