Parameter Estimation, Model Diagnostics and Forecasting

YENNI ANGRAINI

Outline

- Parameter estimation method for AR, MA and ARMA :
  - The method of moments,
  - Least Square Estimation and
  - Maximum Likelihood
- Model Diagnostics : Residual Analysis and Overfitting
- Forecasting
- Illustration

Parameter estimation : The method of moments

- The method of moments is frequently one of the easiest
- The method consists of equating sample moments to corresponding theoretical moments and solving the resulting equations to obtain estimates of any unknown parameters

Parameter estimation :
The method of moments for **Autoregressive Models**

- Consider first the AR(1) case.
- For this process, we have the simple relationship $\rho_1 = \phi$.
- In the method of moments, ρ_1 is equated to r_1, the lag 1 sample autocorrelation.
- Thus we can estimate ϕ by $\hat{\phi} = r_1$

Parameter estimation : The method of moments for **Autoregressive Models**

Now consider the AR(2) case.

The relationships between the parameters ϕ_1 and ϕ_2 and various moments are given by the Yule-Walker equations

$$\rho_1 = \phi_1 + \rho_1\phi_2 \; and \; \rho_2 = \rho_1\phi_1 + \phi_2 \;\rightarrow r_1 = \phi_1 + r_1\phi_2 \; and \; r_2 = r_1\phi_1 + \phi_2$$

which are then solved to obtain $\hat{\phi}_1 = \frac{r_1(1-r_2)}{1-r_1^2} \; and \; \hat{\phi}_2 = \frac{r_2-r_1^2}{1-r_1^2}$

Parameter estimation :
The method of moments for **Autoregressive Models**

The general AR(p) case proceeds similarly. Replace ρ_k by r_k throughout the Yule-Walker equations to obtain

$$\left.\begin{array}{rrrcl} \phi_1 + & r_1\phi_2 + & r_2\phi_3 & + \cdots + r_{p-1}\phi_p & = r_1 \\ r_1\phi_1 + & \phi_2 + & r_1\phi_3 & + \cdots + r_{p-2}\phi_p & = r_2 \\ & & & & \vdots \\ r_{p-1}\phi_1 + & r_{p-2}\phi_2 + & r_{p-3}\phi_3 & + \cdots + \quad \phi_p & = r_p \end{array}\right\}$$

Parameter estimation : The method of moments for **Moving Average Models**

The method of moments is not nearly as convenient when applied to moving average models.

Consider the simple MA(1) case $\rho_1 = -\frac{\theta}{1+\theta^2}$, equating ρ_1 and r_1, we are led to solve a quadratic equation in θ.

$$\hat{\theta} = \frac{-1 + \sqrt{1 - 4r_1^2}}{2r_1}$$

For higher-order MA models, the method of moments quickly gets complicated

Parameter estimation :
The method of moments for **ARMA Models**

The ARMA(1,1) case

$$\rho_k = \frac{(1-\theta\phi)(\phi-\theta)}{1-2\theta\phi+\theta^2}\phi^{k-1} \quad \text{for } k \geq 1$$

Noting that $\rho_2/\rho_1 = \phi$, we can first estimate ϕ as

$$\hat{\phi} = \frac{r_2}{r_1}$$

Having done so, we can then use

$$r_1 = \frac{(1-\theta\hat{\phi})(\hat{\phi}-\theta)}{1-2\theta\hat{\phi}+\theta^2}$$

Note again that a quadratic equation must be solved.

Parameter estimation :
The method of moments for **the Noise Variance**

For the AR(p) models

$$\boxed{\hat{\sigma}_e^2 = (1 - \hat{\phi}_1 r_1 - \hat{\phi}_2 r_2 - \cdots - \hat{\phi}_p r_p) s^2}$$

$$\boxed{s^2 = \frac{1}{n-1} \sum_{t=1}^{n} (Y_t - \bar{Y})^2}$$

In particular, for an AR(1) process,

$$\boxed{\hat{\sigma}_e^2 = (1 - r_1^2) s^2}$$

since $\hat{\phi} = r_1$

Parameter estimation :
The method of moments for **the Noise Variance**

For the MA(q) case

$$\hat{\sigma}_e^2 = \frac{s^2}{1 + \hat{\theta}_1^2 + \hat{\theta}_2^2 + \cdots + \hat{\theta}_q^2}$$

$$s^2 = \frac{1}{n-1} \sum_{t=1}^{n} (Y_t - \bar{Y})^2$$

For the ARMA(1,1) process

$$\hat{\sigma}_e^2 = \frac{1 - \hat{\phi}^2}{1 - 2\hat{\phi}\hat{\theta} + \hat{\theta}^2} s^2$$

Numerical example for method of moments parameter estimation

	True Parameters			Method-of-Moments Estimates			
Model	θ	ϕ_1	ϕ_2	θ	ϕ_1	ϕ_2	***n***
MA(1)	–0.9			–0.554			120
MA(1)	0.9			0.719			120
MA(1)	–0.9			NA†			60
MA(1)	0.5			–0.314			60
AR(1)		0.9			0.831		60
AR(1)		0.4			0.470		60
AR(2)		1.5	–0.75		1.472	–0.767	120

† No method-of-moments estimate exists since $r_1 = 0.544$ for this simulation.

Parameter estimation :
Least Squares Estimation for **Autoregressive Models**

Consider the first-order case where

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + e_t$$

AR(1)

We can view this as a regression model with predictor variable Y_{t-1} and response variable Y_t. Least squares estimation then proceeds by minimizing the sum of squares of the differences

$$(Y_t - \mu) - \phi(Y_{t-1} - \mu)$$

Since only $Y_1, Y_2, \ldots, Y_n$ are observed, we can only sum from $t = 2$ to $t = n$. Let

$$S_c(\phi, \mu) = \sum_{t=2}^{n} [(Y_t - \mu) - \phi(Y_{t-1} - \mu)]^2$$

Conditional sum-of squares function

Parameter estimation :
Least Squares Estimation for **Autoregressive Models**

Consider the equation $\partial S_c / \partial \mu = 0$. We have

$$\frac{\partial S_c}{\partial \mu} = \sum_{t=2}^{n} 2[(Y_t - \mu) - \phi(Y_{t-1} - \mu)](-1 + \phi) = 0$$

or, simplifying and solving for μ,

$$\boxed{\mu = \frac{1}{(n-1)(1-\phi)}\left[\sum_{t=2}^{n} Y_t - \phi \sum_{t=2}^{n} Y_{t-1}\right]}$$

Parameter estimation :
Least Squares Estimation for **Autoregressive Models**

Now, for large n,

$$\frac{1}{n-1}\sum_{t=2}^{n} Y_t \approx \frac{1}{n-1}\sum_{t=2}^{n} Y_{t-1} \approx \bar{Y}$$

$$\hat{\mu} \approx \frac{1}{1-\phi}(\bar{Y} - \phi\bar{Y}) = \bar{Y}$$

We sometimes say, except for end effects, $\hat{\mu} = \bar{Y}$.

Parameter estimation : Least Squares Estimation for **Autoregressive Models**

Consider now the minimization of $S_c(\phi, \bar{Y})$ with respect to ϕ. We have

$$\frac{\partial S_c(\phi, \bar{Y})}{\partial \phi} = \sum_{t=2}^{n} 2[(Y_t - \bar{Y}) - \phi(Y_{t-1} - \bar{Y})](Y_{t-1} - \bar{Y})$$

Setting this equal to zero and solving for ϕ yields

$$\boxed{\hat{\phi} = \frac{\sum_{t=2}^{n} (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=2}^{n} (Y_{t-1} - \bar{Y})^2}}$$

Parameter estimation :
Least Squares Estimation for **Moving Average Models**

Consider now the least-squares estimation of θ in the MA(1) model:

$$Y_t = e_t - \theta e_{t-1}$$

$$Y_t = -\theta Y_{t-1} - \theta^2 Y_{t-2} - \theta^3 Y_{t-3} - \cdots + e_t$$

an autoregressive model but of infinite order. Thus least squares can be meaningfully carried out by choosing a value of θ that minimizes

Parameter estimation :
Least Squares Estimation for **Moving Average Models**

$$S_c(\theta) = \sum(e_t)^2 = \sum[Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} + \theta^3 Y_{t-3} + \cdots]^2$$

where, implicitly, $e_t = e_t(\theta)$ is a function of the observed series and the unknown parameter θ.

- It is clear from this equation that the least squares problem is *nonlinear* in the parameters.
- We will not be able to minimize $Sc(\theta)$ by taking a derivative with respect to θ, setting it to zero, and solving→ numerical optimization

Parameter estimation : Maximum Likelihood

- For any set of observations, Y_1, Y_2,..., Y_n, time series or not, the likelihood function L is defined to be the joint probability density of obtaining the data actually observed
- It is considered as a function of the unknown parameters in the model with the observed data held fixed
- For ARIMA models, L will be a function of the ϕ's, θ's, μ, and σ_e^2 given the observations Y_1, Y_2,..., Y_n.
- The maximum likelihood estimators are then defined as those values of the parameters for which the data actually observed are *most likely*, that is, the values that maximize the likelihood function.

Parameter estimation : Maximum Likelihood AR(1)

The likelihood function for an AR(1) model is given by

$$L(\phi, \mu, \sigma_e^2) = (2\pi\sigma_e^2)^{-\frac{n}{2}}(1-\phi^2)^{1/2} exp\left[-\frac{1}{2\sigma_e^2} S(\phi, \mu)\right]$$

Where

$$S(\phi, \mu) = \sum_{t=2}^{n} [(Y_t - \mu) - \phi(Y_{t-1} - \mu)]^2 + (1-\phi^2)(Y_1 - \mu)$$

Unconditional sum-of squares function

The log-likelihood function for an AR(1) model is given by

$$l(\phi, \mu, \sigma_e^2) = -\frac{n}{2}\log(2\pi) - -\frac{n}{2}\log(\sigma_e^2) + \frac{1}{2}\log(1-\phi^2) - \frac{1}{2\sigma_e^2} S(\phi, \mu)$$

Properties of the Estimates

For large n, the estimators are approximately unbiased and normally distributed. The variances and correlations are as follows:

$$\text{AR(1)}: \; Var(\hat{\phi}) \approx \frac{1-\phi^2}{n}$$

$$\text{AR(2)}: \begin{cases} Var(\hat{\phi}_1) \approx Var(\hat{\phi}_2) \approx \dfrac{1-\phi_2^2}{n} \\ Corr(\hat{\phi}_1, \hat{\phi}_2) \approx -\dfrac{\phi_1}{1-\phi_2} = -\rho_1 \end{cases}$$

$$\text{MA(1)}: \; Var(\hat{\theta}) \approx \frac{1-\theta^2}{n}$$

$$\text{MA(2)}: \begin{cases} Var(\hat{\theta}_1) \approx Var(\hat{\theta}_2) \approx \dfrac{1-\theta_2^2}{n} \\ Corr(\hat{\theta}_1, \hat{\theta}_2) \approx -\dfrac{\theta_1}{1-\theta_2} \end{cases}$$

$$\text{ARMA(1,1)}: \begin{cases} Var(\hat{\phi}) \approx \left[\dfrac{1-\phi^2}{n}\right]\left[\dfrac{1-\phi\theta}{\phi-\theta}\right]^2 \\ Var(\hat{\theta}) \approx \left[\dfrac{1-\theta^2}{n}\right]\left[\dfrac{1-\phi\theta}{\phi-\theta}\right]^2 \\ Corr(\hat{\phi}, \hat{\theta}) \approx \dfrac{\sqrt{(1-\phi^2)(1-\theta^2)}}{1-\phi\theta} \end{cases}$$

Illustrations of Parameter Estimation

Parameter Estimation for Simulated AR(1) Models

Parameter ϕ	Method-of-Moments Estimate	Conditional SS Estimate	Unconditional SS Estimate	Maximum Likelihood Estimate	n
0.9	0.831	0.857	0.911	0.892	60
0.4	0.470	0.473	0.473	0.465	60

$$\sqrt{\hat{Var}(\hat{\phi})} \approx \sqrt{\frac{1-\hat{\phi}^2}{n}} = \sqrt{\frac{1-(0.831)^2}{60}} \approx 0.07$$

$$\sqrt{\hat{Var}(\hat{\phi})} = \sqrt{\frac{1-(0.470)^2}{60}} \approx 0.11$$

Illustrations of Parameter Estimation

Parameter Estimation for a Simulated AR(2) Model

Parameters	Method-of-Moments Estimates	Conditional SS Estimates	Unconditional SS Estimates	Maximum Likelihood Estimate	n
$\phi_1 = 1.5$	1.472	1.5137	1.5183	1.5061	120
$\phi_2 = -0.75$	–0.767	–0.8050	–0.8093	–0.7965	120

$$\sqrt{\hat{Var}(\hat{\phi}_1)} \approx \sqrt{\hat{Var}(\hat{\phi}_2)} \approx \sqrt{\frac{1-\phi_2^2}{n}} = \sqrt{\frac{1-(0.75)^2}{120}} \approx 0.06$$

Illustrations of Parameter Estimation

Parameter Estimation for a Simulated ARMA(1,1) Model

Parameters	Method-of-Moments Estimates	Conditional SS Estimates	Unconditional SS Estimates	Maximum Likelihood Estimate	n
$\phi = 0.6$	0.637	0.5586	0.5691	0.5647	100
$\theta = -0.3$	−0.2066	−0.3669	−0.3618	−0.3557	100

Illustrations of Parameter Estimation

Parameter Estimation for the Color Property Series

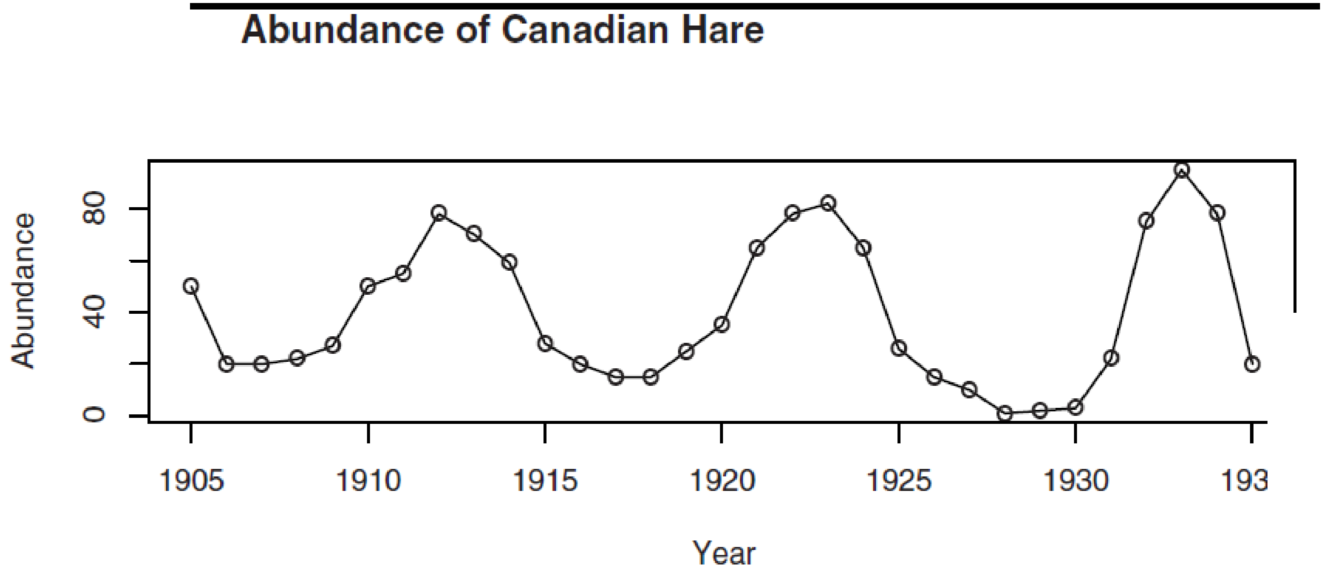
Parameter	Method-of-Moments Estimate	Conditional SS Estimate	Unconditional SS Estimate	Maximum Likelihood Estimate	n
ϕ	0.5282	0.5549	0.5890	0.5703	35

the standard error of the estimates is about

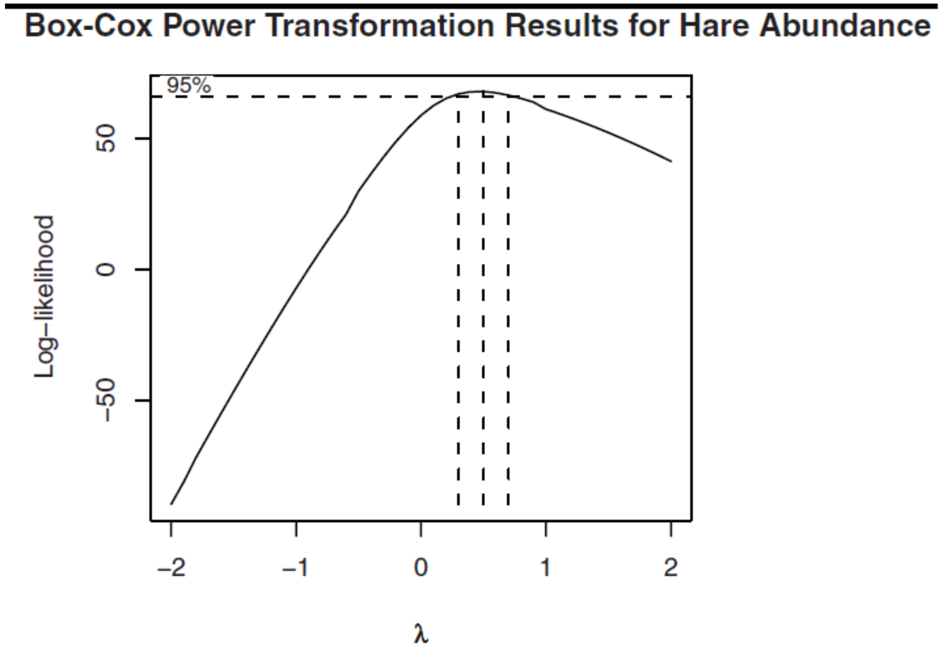
$$\sqrt{\hat{Var}(\hat{\phi})} \approx \sqrt{\frac{1-(0.57)^2}{35}} \approx 0.14$$

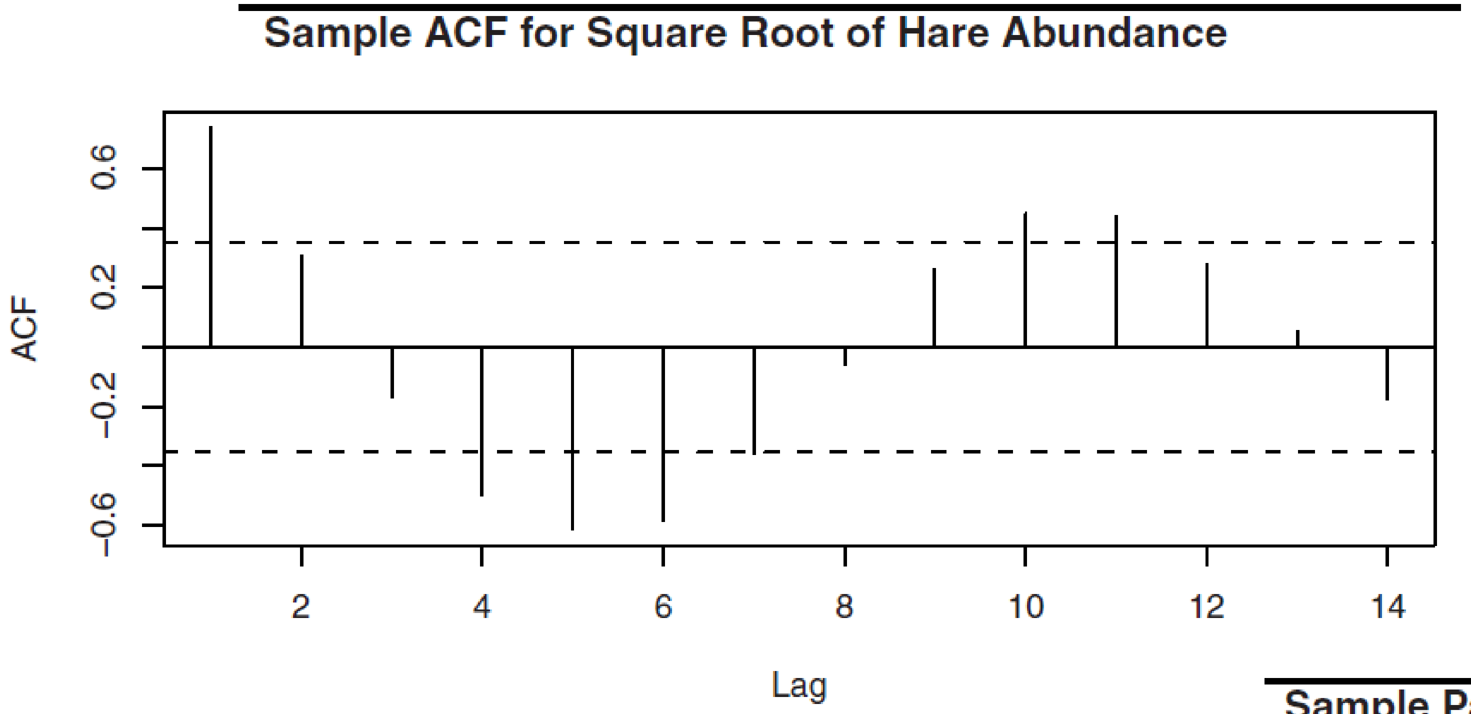
Illustrations of Parameter Estimation

The Annual Abundance of Canadian Hare Series

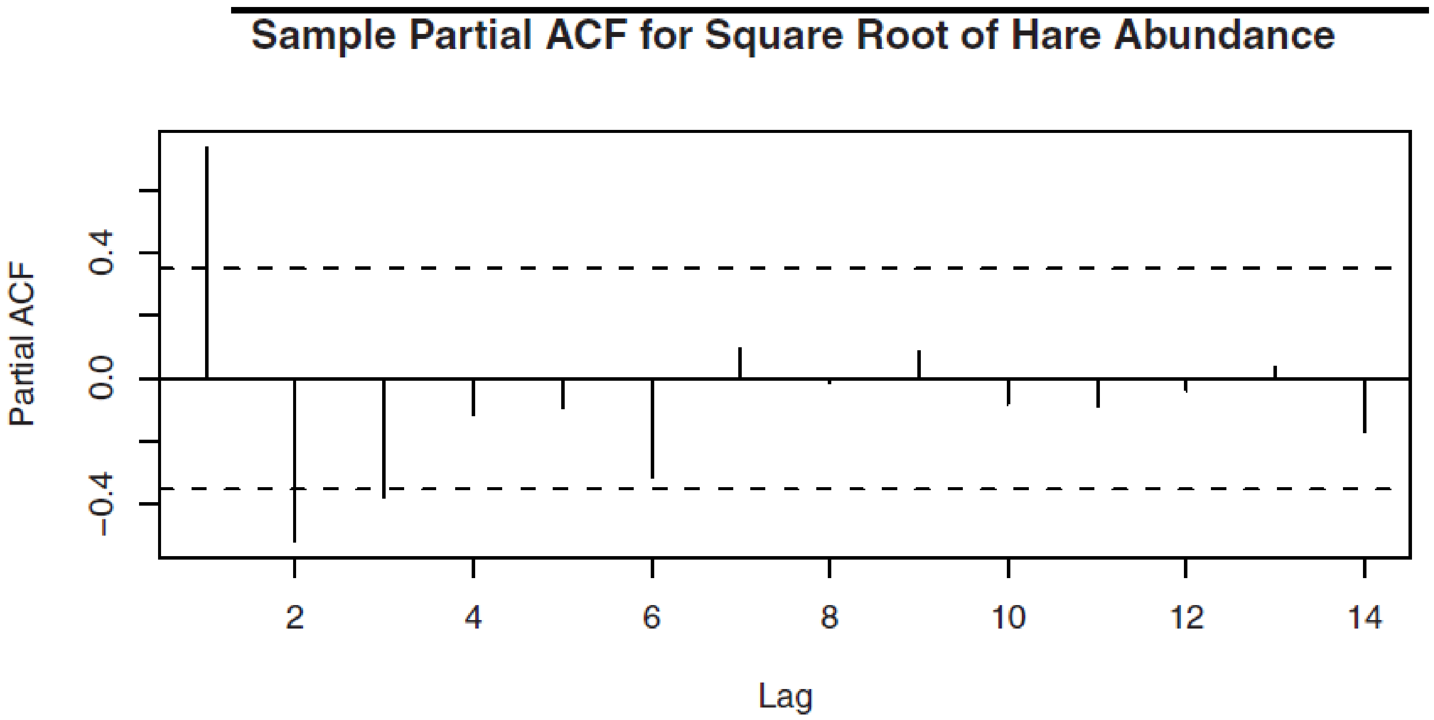


λ	Transformed Data
-2	y^{-2}
-1	y^{-1}
-0.5	$1/\sqrt{y}$
0	$\ln(y)$
$\boxed{0.5}$	$\boxed{\sqrt{y}}$
1	y
2	y^2





- → ACF tails off
- → PACF cuts off after lag 3



Maximum Likelihood Estimates from R Software: Hare Series

Coefficients:	ar1	ar2	ar3	Intercept[†]
	1.0519	–0.2292	–0.3931	5.6923
s.e.	0.1877	0.2942	0.1915	0.3371

```
sigma^2 estimated as 1.066: log-likelihood = -46.54, AIC = 101.08
```

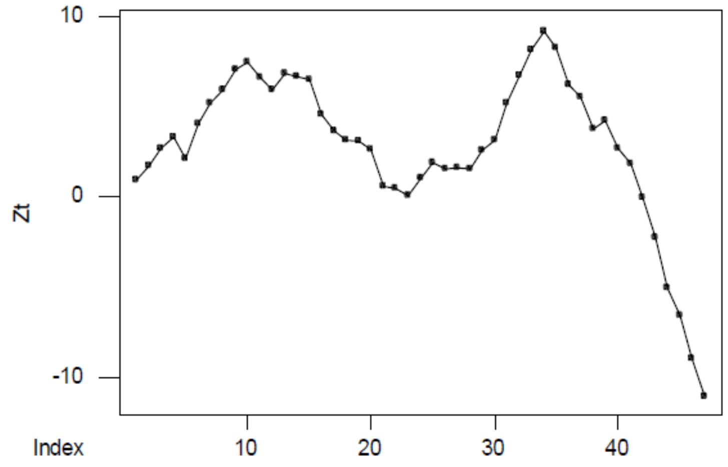
[†] The intercept here is the estimate of the process mean μ—not of θ_0.

$$t_{ar1} = \frac{1.0519}{0.1877} = 5.604$$

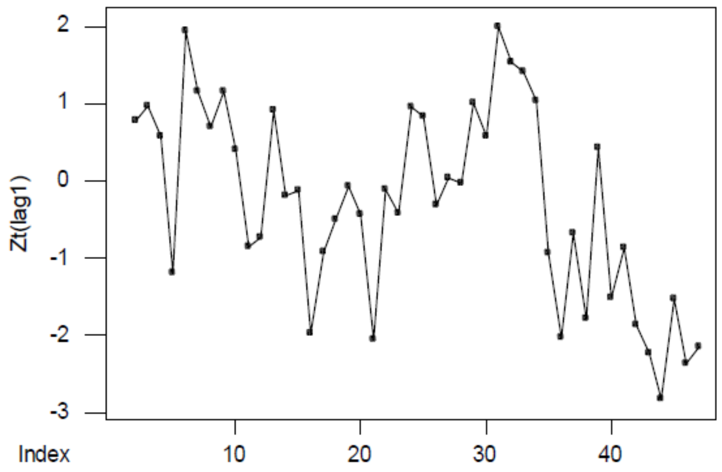
$$\boxed{t_{ar2} = -\frac{0.2292}{0.2941} = -0.779}$$

$$t_{ar3} = -\frac{0.3931}{0.1915} = -2.053$$

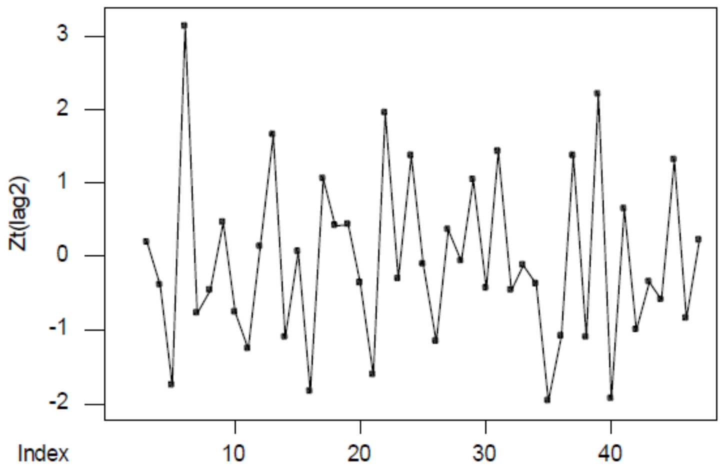
Illustrations of Parameter Estimation



Zt : Data Asal

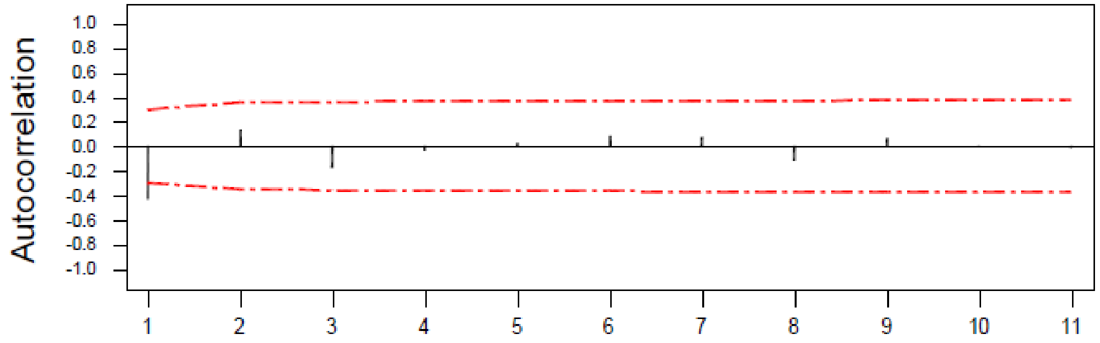


Zt : Data Setelah Differencing Ordo-1

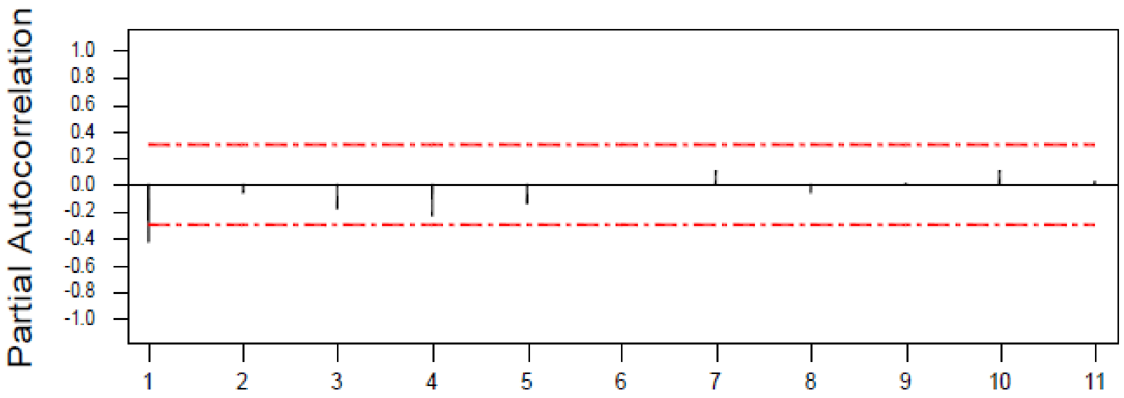


Zt : Data Setelah Differencing Ordo-2

Autocorrelation Function for Zt(lag2)



Partial Autocorrelation Function for Zt(lag2)



ARIMA(0,2,1) & **ARIMA(1,2,0)**

ARIMA(0, 2, 1)

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	-0.4393	0.1371	-3.20	$\boxed{0.003}$
Constant	-0.0995	0.1581	-0.63	0.533

ARIMA(1, 2, 0)

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	0.5958	0.1225	4.86	$\boxed{0.000}$
Constant	-0.06673	0.06299	-1.06	0.295


THANK YOU
MERCI
OBRIGADO
ASANTE
СПАСИБО
SPASIBO
WELALIN
GRAZIE
CHOKRANE
RAIBH MAITH AGAT
MULTUMESC
NIRRINGRAZZJAK
MOCHCHAKKERAM
MATUR NUWUN
KHUBLEI
VINAKA
ZIKOMO
MATONDO
SUWUN
EIO
THANK YOU VERY MUCH
SAAGHA XALDA HWA
どうも
MAAKE
МНОГО ВИ БЛАГОДАРЯ
БАЯРЛАЛАА
QUYANAQ
iStock by Getty Images