Econometric Time Series Analysis|EC5221

Wk2: Stationary Univariate Time Series - Estimation and Inference

Nicky Grant (Semester 2, 2020/2021)

Table of contents

Recap

Distinction between **process** $\{Y_t\}$ and **sample realisation** $\{y_t\}$

White noise & **stationarity**

Measuring statistical dependence- **auto-correlation function [ACF]**

Moments and stationarity conditions of **ARMA** processes

Focused on theoretical properties of ARMA processes assuming model parameters **known**

What We Will Cover

Sample moments & asymptotic properties

Quick refresher on convergence in probability and distribution

Maximum likelihood estimation of ARMA models

Sample Moments

Estimating Population Moments

Distribution of process generating sample data $\{Y_t : t = 1, .., T\}$ **unknown**

Estimating Population Moments

Distribution of process generating sample data $\{Y_t : t = 1, .., T\}$ **unknown**

$\Rightarrow$ population moments of $\{Y_t\}$ **unknown** (e.g. mean, variance, auto-covariances)

Estimating Population Moments

Distribution of process generating sample data $\{Y_t : t = 1, .., T\}$ **unknown**

$\Rightarrow$ population moments of $\{Y_t\}$ **unknown** (e.g. mean, variance, auto-covariances)

Construct **estimators** of population moments using realisations from $\{Y_t : t = 1, .., T\}$

Estimating Population Moments

Distribution of process generating sample data $\{Y_t : t = 1, .., T\}$ **unknown**

$\Rightarrow$ population moments of $\{Y_t\}$ **unknown** (e.g. mean, variance, auto-covariances)

Construct **estimators** of population moments using realisations from $\{Y_t : t = 1, .., T\}$

E.g. How to form a consistent estimator of the population mean $\mathbb{E}[Y_t]$?

Estimating Population Moments

Distribution of process generating sample data $\{Y_t : t = 1, .., T\}$ **unknown**

$\Rightarrow$ population moments of $\{Y_t\}$ **unknown** (e.g. mean, variance, auto-covariances)

Construct **estimators** of population moments using realisations from $\{Y_t : t = 1, .., T\}$

E.g. How to form a consistent estimator of the population mean $\mathbb{E}[Y_t]$?

i.e. an estimator **converging in probability** to $\mathbb{E}[Y_t]$ as $T \to \infty$

Refresher: Convergence in Probability and Distribution

Definition: Convergence in Probability

A stochastic sequence X_T on $T = 1, 2, \cdots$ converges in probability to X if $\forall\ \epsilon > 0$

$$\Pr\{|X_T - X| < \epsilon\} \underset{T \to \infty}{\to} 1$$

Refresher: Convergence in Probability and Distribution

Definition: Convergence in Probability

A stochastic sequence X_T on $T = 1, 2, \cdots$ converges in probability to X if $\forall\ \epsilon > 0$

$$\Pr\{|X_T - X| < \epsilon\} \underset{T \to \infty}{\to} 1$$

Definition: Convergence in Distribution

A stochastic sequence X_T on $T = 1, 2, \cdots$ converges in distribution to X when

$$F_{X_T}(x) \underset{T \to \infty}{\to} F_X(x) \quad \text{for x where } F_X(x) \text{ is continuous}$$

Refresher: Convergence in Probability and Distribution

Definition: Convergence in Probability

A stochastic sequence X_T on $T = 1, 2, \cdots$ converges in probability to X if $\forall\ \epsilon > 0$

$$\Pr\{|X_T - X| < \epsilon\} \underset{T \to \infty}{\to} 1$$

Definition: Convergence in Distribution

A stochastic sequence X_T on $T = 1, 2, \cdots$ converges in distribution to X when

$$F_{X_T}(x) \underset{T \to \infty}{\to} F_X(x) \quad \text{for x where } F_X(x) \text{ is continuous}$$

Shorthand: $X_T \xrightarrow{p} X$ and $X_T \xrightarrow{d} X$ respectively

Time Averages

Economics time series data contain **one realisation** from the unknown data generating process

Time Averages

Economics time series data contain **one realisation** from the unknown data generating process

Only sample data $(y_1 \cdots, y_T)'$ is observed

Time Averages

Economics time series data contain **one realisation** from the unknown data generating process

Only sample data $(y_1 \cdots , y_T)'$ is observed

Can constructs estimators of population moments by **time averages** based on sample realisation

Time Averages

Economics time series data contain **one realisation** from the unknown data generating process

Only sample data $(y_1 \cdots, y_T)'$ is observed

Can constructs estimators of population moments by **time averages** based on sample realisation

A crucial assumption for consistency is the process generating the sample is stationary

Sample Mean - Asymptotic Properties

Time average in sample $\bar{y}_T := \frac{1}{T}\sum_{t=1}^{T} y_t$ drawn from the r.v. $\bar{Y}_T := \frac{1}{T}\sum_{t=1}^{T} Y_t$

Sample Mean - Asymptotic Properties

Time average in sample $\bar{y}_T := \frac{1}{T}\sum_{t=1}^{T} y_t$ drawn from the r.v. $\bar{Y}_T := \frac{1}{T}\sum_{t=1}^{T} Y_t$

Define $\mathbb{E}[Y_t] = \mu$

Sample Mean - Asymptotic Properties

Time average in sample $\bar{y}_T := \frac{1}{T}\sum_{t=1}^{T} y_t$ drawn from the r.v. $\bar{Y}_T := \frac{1}{T}\sum_{t=1}^{T} Y_t$

Define $\mathbb{E}[Y_t] = \mu$

Theorem 1.5: Consistency of Sample Mean for Serial Dependent Processes

Any weakly-stationary process $\{Y_t\}$ with absolute summable covariances

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

Exercise Prove the above result [*Hint: use Chebychev's Inequality*]

Sample Mean - Asymptotic Properties

Time average in sample $\bar{y}_T := \frac{1}{T}\sum_{t=1}^{T} y_t$ drawn from the r.v. $\bar{Y}_T := \frac{1}{T}\sum_{t=1}^{T} Y_t$

Define $\mathbb{E}[Y_t] = \mu$

Theorem 1.7: Consistency of Sample Mean for Serial Dependent Processes

Any weakly-stationary process $\{Y_t\}$ with absolute summable covariances

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

Exercise Prove the above result [*Hint: use Chebychev's Inequality*]

Theorem 1.8: Central Limit Theorem for Serially Dependent Processes

If $Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ where $\{\varepsilon_t\}$ is i.i.d, $\mathbb{E}[\varepsilon_t] = 0$, $\mathbb{E}[\varepsilon_t^2] < \infty$ and $\sum_{j=0}^{\infty} |\psi_j| < \infty$

$$\sqrt{T}(\bar{Y}_T - \mu) \xrightarrow{d} N\left(0, \sum_{j=-\infty}^{\infty} \gamma(j)\right)$$

Sample Second Order Moments

Definition: Sample Auto-Covariance Function

$$\hat{\gamma}_T(k) = \frac{1}{T-k} \sum_{t=k+1}^{T} \left(Y_t - \bar{Y}_T\right)\left(Y_{t-k} - \bar{Y}_T\right) \qquad \forall \;\; 0 \leq k < T$$

Sample Second Order Moments

Definition: Sample Auto-Covariance Function

$$\hat{\gamma}_T(k) = \frac{1}{T-k} \sum_{t=k+1}^{T} \left(Y_t - \bar{Y}_T\right)\left(Y_{t-k} - \bar{Y}_T\right) \qquad \forall\ 0 \leq k < T$$

Definition: Sample Auto-Correlation Function

$$\hat{\rho}_T(k) = \frac{\hat{\gamma}_T(k)}{\hat{\gamma}_T(0)} \qquad \forall\ 0 \leq k < T$$

Sample Second Order Moments

Definition: Sample Auto-Covariance Function

$$\hat{\gamma}_T(k) = \frac{1}{T-k} \sum_{t=k+1}^{T} \left(Y_t - \bar{Y}_T\right)\left(Y_{t-k} - \bar{Y}_T\right) \qquad \forall \; 0 \leq k < T$$

Definition: Sample Auto-Correlation Function

$$\hat{\rho}_T(k) = \frac{\hat{\gamma}_T(k)}{\hat{\gamma}_T(0)} \qquad \forall \; 0 \leq k < T$$

$\hat{\gamma}_T(k)$ and $\hat{\rho}_T(k)$ evaluated at sample realisation $(y_1, \cdots, y_T)'$ to form **sample estimate**

Asymptotic Properties of Second Order Sample Moments

For linear processes $Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ where $\{\varepsilon_t\}$ is i.i.d and $\sum_{j=0}^{\infty} |\psi_j| < \infty$

Asymptotic Properties of Second Order Sample Moments

For linear processes $Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ where $\{\varepsilon_t\}$ is i.i.d and $\sum_{j=0}^{\infty} |\psi_j| < \infty$

Theorem 1.11: Consistency of Sample Auto-covariances for Stationary Processes

If $E[|\varepsilon_t|^r] < \infty$ for some $r > 2$

$$\hat{\gamma}_T(k) \xrightarrow{p} \gamma(k)$$

Asymptotic Properties of Second Order Sample Moments

For linear processes $Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ where $\{\varepsilon_t\}$ is i.i.d and $\sum_{j=0}^{\infty} |\psi_j| < \infty$

Theorem 1.13: Consistency of Sample Auto-covariances for Stationary Processes

If $E[|\varepsilon_t|^r] < \infty$ for some $r > 2$

$$\hat{\gamma}_T(k) \xrightarrow{p} \gamma(k)$$

Proven in **H** pp.192-193 **[Not Examinable]**

Asymptotic Properties of Second Order Sample Moments

For linear processes $Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ where $\{\varepsilon_t\}$ is i.i.d and $\sum_{j=0}^{\infty} |\psi_j| < \infty$

Theorem 1.15: Consistency of Sample Auto-covariances for Stationary Processes

If $E[|\varepsilon_t|^r] < \infty$ for some $r > 2$

$$\hat{\gamma}_T(k) \xrightarrow{p} \gamma(k)$$

Proven in **H** pp.192-193 **[Not Examinable]**

Theorem 1.16: Limit Distn. of Sample Auto-covariances for Stationary Processes

If $E[\varepsilon_t^4] < \infty$

$$\sqrt{T}(\hat{\gamma}_T(k) - \gamma(k)) \xrightarrow{d} N\left(0, \sum_{s=0}^{\infty} (\rho(s+k) + \rho(s-k) - \rho(s)\rho(k))^2\right)$$

Asymptotic Properties of Second Order Sample Moments

For linear processes $Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ where $\{\varepsilon_t\}$ is i.i.d and $\sum_{j=0}^{\infty} |\psi_j| < \infty$

Theorem 1.17: Consistency of Sample Auto-covariances for Stationary Processes

If $E[|\varepsilon_t|^r] < \infty$ for some $r > 2$

$$\hat{\gamma}_T(k) \xrightarrow{p} \gamma(k)$$

Proven in **H** pp.192-193 **[Not Examinable]**

Theorem 1.18: Limit Distn. of Sample Auto-covariances for Stationary Processes

If $E[\varepsilon_t^4] < \infty$

$$\sqrt{T}(\hat{\gamma}_T(k) - \gamma(k)) \xrightarrow{d} N\left(0, \sum_{s=0}^{\infty} (\rho(s+k) + \rho(s-k) - \rho(s)\rho(k))^2\right)$$

We do **not** prove Theorem 1.18

Ergodicity

Definition: Ergodicity (mean)

A stationary process is *ergodic for the mean* if $\frac{1}{T}\sum_{t=1}^{T} Y_t \xrightarrow{p} \mu$

Ergodicity

Definition: Ergodicity (mean)

A stationary process is *ergodic for the mean* if $\frac{1}{T}\sum_{t=1}^{T} Y_t \xrightarrow{p} \mu$

Stationary process ergodic for mean if $\sum_{j=0}^{\infty} |\gamma(j)| < \infty$ (follows from Theorem 1.7)

Definition: Ergodicity (second moments)

A stationary process is *ergodic for the second moment* if $\hat{\gamma}_T(k) \xrightarrow{p} \gamma(k)\ \forall$ k

Ergodicity

Definition: Ergodicity (mean)

A stationary process is *ergodic for the mean* if $\frac{1}{T}\sum_{t=1}^{T} Y_t \xrightarrow{p} \mu$

Stationary process ergodic for mean if $\sum_{j=0}^{\infty} |\gamma(j)| < \infty$ (follows from Theorem 1.7)

Definition: Ergodicity (second moments)

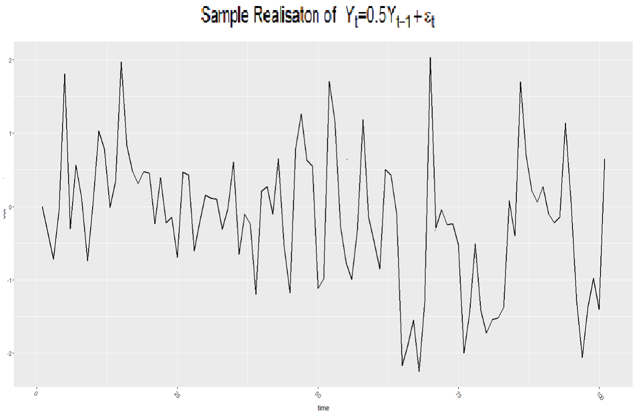
A stationary process is *ergodic for the second moment* if $\hat{\gamma}_T(k) \xrightarrow{p} \gamma(k)\ \forall$ k

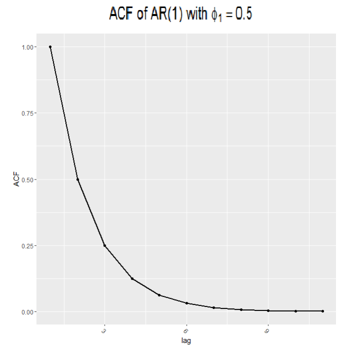
Ergodicity & stationarity closely linked- but a process can be stationary & not ergodic (see **H** p. 47)

Statistical Properties of Sample ACF

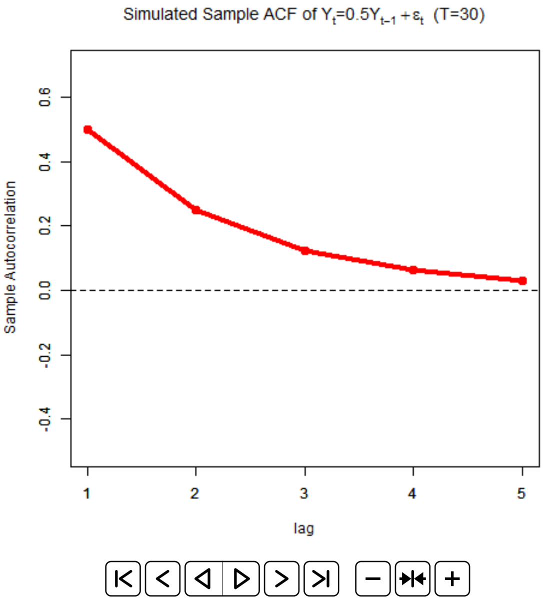
Simulation Example: Properties of Sample Autocorrelation of AR(1)

Suppose the true process is $Y_t = 0.5Y_{t-1} + \varepsilon_t \quad \varepsilon_t \overset{i.i.d}{\sim}$ N(0,1)





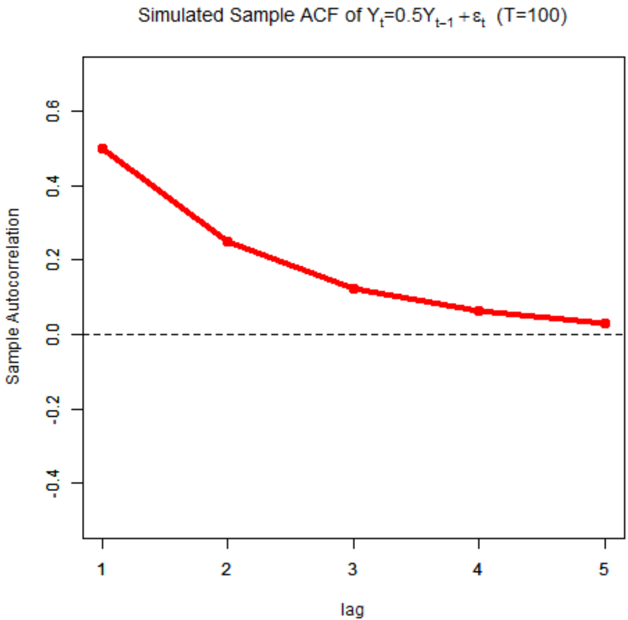
By Theorems 1.7 $\hat{\rho}_T(k) \xrightarrow{p} \rho(k)$



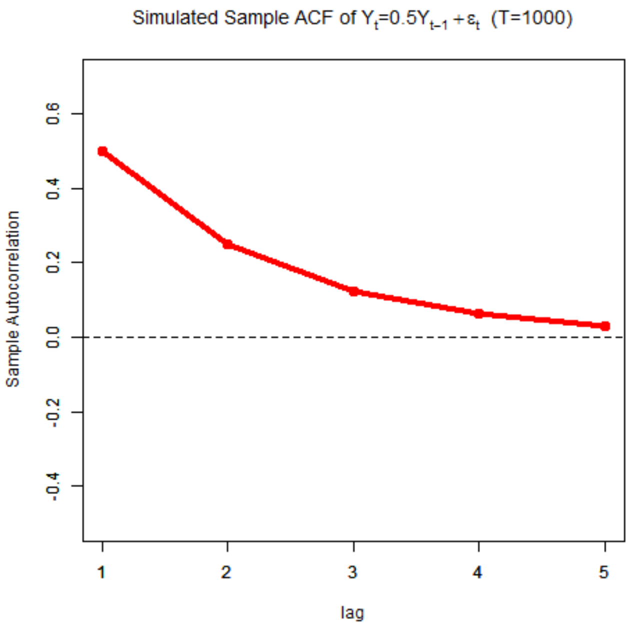
1. Generate a sample $\{y_1^b, .., y_T^b\}$ from $Y_t = 0.5Y_{t-1} + \varepsilon_t$ where $\varepsilon_t \overset{i.i.d}{\sim} N(0,1)$ for $T = 30$

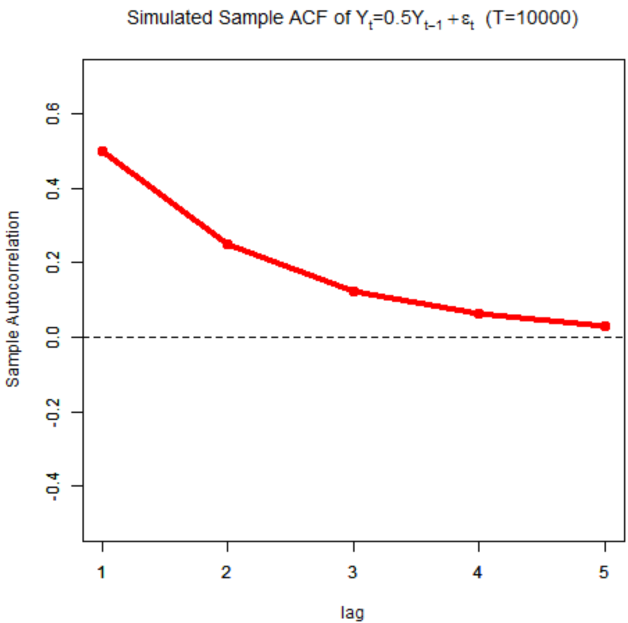
2. Calculate $\hat{\rho}_T^b(1)$,..,$\hat{\rho}_T^b(1)$ for each sample b

3. Plot Sample ACF $\hat{\rho}_T^b(1)$,..,$\hat{\rho}_T^b(5)$ for $b = \{1, .., 100\}$.

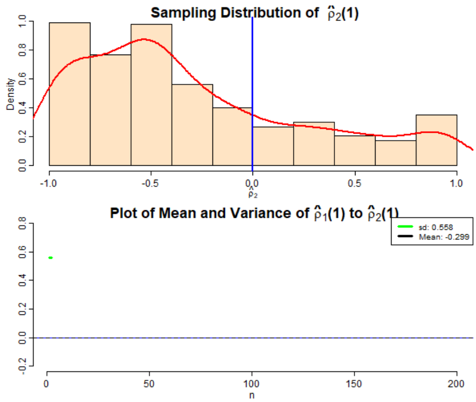


Simulated Sample ACF of Y_t=0.5Y_{t-1} + ε_t (T=100)
Sample Autocorrelation
-0.4
-0.2
0.0
0.2
0.4
0.6
1
2
3
4
5
lag


Simulated Sample ACF of Yt=0.5Yt-1 + εt (T=1000)
Sample Autocorrelation
0.6
0.4
0.2
0.0
-0.2
-0.4
1
2
3
4
5
lag


Simulated Sample ACF of Y_t=0.5Y_{t-1}+ε_t (T=10000)
Sample Autocorrelation
lag
-0.4
-0.2
0.0
0.2
0.4
0.6
1
2
3
4
5

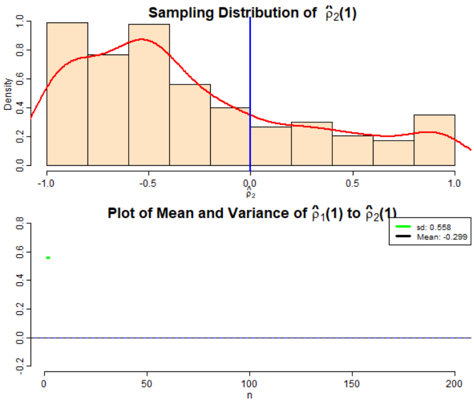
Simulation: Sampling Distribution of Sample Autocorrelation



Consider $Y_t = \varepsilon_t$ [i.e. $\phi_1 = 1$]

Theorem 1.8 $\Rightarrow \hat{\rho}_T(1) \overset{a}{\sim} N(0, 1/T)$

Sampling Distribution of $\hat{\rho}_T(1)$ for $T = 2, .., 200$



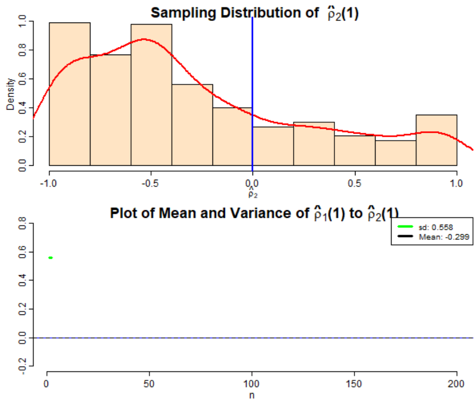
Consider $Y_t = \varepsilon_t$ [i.e. $\phi_1 = 1$]

Theorem 1.8 $\Rightarrow \hat{\rho}_T(1) \overset{a}{\sim} N(0, 1/T)$

Sampling Distribution of $\hat{\rho}_T(1)$ for $T = 2, .., 200$

1. Generate 1000 samples $\{y_1^b, .., y_T^b\}$ from $Y_t = 0.5Y_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim N(0, 1)$

Simulation: Sampling Distribution of Sample Autocorrelation



Consider $Y_t = \varepsilon_t$ [i.e. $\phi_1 = 1$]

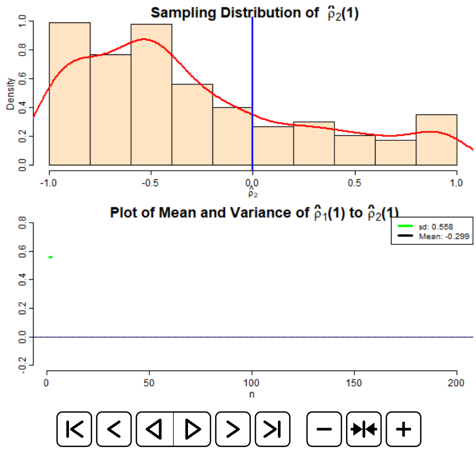
Theorem 1.8 $\Rightarrow \hat{\rho}_T(1) \overset{a}{\sim} N(0, 1/T)$

Sampling Distribution of $\hat{\rho}_T(1)$ for $T = 2, .., 200$

1. Generate 1000 samples $\{y_1^b, .., y_T^b\}$ from $Y_t = 0.5Y_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim N(0, 1)$

2. Calculate $\hat{\rho}_T^b(1)$ for samples $b = \{1, .., 1000\}$

Simulation: Sampling Distribution of Sample Autocorrelation



Consider $Y_t = \varepsilon_t$ [i.e. $\phi_1 = 1$]

Theorem 1.8 $\Rightarrow \hat{\rho}_T(1) \overset{a}{\sim} N(0, 1/T)$

Sampling Distribution of $\hat{\rho}_T(1)$ for $T = 2, .., 200$

1. Generate 1000 samples $\{y_1^b, .., y_T^b\}$ from $Y_t = 0.5Y_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim N(0, 1)$
2. Calculate $\hat{\rho}_T^b(1)$ for samples $b = \{1, .., 1000\}$
3. Plot distribution of $\hat{\rho}_T^b(1)$ for $b = \{1, .., 1000\}$

Maximum Likelihood Estimation of ARMA Processes

Refresher: Maximum Likelihood [ML]

$\{Y_t : t = 1, \cdots, T\}$ has **unknown** joint density function $h_{Y_1,\cdots,Y_T}(c_1, \cdots, c_T)$

Refresher: Maximum Likelihood [ML]

$\{Y_t : t = 1, \cdots, T\}$ has **unknown** joint density function $h_{Y_1,\cdots,Y_T}(c_1, \cdots, c_T)$

Assume joint density fn. known up to parameter vector θ in parameter space Θ

Refresher: Maximum Likelihood [ML]

$\{Y_t : t = 1, \cdots, T\}$ has **unknown** joint density function $h_{Y_1,\cdots,Y_T}(c_1, \cdots, c_T)$

Assume joint density fn. known up to parameter vector θ in parameter space Θ

$h_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1) \equiv f_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1; \theta)$ for some $\theta = \theta_0 \in \Theta$

Refresher: Maximum Likelihood [ML]

$\{Y_t : t = 1, \cdots, T\}$ has **unknown** joint density function $h_{Y_1,\cdots,Y_T}(c_1, \cdots, c_T)$

Assume joint density fn. known up to parameter vector θ in parameter space Θ

$$h_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1) \equiv f_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1; \theta) \text{ for some } \theta = \theta_0 \in \Theta$$

Define **likelihood function** $L(\theta) = f_{Y_T,\cdots,Y_1}(y_T, \cdots, y_1; \theta)$

Refresher: Maximum Likelihood [ML]

$\{Y_t : t = 1, \cdots, T\}$ has **unknown** joint density function $h_{Y_1,\cdots,Y_T}(c_1, \cdots, c_T)$

Assume joint density fn. known up to parameter vector θ in parameter space Θ

$$h_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1) \equiv f_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1; \theta) \text{ for some } \theta = \theta_0 \in \Theta$$

Define **likelihood function** $L(\theta) = f_{Y_T,\cdots,Y_1}(y_T, \cdots, y_1; \theta)$

ML Estimator: $\hat{\theta}_{ML} := \arg\max_{\theta \in \Theta} \ \ln L(\theta)$

Refresher: Maximum Likelihood [ML]

$\{Y_t : t = 1, \cdots, T\}$ has **unknown** joint density function $h_{Y_1,\cdots,Y_T}(c_1, \cdots, c_T)$

Assume joint density fn. known up to parameter vector θ in parameter space Θ

$$h_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1) \equiv f_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1; \theta) \text{ for some } \theta = \theta_0 \in \Theta$$

Define **likelihood function** $L(\theta) = f_{Y_T,\cdots,Y_1}(y_T, \cdots, y_1; \theta)$

ML Estimator: $\hat{\theta}_{ML} := \underset{\theta \in \Theta}{\arg\max} \; \ln L(\theta)$

If θ_0 unique, $\{Y_t\}$ strictly stationary (+ regularity conditions)

Refresher: Maximum Likelihood [ML]

$\{Y_t : t = 1, \cdots, T\}$ has **unknown** joint density function $h_{Y_1,\cdots,Y_T}(c_1, \cdots, c_T)$

Assume joint density fn. known up to parameter vector θ in parameter space Θ

$$h_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1) \equiv f_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1; \theta) \text{ for some } \theta = \theta_0 \in \Theta$$

Define **likelihood function** $L(\theta) = f_{Y_T,\cdots,Y_1}(y_T, \cdots, y_1; \theta)$

ML Estimator: $\hat{\theta}_{ML} := \underset{\theta\in\Theta}{\arg\max} \; \ln L(\theta)$

If θ_0 unique, $\{Y_t\}$ strictly stationary (+ regularity conditions)

$$\hat{\theta}_{ML} \xrightarrow{p} \theta_0$$

Refresher: Maximum Likelihood [ML]

$\{Y_t : t = 1, \cdots, T\}$ has **unknown** joint density function $h_{Y_1,\cdots,Y_T}(c_1, \cdots, c_T)$

Assume joint density fn. known up to parameter vector θ in parameter space Θ

$$h_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1) \equiv f_{Y_T,\cdots,Y_1}(c_T, \cdots, c_1; \theta) \text{ for some } \theta = \theta_0 \in \Theta$$

Define **likelihood function** $L(\theta) = f_{Y_T,\cdots,Y_1}(y_T, \cdots, y_1; \theta)$

ML Estimator: $\hat{\theta}_{ML} := \underset{\theta \in \Theta}{\arg\max} \; \ln L(\theta)$

If θ_0 unique, $\{Y_t\}$ strictly stationary (+ regularity conditions)

$$\hat{\theta}_{ML} \xrightarrow{p} \theta_0$$

$$\sqrt{T}(\hat{\theta}_{ML} - \theta_0) \xrightarrow{d} N(0, I(\theta_0)^{-1}) \quad \text{where } I(\theta) := -\mathbb{E}\left[\frac{\partial^2}{\partial\theta\partial\theta'} \ln L(\theta)\right]$$

Deriving Likelihood Function for Dependent Process

Generally $\{Y_t\}$ is not independent $\Rightarrow$ joint distribution **not** product of marginals

Deriving Likelihood Function for Dependent Process

Generally $\{Y_t\}$ is not independent $\Rightarrow$ joint distribution **not** product of marginals

Suppose joint density of $\{Y_t : t = 1, \cdots, T\}$ is $f_{Y_T,\cdots,Y_1}(y_1, \cdots, y_T; \theta)$ at $\theta = \theta_0$

Deriving Likelihood Function for Dependent Process

Generally $\{Y_t\}$ is not independent $\Rightarrow$ joint distribution **not** product of marginals

Suppose joint density of $\{Y_t : t = 1, \cdots, T\}$ is $f_{Y_T,\cdots,Y_1}(y_1, \cdots, y_T; \theta)$ at $\theta = \theta_0$

T=2 $\quad f_{Y_2,Y_1}(y_2, y_1; \theta) = f_{Y_2|Y_1}(y_2|y_1; \theta) f_{Y_1}(y_1; \theta)$

Deriving Likelihood Function for Dependent Process

Generally $\{Y_t\}$ is not independent $\Rightarrow$ joint distribution **not** product of marginals

Suppose joint density of $\{Y_t : t = 1, \cdots, T\}$ is $f_{Y_T, \cdots, Y_1}(y_1, \cdots, y_T; \theta)$ at $\theta = \theta_0$

T=2 $\quad f_{Y_2,Y_1}(y_2, y_1; \theta) = f_{Y_2|Y_1}(y_2|y_1; \theta) f_{Y_1}(y_1; \theta)$

T=3 $\quad f_{Y_3,Y_2,Y_1}(y_3, y_2, y_1; \theta) = f_{Y_3|Y_2,Y_1}(y_3|y_2, y_1; \theta) f_{Y_2|Y_1}(y_2|y_1; \theta) f_{Y_1}(y_1; \theta)$

Deriving Likelihood Function for Dependent Process

Generally $\{Y_t\}$ is not independent $\Rightarrow$ joint distribution **not** product of marginals

Suppose joint density of $\{Y_t : t = 1, \cdots, T\}$ is $f_{Y_T,\cdots,Y_1}(y_1, \cdots, y_T; \theta)$ at $\theta = \theta_0$

T=2 $\quad f_{Y_2,Y_1}(y_2, y_1; \theta) = f_{Y_2|Y_1}(y_2|y_1; \theta) f_{Y_1}(y_1; \theta)$

T=3 $\quad f_{Y_3,Y_2,Y_1}(y_3, y_2, y_1; \theta) = f_{Y_3|Y_2,Y_1}(y_3|y_2, y_1; \theta) f_{Y_2|Y_1}(y_2|y_1; \theta) f_{Y_1}(y_1; \theta)$

T $>$ **3** $\quad f_{Y_T,\cdots,Y_1}(y_T, \cdots, y_1; \theta) = \prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1}(y_T|y_{T-1}, \cdots, y_1); \theta) \times f_{Y_1}(y_1; \theta)$

ML Estimation of Stationary Gaussian AR(1) Process

Gaussian AR(1): $Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

ML Estimation of Stationary Gaussian AR(1) Process

Gaussian AR(1): $Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True parameter: $\theta = \theta_0$ where $\theta_0 = (c_0, \phi_{10}, \sigma_0^2)'$ unknown

ML Estimation of Stationary Gaussian AR(1) Process

Gaussian AR(1): $Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True parameter: $\theta = \theta_0$ where $\theta_0 = (c_0, \phi_{10}, \sigma_0^2)'$ unknown

Construct **likelihood function** $L(\theta) = f_{Y_1}(y_1; \theta) \prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1}(y_t|y_{t-1}, \cdots, y_1; \theta)$

ML Estimation of Stationary Gaussian AR(1) Process

Gaussian AR(1): $Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True parameter: $\theta = \theta_0$ where $\theta_0 = (c_0, \phi_{10}, \sigma_0^2)'$ unknown

Construct **likelihood function** $L(\theta) = f_{Y_1}(y_1; \theta) \prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1}(y_t|y_{t-1}, \cdots, y_1; \theta)$

$(Y_t|Y_{t-1} = y_{t-1}, \cdots, Y_1 = y_1)$ same distribution as $(Y_t|Y_{t-1} = y_{t-1})$

ML Estimation of Stationary Gaussian AR(1) Process

Gaussian AR(1): $Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True parameter: $\theta = \theta_0$ where $\theta_0 = (c_0, \phi_{10}, \sigma_0^2)'$ unknown

Construct **likelihood function** $L(\theta) = f_{Y_1}(y_1; \theta) \prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1}(y_t|y_{t-1}, \cdots, y_1; \theta)$

$(Y_t|Y_{t-1} = y_{t-1}, \cdots, Y_1 = y_1)$ same distribution as $(Y_t|Y_{t-1} = y_{t-1})$

$$\Rightarrow L(\theta) = f_{Y_1}(y_1; \theta) \prod_{t=2}^{T} f_{Y_t|Y_{t-1}}(y_t|y_{t-1}; \theta)$$

ML Estimation - Stationary Gaussian AR(1) Process

Conditional density of Y_t given Y_{t-1}

$$(Y_t | Y_{t-1} = y_{t-1}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$$

ML Estimation - Stationary Gaussian AR(1) Process

Conditional density of Y_t given Y_{t-1}

$$(Y_t | Y_{t-1} = y_{t-1}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$$

i.e. $$f_{Y_t|Y_{t-1}}(y_t; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

ML Estimation - Stationary Gaussian AR(1) Process

Conditional density of Y_t given Y_{t-1}

$$(Y_t|Y_{t-1} = y_{t-1}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$$

$$\text{i.e. } f_{Y_t|Y_{t-1}}(y_t;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

ML Estimation - Stationary Gaussian AR(1) Process

Conditional density of Y_t given Y_{t-1}

$$(Y_t | Y_{t-1} = y_{t-1}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$$

i.e. $$f_{Y_t|Y_{t-1}}(y_t; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2} \right)$$

Marginal density of Y_1

Stationarity($|\phi_1| < 1$) $\Rightarrow$ $\mathbb{E}[Y_1] = c/(1 - \phi_1)$ & Var$(Y_1) = \sigma^2/(1 - \phi_1^2)$

ML Estimation - Stationary Gaussian AR(1) Process

Conditional density of Y_t given Y_{t-1}

$$(Y_t | Y_{t-1} = y_{t-1}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$$

i.e. $f_{Y_t|Y_{t-1}}(y_t; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$

Marginal density of Y_1

Stationarity($|\phi_1| < 1$) $\Rightarrow$ $\mathbb{E}[Y_1] = c/(1-\phi_1)$ & Var$(Y_1) = \sigma^2/(1-\phi_1^2)$

Y_1 is linear function of process $\{\varepsilon_t\}$ **[MA(∞) representation]**

ML Estimation - Stationary Gaussian AR(1) Process

Conditional density of Y_t given Y_{t-1}

$$(Y_t|Y_{t-1} = y_{t-1}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$$

i.e. $f_{Y_t|Y_{t-1}}(y_t;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$

Marginal density of Y_1

Stationarity($|\phi_1| < 1$) $\Rightarrow$ $\mathbb{E}[Y_1] = c/(1-\phi_1)$ & Var$(Y_1) = \sigma^2/(1-\phi_1^2)$

Y_1 is linear function of process $\{\varepsilon_t\}$ **[MA(∞) representation]**

Linear combination of Gaussian random variates is also Gaussian distributed

ML Estimation - Stationary Gaussian AR(1) Process

Conditional density of Y_t given Y_{t-1}

$$(Y_t | Y_{t-1} = y_{t-1}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$$

i.e. $$f_{Y_t|Y_{t-1}}(y_t;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Marginal density of Y_1

Stationarity($|\phi_1| < 1$) $\Rightarrow$ $\mathbb{E}[Y_1] = c/(1-\phi_1)$ & Var$(Y_1) = \sigma^2/(1-\phi_1^2)$

Y_1 is linear function of process $\{\varepsilon_t\}$ **[MA(∞) representation]**

Linear combination of Gaussian random variates is also Gaussian distributed

$$\Rightarrow Y_1 \sim N(c/(1-\phi_1), \sigma^2/(1-\phi_1^2))$$

ML Estimation - Stationary Gaussian AR(1) Process

Conditional density of Y_t given Y_{t-1}

$$(Y_t | Y_{t-1} = y_{t-1}) \sim N(c + \phi_1 y_{t-1}, \sigma^2)$$

$$\text{i.e. } f_{Y_t|Y_{t-1}}(y_t; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Marginal density of Y_1

Stationarity($|\phi_1| < 1$) $\Rrightarrow$ $\mathbb{E}[Y_1] = c/(1-\phi_1)$ & Var$(Y_1) = \sigma^2/(1-\phi_1^2)$

Y_1 is linear function of process $\{\varepsilon_t\}$ **[MA(∞) representation]**

Linear combination of Gaussian random variates is also Gaussian distributed

$$\Rrightarrow Y_1 \sim N(c/(1-\phi_1), \sigma^2/(1-\phi_1^2))$$

$$\text{i.e. } f_{Y_1}(y_1; \theta) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2/(1-\phi_1^2)}} \exp\left(-\frac{(y_t - c/(1-\phi_1))^2}{2\sigma^2/(1-\phi_1^2)}\right)$$

ML Estimation - Stationary Gaussian AR(1) Process

Likelihood function

$$L(\theta) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2/(1-\phi_1^2)}} \exp\left(-\frac{(y_t - c/(1-\phi_1))^2}{2\sigma^2/(1-\phi_1^2)}\right) \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

ML Estimation - Stationary Gaussian AR(1) Process

Likelihood function

$$L(\theta) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2/(1-\phi_1^2)}} \exp\left(-\frac{(y_t - c/(1-\phi_1))^2}{2\sigma^2/(1-\phi_1^2)}\right) \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Log-likelihood function

$$\ln L(\theta) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma^2/(1-\phi_1^2)) - \frac{(y_1 - c/(1-\phi_1))^2}{2\sigma^2/(1-\phi_1^2)}$$
$$-\frac{T-1}{2}\ln(2\pi) - \frac{T-1}{2}\ln(\sigma^2) - \sum_{t=2}^{T}\left(\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

ML Estimation - Stationary Gaussian AR(1) Process

Likelihood function

$$L(\theta) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2/(1-\phi_1^2)}} \exp\left(-\frac{(y_t - c/(1-\phi_1))^2}{2\sigma^2/(1-\phi_1^2)}\right) \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Log-likelihood function

$$\ln L(\theta) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma^2/(1-\phi_1^2)) - \frac{(y_1 - c/(1-\phi_1))^2}{2\sigma^2/(1-\phi_1^2)}$$
$$-\frac{T-1}{2}\ln(2\pi) - \frac{T-1}{2}\ln(\sigma^2) - \sum_{t=2}^{T}\left(\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Can derive directly noting $(Y_1, \cdots, Y_T)' \sim N(\mathbf{c}, \mathbf{\Sigma})$ for some $\mathbf{c}$ & $\mathbf{\Sigma}$

ML Estimation - Stationary Gaussian AR(1) Process

Likelihood function

$$L(\theta) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2/(1-\phi_1^2)}} \exp\left(-\frac{(y_t - c/(1-\phi_1))^2}{2\sigma^2/(1-\phi_1^2)}\right) \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Log-likelihood function

$$\ln L(\theta) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma^2/(1-\phi_1^2)) - \frac{(y_1 - c/(1-\phi_1))^2}{2\sigma^2/(1-\phi_1^2)}$$
$$-\frac{T-1}{2}\ln(2\pi) - \frac{T-1}{2}\ln(\sigma^2) - \sum_{t=2}^{T}\left(\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Can derive directly noting $(Y_1, \cdots, Y_T)' \sim N(\mathbf{c}, \mathbf{\Sigma})$ for some $\mathbf{c}$ & $\mathbf{\Sigma}$

Use joint density of vector Gaussian variable [*Do as an exercise*, see **H** pp. 119-122]

ML Estimation - Stationary Gaussian AR(1) Process

Log likelihood function is non-linear in θ

ML Estimation - Stationary Gaussian AR(1) Process

Log likelihood function is non-linear in θ

Finding maximiser of $\ln L(\theta)$ w.r.t. θ (i.e. ML estimator) can be computationally challenging

ML Estimation - Stationary Gaussian AR(1) Process

Log likelihood function is non-linear in θ

Finding maximiser of $\ln L(\theta)$ w.r.t. θ (i.e. ML estimator) can be computationally challenging

This is known as the **exact maximum likelihood estimator**

ML Estimation - Stationary Gaussian AR(1) Process

Log likelihood function is non-linear in θ

Finding maximiser of $\ln L(\theta)$ w.r.t. θ (i.e. ML estimator) can be computationally challenging

This is known as the **exact maximum likelihood estimator**

Exact ML estimator of ϕ_{10} biased when $|\phi_{10}| > 1$ - **why?**

ML Estimation - Stationary Gaussian AR(1) Process

Log likelihood function is non-linear in θ

Finding maximiser of $\ln L(\theta)$ w.r.t. θ (i.e. ML estimator) can be computationally challenging

This is known as the **exact maximum likelihood estimator**

Exact ML estimator of ϕ_{10} biased when $|\phi_{10}| > 1$ - **why?**

Find **conditional maximum likelihood estimator** instead

ML Estimation - Stationary Gaussian AR(1) Process

Log likelihood function is non-linear in θ

Finding maximiser of $\ln L(\theta)$ w.r.t. θ (i.e. ML estimator) can be computationally challenging

This is known as the **exact maximum likelihood estimator**

Exact ML estimator of ϕ_{10} biased when $|\phi_{10}| > 1$ - **why?**

Find **conditional maximum likelihood estimator** instead

Doesn't rely on distributional assumptions on Y_1

ML Estimation - Stationary Gaussian AR(1) Process

Log likelihood function is non-linear in θ

Finding maximiser of $\ln L(\theta)$ w.r.t. θ (i.e. ML estimator) can be computationally challenging

This is known as the **exact maximum likelihood estimator**

- Exact ML estimator of ϕ_{10} biased when $|\phi_{10}| > 1$ - **why?**

Find **conditional maximum likelihood estimator** instead

- Doesn't rely on distributional assumptions on Y_1
- Conditional ML estimator has **closed form** solution

Conditional ML Estimation - Stationary Gaussian AR(1) Process

Exact ML requires knowledge of whole joint density fn. of $(Y_1, \cdots, Y_T)'$ up to some parameter θ

Conditional ML Estimation - Stationary Gaussian AR(1) Process

Exact ML requires knowledge of whole joint density fn. of $(Y_1, \cdots, Y_T)'$ up to some parameter θ

Conditional ML (for AR(1)) conditions on Y_1 equal to its realisation y_1

Conditional ML Estimation - Stationary Gaussian AR(1) Process

Exact ML requires knowledge of whole joint density fn. of $(Y_1, \cdots, Y_T)'$ up to some parameter θ

Conditional ML (for AR(1)) conditions on Y_1 equal to its realisation y_1

Conditional likelihood: $L(\theta|Y_1) = f_{Y_T,\cdots Y_2|Y_1}(y_T, \cdots, y_2|y_1; \theta)$

Conditional ML Estimation - Stationary Gaussian AR(1) Process

Exact ML requires knowledge of whole joint density fn. of $(Y_1, \cdots, Y_T)'$ up to some parameter θ

Conditional ML (for AR(1)) conditions on Y_1 equal to its realisation y_1

Conditional likelihood: $L(\theta|Y_1) = f_{Y_T,\cdots Y_2|Y_1}(y_T, \cdots, y_2|y_1; \theta)$

$Y_1 = y_1$ is fixed and conditional density doesn't require assumptions on distribution of Y_1

Conditional ML Estimation - Stationary Gaussian AR(1) Process

Exact ML requires knowledge of whole joint density fn. of $(Y_1, \cdots, Y_T)'$ up to some parameter θ

Conditional ML (for AR(1)) conditions on Y_1 equal to its realisation y_1

Conditional likelihood: $L(\theta|Y_1) = f_{Y_T,\cdots Y_2|Y_1}(y_T, \cdots, y_2|y_1; \theta)$

$Y_1 = y_1$ is fixed and conditional density doesn't require assumptions on distribution of Y_1

$$L(\theta|Y_1) = \prod_{t=2}^{T} f_{Y_t|Y_{t-1}}(y_t|y_{t-1}; \theta)$$

Conditional ML Estimation - Stationary Gaussian AR(1) Process

Exact ML requires knowledge of whole joint density fn. of $(Y_1, \cdots, Y_T)'$ up to some parameter θ

Conditional ML (for AR(1)) conditions on Y_1 equal to its realisation y_1

Conditional likelihood: $L(\theta|Y_1) = f_{Y_T,\cdots Y_2|Y_1}(y_T, \cdots, y_2|y_1;\theta)$

$Y_1 = y_1$ is fixed and conditional density doesn't require assumptions on distribution of Y_1

$$L(\theta|Y_1) = \prod_{t=2}^{T} f_{Y_t|Y_{t-1}}(y_t|y_{t-1};\theta) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Conditional ML Estimation - Stationary Gaussian AR(1) Process

Exact ML requires knowledge of whole joint density fn. of $(Y_1, \cdots, Y_T)'$ up to some parameter θ

Conditional ML (for AR(1)) conditions on Y_1 equal to its realisation y_1

Conditional likelihood: $L(\theta|Y_1) = f_{Y_T,\cdots Y_2|Y_1}(y_T, \cdots, y_2|y_1;\theta)$

$Y_1 = y_1$ is fixed and conditional density doesn't require assumptions on distribution of Y_1

$$L(\theta|Y_1) = \prod_{t=2}^{T} f_{Y_t|Y_{t-1}}(y_t|y_{t-1};\theta) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Conditional log-likelihood:

$$\ln L(\theta|Y_1) = -\frac{T-1}{2}\ln(2\pi) - \frac{T-1}{2}\ln(\sigma^2) - \sum_{t=2}^{T}\left(\frac{(y_t - c - \phi_1 y_{t-1})^2}{2\sigma^2}\right)$$

Exact ML Estimation - Stationary Gaussian AR(p) Process

Gaussian AR(p): $Y_t = c + \phi_1 Y_{t-1} + \cdots \phi_p Y_{t-p} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

Exact ML Estimation - Stationary Gaussian AR(p) Process

Gaussian AR(p): $Y_t = c + \phi_1 Y_{t-1} + \cdots \phi_p Y_{t-p} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True parameter: $\theta = \theta_0$ where $\theta_0 = (c_0, \phi_{10}, \cdots, \phi_{p0}, \sigma_0^2)'$ **unknown**

Exact ML Estimation - Stationary Gaussian AR(p) Process

Gaussian AR(p): $Y_t = c + \phi_1 Y_{t-1} + \cdots \phi_p Y_{t-p} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True parameter: $\theta = \theta_0$ where $\theta_0 = (c_0, \phi_{10}, \cdots, \phi_{p0}, \sigma_0^2)'$ **unknown**

$$L(\theta) = f_{Y_p,\cdots,Y_1}(y_p, \cdots, y_1; \theta) \prod_{t=p+1}^{T} f_{Y_t|Y_{t-1},\cdots,Y_{t-p}}(y_t|y_{t-1}; \cdots, y_{t-p}; \theta)$$

Exact ML Estimation - Stationary Gaussian AR(p) Process

Gaussian AR(p): $Y_t = c + \phi_1 Y_{t-1} + \cdots \phi_p Y_{t-p} + \varepsilon_t$ and $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True parameter: $\theta = \theta_0$ where $\theta_0 = (c_0, \phi_{10}, \cdots, \phi_{p0}, \sigma_0^2)'$ **unknown**

$$L(\theta) = f_{Y_p, \cdots, Y_1}(y_p, \cdots, y_1; \theta) \prod_{t=p+1}^{T} f_{Y_t|Y_{t-1}, \cdots, Y_{t-p}}(y_t|y_{t-1}; \cdots, y_{t-p}; \theta)$$

See **H** pp. 123-125 for full derivation of log-likelihood fn. of Gaussian AR(p)

Full derivation of exact ML log-likelihood fn. of Gaussian AR(p) **not** examinable for $p > 2$

Conditional ML Estimation - Stationary Gaussian AR(p) Process

Condition on $(Y_1 = y_1, \cdots, Y_p = y_p)$

Conditional ML Estimation - Stationary Gaussian AR(p) Process

Condition on $(Y_1 = y_1, \cdots, Y_p = y_p)$

$(Y_t | Y_{t-1} = y_{t-1}, \cdots, Y_{t-p} = y_{t-p}) \sim N(c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p}, \sigma^2)$ for $t = p+1, \cdots T$

Conditional ML Estimation - Stationary Gaussian AR(p) Process

Condition on $(Y_1 = y_1, \cdots, Y_p = y_p)$

$(Y_t | Y_{t-1} = y_{t-1}, \cdots, Y_{t-p} = y_{t-p}) \sim N(c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p}, \sigma^2)$ for $t = p+1, \cdots T$

Conditional likelihood:

$$L(\theta | Y_1, \cdots Y_p) = \prod_{t=p+1}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p})^2}{2\sigma^2} \right)$$

Conditional ML Estimation - Stationary Gaussian AR(p) Process

Condition on $(Y_1 = y_1, \cdots, Y_p = y_p)$

$(Y_t | Y_{t-1} = y_{t-1}, \cdots, Y_{t-p} = y_{t-p}) \sim N(c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p}, \sigma^2)$ for $t = p+1, \cdots T$

Conditional likelihood:

$$L(\theta | Y_1, \cdots Y_p) = \prod_{t=p+1}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_t - c - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p})^2}{2\sigma^2} \right)$$

Conditional log-likelihood:

$$\ln L(\theta | Y_1, \cdots, Y_p) = -\frac{T-1}{2} \ln(2\pi) - \frac{T-1}{2} \ln(\sigma^2) - \sum_{t=p+1}^{T} \left(\frac{(y_t - c - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p})^2}{2\sigma^2} \right)$$

Summary of Properties of Exact and Conditional ML of Stationary Gaussian AR(p)

Conditional ML estimator of $(c, \phi_1, \cdots, \phi_p)'$ equivalent to **ordinary least squares**

Summary of Properties of Exact and Conditional ML of Stationary Gaussian AR(p)

Conditional ML estimator of $(c, \phi_1, \cdots, \phi_p)'$ equivalent to **ordinary least squares**

$\Rightarrow$ closed form solution (easier to solve than exact ML estimator)

Summary of Properties of Exact and Conditional ML of Stationary Gaussian AR(p)

Conditional ML estimator of $(c, \phi_1, \cdots, \phi_p)'$ equivalent to **ordinary least squares**

$\Rrightarrow$ closed form solution (easier to solve than exact ML estimator)

$\Rrightarrow$Conditional ML estimator consistent estimator of $(c, \phi_1, \cdots, \phi_p)'$ when ε_t **not** Gaussian

Summary of Properties of Exact and Conditional ML of Stationary Gaussian AR(p)

Conditional ML estimator of $(c, \phi_1, \cdots, \phi_p)'$ equivalent to **ordinary least squares**

$\Rrightarrow$ closed form solution (easier to solve than exact ML estimator)

$\Rrightarrow$Conditional ML estimator consistent estimator of $(c, \phi_1, \cdots, \phi_p)'$ when ε_t **not** Gaussian

Conditional ML estimator **consistent** when AR(p) does **not** satisfy stationary condition

Summary of Properties of Exact and Conditional ML of Stationary Gaussian AR(p)

Conditional ML estimator of $(c, \phi_1, \cdots, \phi_p)'$ equivalent to **ordinary least squares**

$\Rrightarrow$ closed form solution (easier to solve than exact ML estimator)

$\Rrightarrow$Conditional ML estimator consistent estimator of $(c, \phi_1, \cdots, \phi_p)'$ when ε_t **not** Gaussian

Conditional ML estimator **consistent** when AR(p) does **not** satisfy stationary condition

Not the case for exact ML estimator (inconsistent)

Conditional ML - Gaussian MA(1) Process

Gaussian MA(1): $Y_t = \alpha + \theta_1 \varepsilon_{t-1} + \varepsilon_t$ & $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

Conditional ML - Gaussian MA(1) Process

Gaussian MA(1): $Y_t = \alpha + \theta_1 \varepsilon_{t-1} + \varepsilon_t$ & $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True Parameter: $\theta = \theta_0 = (\alpha_0, \theta_{10}, \sigma_0^2)$ **unknown**

Conditional ML - Gaussian MA(1) Process

Gaussian MA(1): $Y_t = \alpha + \theta_1 \varepsilon_{t-1} + \varepsilon_t$ & $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True Parameter: $\theta = \theta_0 = (\alpha_0, \theta_{10}, \sigma_0^2)$ **unknown**

Derive likelihood function similar to the AR(1) using Bayes Rule

Conditional ML - Gaussian MA(1) Process

Gaussian MA(1): $Y_t = \alpha + \theta_1 \varepsilon_{t-1} + \varepsilon_t$ & $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True Parameter: $\theta = \theta_0 = (\alpha_0, \theta_{10}, \sigma_0^2)$ **unknown**

Derive likelihood function similar to the AR(1) using Bayes Rule

Difference with AR(1): the 'regressor' ϵ_{t-1} is **not observed**

Conditional ML - Gaussian MA(1) Process

Gaussian MA(1): $Y_t = \alpha + \theta_1 \varepsilon_{t-1} + \varepsilon_t$ & $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True Parameter: $\theta = \theta_0 = (\alpha_0, \theta_{10}, \sigma_0^2)$ **unknown**

Derive likelihood function similar to the AR(1) using Bayes Rule

Difference with AR(1): the 'regressor' ϵ_{t-1} is **not observed**

Derive **conditional likelihood** $f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T, \cdots, y_1|\varepsilon_0 = 0, \theta)$

Conditional ML - Gaussian MA(1) Process

Gaussian MA(1): $Y_t = \alpha + \theta_1 \varepsilon_{t-1} + \varepsilon_t$ & $\varepsilon_t \overset{i.i.d}{\sim} N(0, \sigma^2)$

True Parameter: $\theta = \theta_0 = (\alpha_0, \theta_{10}, \sigma_0^2)$ **unknown**

Derive likelihood function similar to the AR(1) using Bayes Rule

Difference with AR(1): the 'regressor' ϵ_{t-1} is **not observed**

Derive **conditional likelihood** $f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1|\varepsilon_0 = 0, \theta)$

Express **realised shocks** ε_t as function of **observable** $(y_{t-1},\cdots,y_1)$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional on $\varepsilon_0 = 0$ can express realised shocks ε_t as function of $(y_{t-1},\cdots,y_1)$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional on $\varepsilon_0 = 0$ can express realised shocks ε_t as function of $(y_{t-1},\cdots,y_1)$

$$\varepsilon_1 = y_1 - \alpha \qquad [Y_1 = y_1]$$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional on $\varepsilon_0 = 0$ can express realised shocks ε_t as function of $(y_{t-1},\cdots,y_1)$

$$\varepsilon_1 = y_1 - \alpha \qquad [Y_1 = y_1]$$

$$\varepsilon_2 = y_2 - \alpha - \theta_1\varepsilon_1 \qquad [Y_1 = y_1, Y_2 = y_2]$$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional on $\varepsilon_0 = 0$ can express realised shocks ε_t as function of $(y_{t-1},\cdots,y_1)$

$$\varepsilon_1 = y_1 - \alpha \qquad [Y_1 = y_1]$$

$$\varepsilon_2 = y_2 - \alpha - \theta_1\varepsilon_1 \qquad [Y_1 = y_1, Y_2 = y_2]$$

$$\vdots$$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional on $\varepsilon_0 = 0$ can express realised shocks ε_t as function of $(y_{t-1},\cdots,y_1)$

$$\varepsilon_1 = y_1 - \alpha \qquad [Y_1 = y_1]$$

$$\varepsilon_2 = y_2 - \alpha - \theta_1\varepsilon_1 \qquad [Y_1 = y_1, Y_2 = y_2]$$

$$\vdots$$

$$\varepsilon_T = y_T - \alpha - \theta_1\varepsilon_{T-1} \qquad [Y_1 = y_1,\cdots,Y_T = y_T]$$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional on $\varepsilon_0 = 0$ can express realised shocks ε_t as function of $(y_{t-1},\cdots,y_1)$

$$\varepsilon_1 = y_1 - \alpha \qquad [Y_1 = y_1]$$

$$\varepsilon_2 = y_2 - \alpha - \theta_1\varepsilon_1 \qquad [Y_1 = y_1, Y_2 = y_2]$$

$$\vdots$$

$$\varepsilon_T = y_T - \alpha - \theta_1\varepsilon_{T-1} \qquad [Y_1 = y_1,\cdots,Y_T = y_T]$$

$Y_1|\varepsilon_0 = 0 \sim N(\alpha,\sigma^2)$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional on $\varepsilon_0 = 0$ can express realised shocks ε_t as function of $(y_{t-1},\cdots,y_1)$

$$\varepsilon_1 = y_1 - \alpha \qquad [Y_1 = y_1]$$

$$\varepsilon_2 = y_2 - \alpha - \theta_1\varepsilon_1 \qquad [Y_1 = y_1, Y_2 = y_2]$$

$$\vdots$$

$$\varepsilon_T = y_T - \alpha - \theta_1\varepsilon_{T-1} \qquad [Y_1 = y_1,\cdots,Y_T = y_T]$$

$Y_1|\varepsilon_0 = 0 \sim N(\alpha,\sigma^2)$

$(Y_t|Y_{t-1} = y_{t-1},\cdots Y_1 = y_1,\varepsilon_0 = 0)$ same distribution as $(Y_t|\varepsilon_{t-1})$ for $t = 2,..,T$

Conditional ML - Gaussian MA(1) Process

Conditional likelihood:

$$f_{Y_T,\cdots,Y_1|\varepsilon_0=0}(y_T,\cdots,y_1,;\theta) = f_{Y_1|\epsilon_0=0}(y_1;\theta)\prod_{t=2}^{T} f_{Y_t|Y_{t-1},\cdots,Y_1,\varepsilon_0=0}(y_t|y_{t-1},\cdots,y_1;\theta)$$

Conditional on $\varepsilon_0 = 0$ can express realised shocks ε_t as function of $(y_{t-1},\cdots,y_1)$

$$\varepsilon_1 = y_1 - \alpha \qquad [Y_1 = y_1]$$

$$\varepsilon_2 = y_2 - \alpha - \theta_1\varepsilon_1 \qquad [Y_1 = y_1, Y_2 = y_2]$$

$$\vdots$$

$$\varepsilon_T = y_T - \alpha - \theta_1\varepsilon_{T-1} \qquad [Y_1 = y_1,\cdots,Y_T = y_T]$$

$Y_1|\varepsilon_0 = 0 \sim N(\alpha,\sigma^2)$

$(Y_t|Y_{t-1} = y_{t-1},\cdots Y_1 = y_1,\varepsilon_0 = 0)$ same distribution as $(Y_t|\varepsilon_{t-1})$ for $t = 2,..,T$

i.e. $(Y_t|Y_{t-1} = y_{t-1},\cdots,Y_1 = y_1,\varepsilon_0 = 0) \sim N(\alpha + \theta_1\varepsilon_{t-1},\sigma^2)$

ML Estimation of ARMA Models (what you need to know)

We will **not** cover exact ML estimation of models with MA component or with AR(p) for $p > 2$

ML Estimation of ARMA Models (what you need to know)

We will **not** cover exact ML estimation of models with MA component or with AR(p) for $p > 2$

All other cases of exact/conditional ML estimation for ARMA models are examinable

ML Estimation of ARMA Models (what you need to know)

We will **not** cover exact ML estimation of models with MA component or with AR(p) for $p > 2$

All other cases of exact/conditional ML estimation for ARMA models are examinable

See **H** p.130 for conditional ML for MA(q)

ML Estimation of ARMA Models (what you need to know)

We will **not** cover exact ML estimation of models with MA component or with AR(p) for $p > 2$

All other cases of exact/conditional ML estimation for ARMA models are examinable

See **H** p.130 for conditional ML for MA(q)

See **H** p.134 for conditional ML for ARMA(p,q)

ML Estimation of ARMA Models (what you need to know)

We will **not** cover exact ML estimation of models with MA component or with AR(p) for $p > 2$

All other cases of exact/conditional ML estimation for ARMA models are examinable

See **H** p.130 for conditional ML for MA(q)

See **H** p.134 for conditional ML for ARMA(p,q)

H pp. 132-139 for ML estimator computation in non-linear models (useful, but **not examinable**)

Reading list

Main Reading

Hamilton

[4.8] Wold's Decomposition (Lecture 1 readinf) and section on sample autocorrelations.

[5.1] Intro to ML. It may be useful to refresh your knowledge of ML estimation from your PG/UG notes.
[5.2] Conditional and exact ML estimation of Gaussian AR(1). Understand two alternative ways of deriving the log-likelihood function. You should be comfortable understanding both methods.
[5.3] Generalises 5.2 to AR(p). You need to understand all of this sub-chapter, aside from deriving the exact log-likelihood of AR(p) for p¿2
[5.4] Conditional ML estimation of MA(1) (read the part on exact ML, not examinable)
[5.5] Extends 5.4 to general MA(q), again only the conditional ML part is examinable.
[5.6] Conditional ARMA(p,q)
[5.7 Numerical optimisation (read over this material [omit the DFP if you wish], it is not examinable for my half of the course)
[5.8] ML estimation and inference. We did not cover this in lecture, a lot of this you saw in EC . You need to be comfortable doing the likelihood ratio and score test and deriving asymptotic standard errors of ML estimator.

[7.1] Review of asymptotic distribution theory
[7.2] Limit theorems for serially dependent processes pp.188-190 and CLT for stationary processes on p.195

Moving On

Covered univariate models and ML estimation of parameters of stationary processes

Moving On

Covered univariate models and ML estimation of parameters of stationary processes

Lects 3-5 extend these results to **multivariate** time series

Moving On

Covered univariate models and ML estimation of parameters of stationary processes

Lects 3-5 extend these results to **multivariate** time series

Model **vector processes**, e.g. the process generating interest rates and inflation together

Moving On

Covered univariate models and ML estimation of parameters of stationary processes

Lects 3-5 extend these results to **multivariate** time series

Model **vector processes**, e.g. the process generating interest rates and inflation together

Study vector autoregression (**VAR**) and **impulse response functions**

Moving On

Covered univariate models and ML estimation of parameters of stationary processes

Lects 3-5 extend these results to **multivariate** time series

Model **vector processes**, e.g. the process generating interest rates and inflation together

Study vector autoregression (**VAR**) and **impulse response functions**

Estimation/inference from VARs (includes univariate as a special case)