

Diagnostic checking e criteri di selezione

Distribuzione Normale

Jarque - Bera (1975)

$$\sqrt{b_2} = \frac{m_3}{m_2^{3/2}} = 0, \quad b_3 = \frac{m_4}{m_2^2} = 3$$

$$N = \frac{nb_2}{6} + \frac{n(b_3 - 3)^2}{24} \sim \chi^2(2)$$

Test per verificare l'ipotesi di normalità

Diagnostic checking e criteri di selezione

Distribuzione Normale

Il test di Jarque - Bera (1975):

H0: la serie ha distribuzione normale

Molto sensibile alla curtosi

Test alternativi sulla distribuzione normale (Smirnov – Kolmogorov, ...)

Diagnostic checking e criteri di selezione

No autocorrelazione

Bartlett (1946)

Box-Pierce (1978)

Ljung-Box

Portmanteau test

$$Q = n(n+2) \sum_{k=1}^K (n-k)^{-1} \hat{\rho}_k^2 \sim \chi^2(K-m)$$

$\pm 2 \sqrt{\frac{1}{n}}$

Test per verificare correlazione nulla (singole o globali)

Test applicato sui residui

Diagnostic checking e criteri di selezione

No eteroschedasticità:

Test ARCH sui residui

H0: i residui sono omoschedastici
(varianza costante)

H1: i residui sono eteroschedastici

Il test si distribuisce come un χ^2

Diagnostic checking e criteri di selezione

Model selection:

Akaike (1973, 1974)

$$\min AIC = n \ln \bar{\sigma}_a^2 + 2M,$$

Schwarz (1978)

$$\min SBC = n \ln \bar{\sigma}_a^2 + M \ln n$$

Hannan-Rissanen (1981)

$$\min HR = \ln \bar{\sigma}_a^2 + \frac{M \ln n}{n}$$

Diagnostic checking e criteri di selezione

Model selection:

I criteri di selezione tengono conto della verosimiglianza, della variabilità e del numero di parametri stimati.

Si penalizzano quei modelli che includono un numero elevato di parametri, se non rilevanti

Simili al concetto di R^2 aggiustato

- Al contrario di quanto avviene nel modello di regressione semplice, l' R^2 non può essere usato per misurare la bontà del modello.
- La scelta del modello deve avvenire tramite attenta diagnostica dei residui.

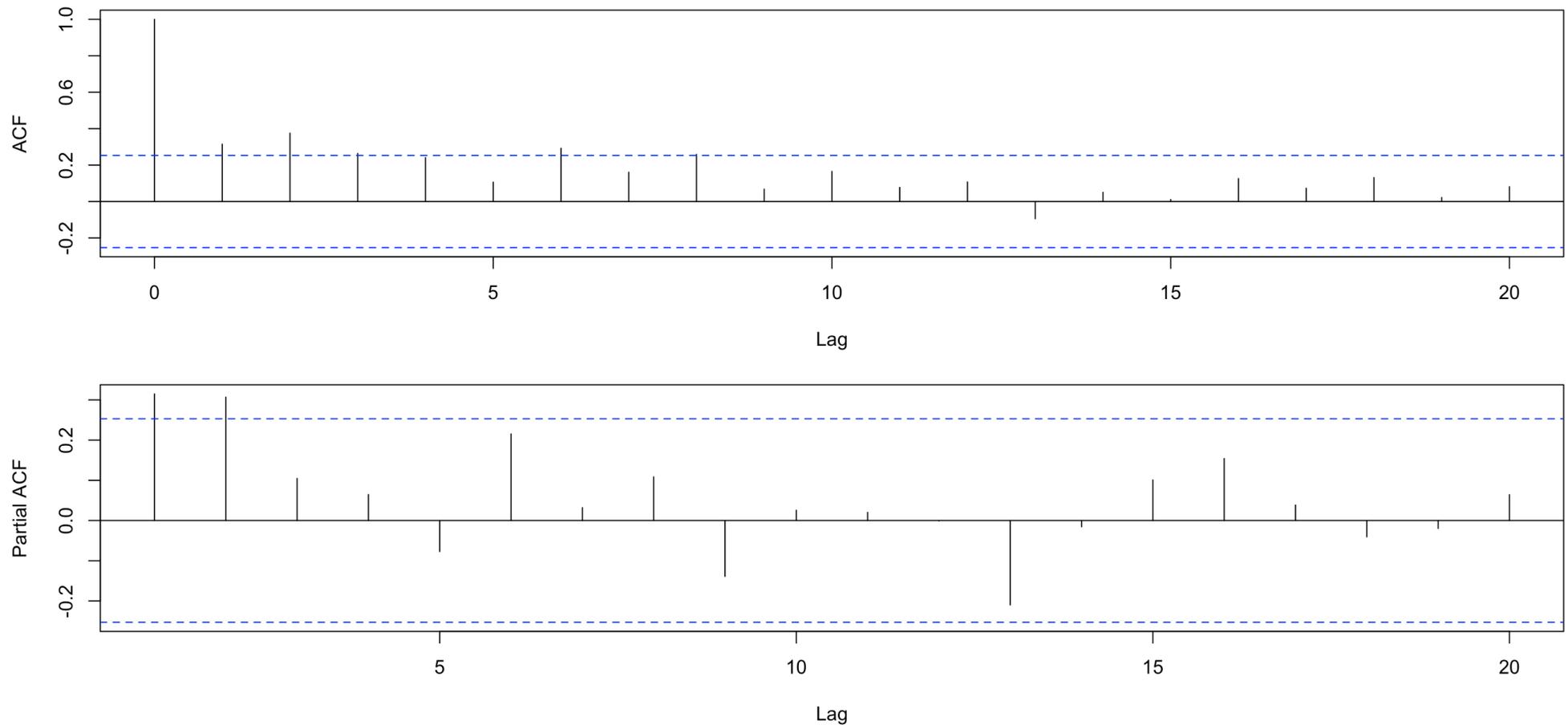
Esempio: New Haven

- Usiamo la serie "Temperature medie annuali in New Haven dal 1912 al 1971"



Esempio: New Haven

- Calcoliamo ACF e PACF, $k=20$



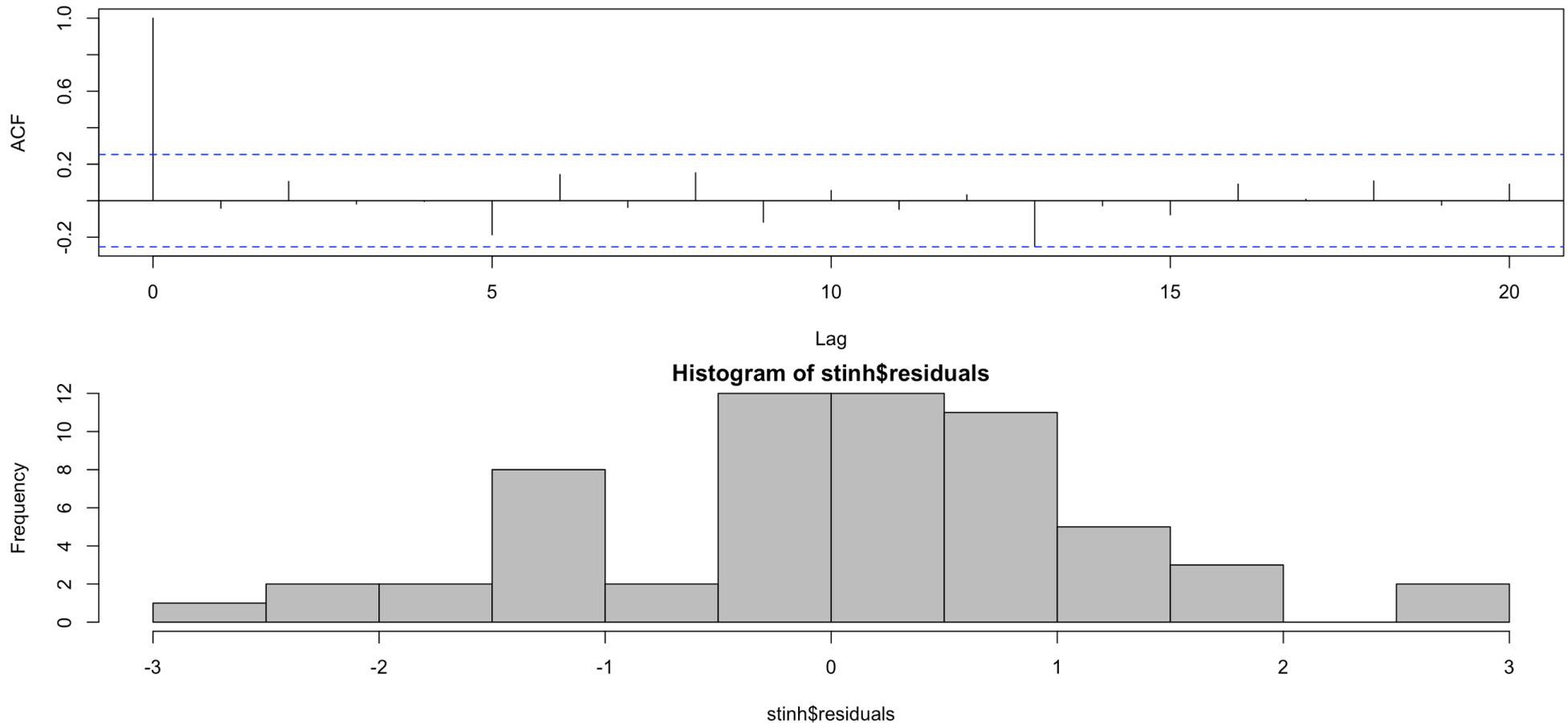
Esempio: New Haven

Stimiamo il modello

```
stinh <- arima(nhtemp,order=c(1,0,2))
stinh
Coefficients:
      ar1      ma1      ma2  intercept
      0.8800 -0.6983  0.0657   51.1879
s.e.  0.1629  0.2178  0.1559    0.4073
sigma^2 estimated as 1.251:  log likelihood = -92.05,  aic = 194.1
coefest(stinh)
z test of coefficients:
      Estimate Std. Error  z value  Pr(>|z|)
ar1      0.879987   0.162897   5.4021 6.586e-08 ***
ma1     -0.698260   0.217831  -3.2055 0.001348 **
ma2      0.065658   0.155930   0.4211 0.673700
intercept 51.187851   0.407264 125.6872 < 2.2e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

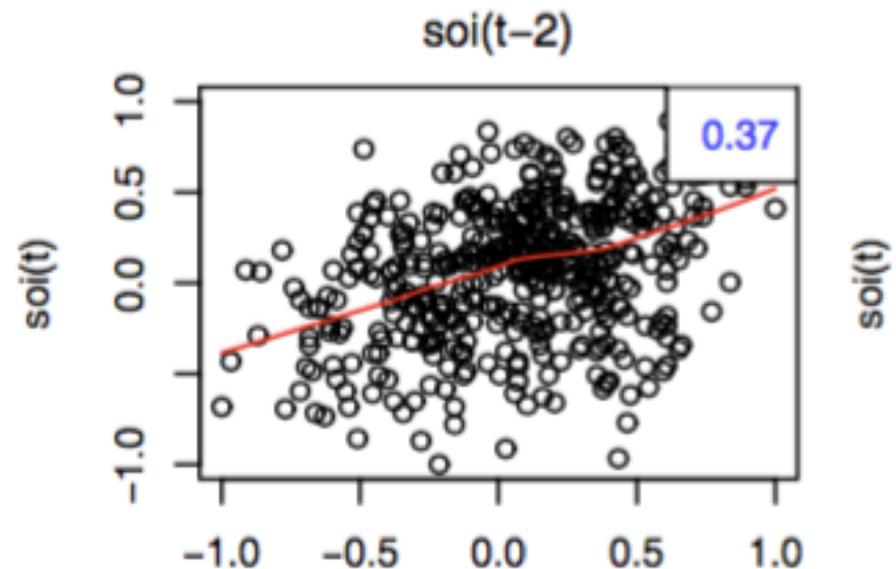
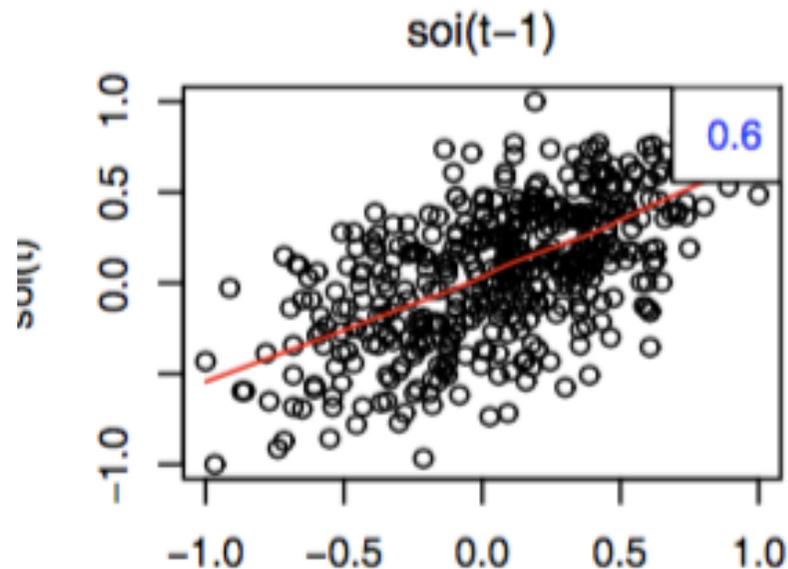
```
stinh <- arima(nhtemp,order=c(1,0,1))
stinh
Coefficients:
      ar1      ma1  intercept
      0.9151 -0.7088   51.1690
s.e.  0.0941  0.1668    0.4403
sigma^2 estimated as 1.254:  log likelihood = -92.15,  aic = 192.29
coefest(stinh)
z test of coefficients:
      Estimate Std. Error  z value  Pr(>|z|)
ar1      0.915069   0.094117   9.7227 < 2.2e-16 ***
ma1     -0.708839   0.166811  -4.2494 2.144e-05 ***
intercept 51.168952   0.440325 116.2074 < 2.2e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Diagnostic checking: New Haven



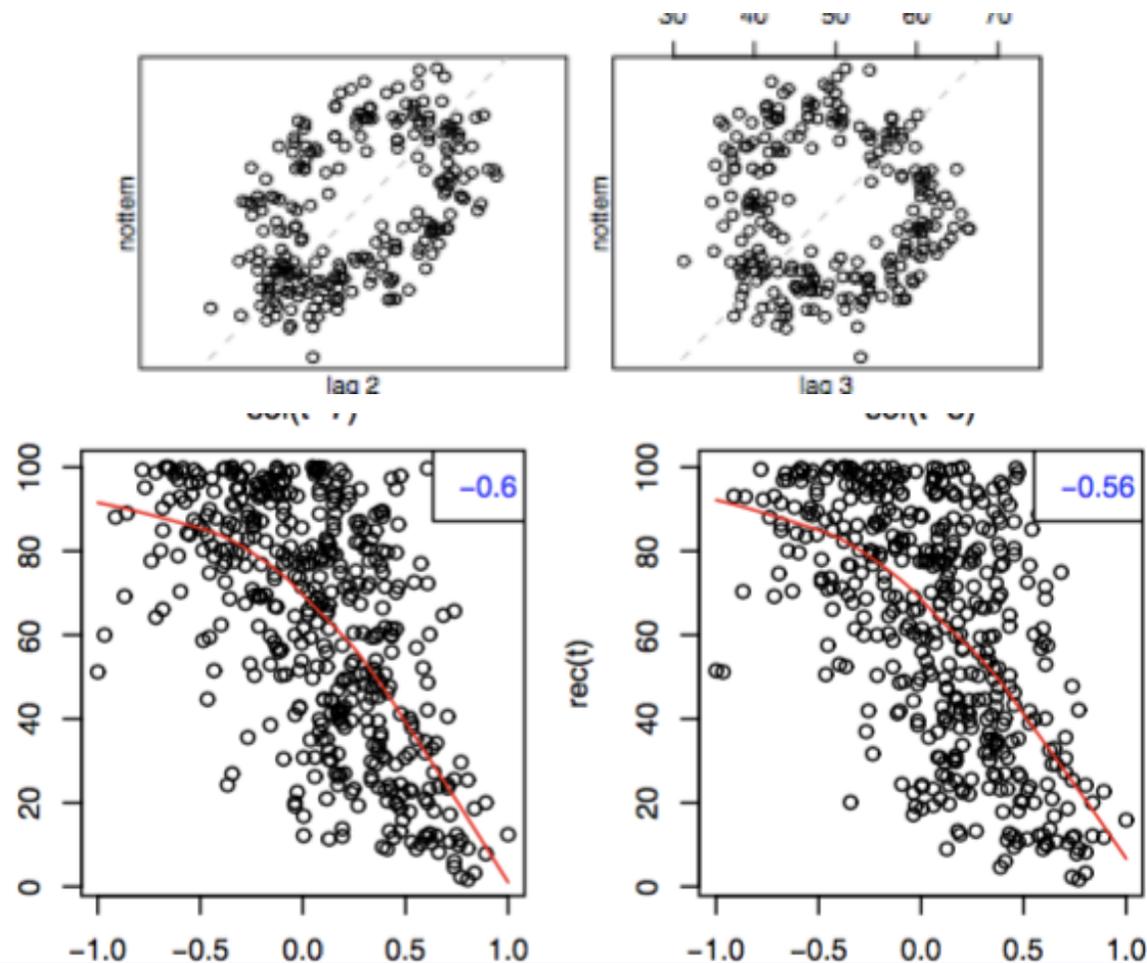
Analisi esplorativa

- Le figure riportano i grafici di y_t disegnato rispetto a y_{t-h} per alcuni valori di h . Questo tipo di grafici è detto di *autodispersione* (*lag plot* nella letteratura anglosassone).



Analisi esplorativa

- Questo tipo di grafico molto utile per individuare andamenti periodici nella serie o non lineari.



Stationary stochastic processes

Necessario porre delle restrizioni su alcune proprietà fondamentali: **stazionarietà debole**

$$\begin{aligned}E(X_t) &= \mu = \text{const}, & \forall t \\E(X_t - \mu)^2 &= \sigma^2 = \text{const} < +\infty, \\E[(X_t - \mu)(X_{t-k} - \mu)] &= \gamma_k\end{aligned}$$

La media e la varianza sono costanti, l'autocorrelazione dipende solo da k e non da t .

Le trasformazioni

La variazione assoluta si ottiene immediatamente come differenza del valore di X al tempo t rispetto allo stesso un tempo precedente:

$$\Delta X_t = X_t - X_{t-1}$$

Il simbolo usato per la differenza assoluta è Δ .
Analizzando ΔX piuttosto che X , il modello specificato spiegherà la variazione di X .

Le trasformazioni

L'altra trasformazione importante è quella che porta al tasso di variazione della serie

$$\frac{X_t - x_{t-1}}{X_{t-1}}$$

Si dimostra facilmente che questa è ben approssimata dalla seguente espressione

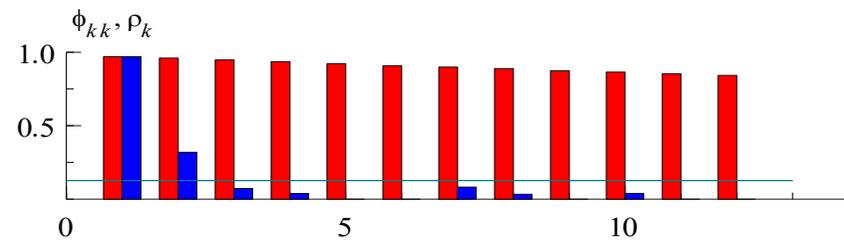
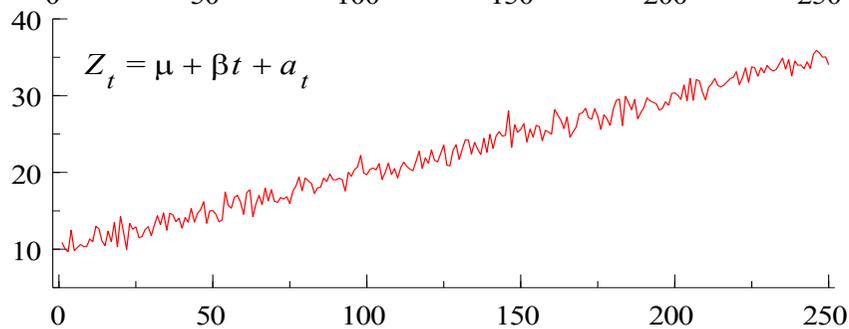
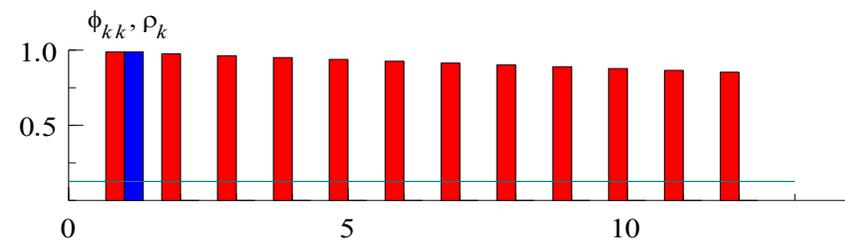
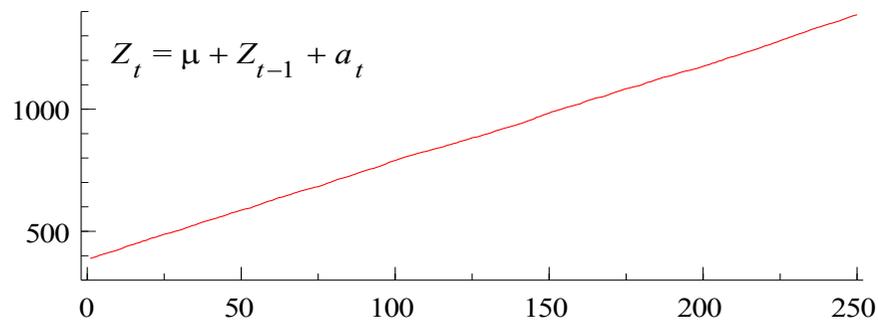
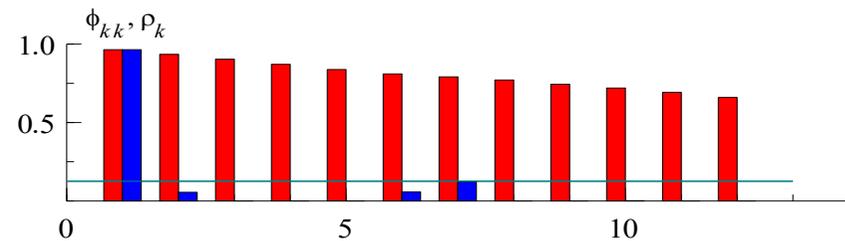
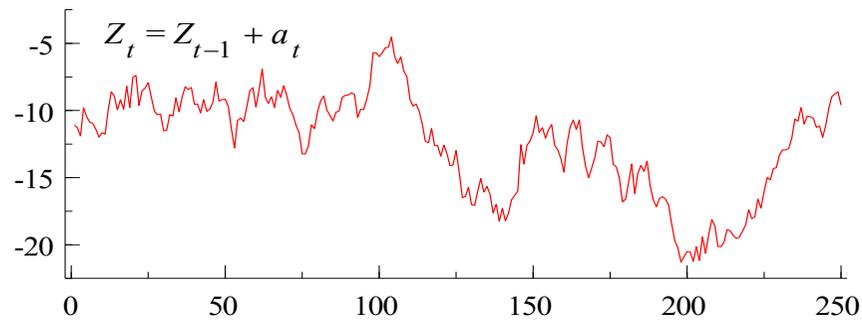
$$\frac{X_t - X_{t-1}}{X_{t-1}} \approx \log X_t - \log X_{t-1} = \Delta \log X_t$$

Molto spesso quindi piuttosto che rilevare Y_t , si studierà $\log Y_t$ e poi, successivamente, la sua variazione.

La trasformazione logaritmica viene ampiamente utilizzata in statistica.

Nonstazionarietà

Random walks & Deterministic trend processes



Nonstazionarietà in media

- I primi due processi hanno trend stocastico, mentre il terzo ha trend deterministico.
- Il trend stocastico è spiegato da un modello **random walk**.
- Esiste una somiglianza tra i modelli AR(1), RW e RWD:

AR(1):

$$Y_t = \varphi_1 Y_{t-1} + a_t \quad |\varphi_1| < 1$$

RW :

$$Y_t = Y_{t-1} + a_t$$

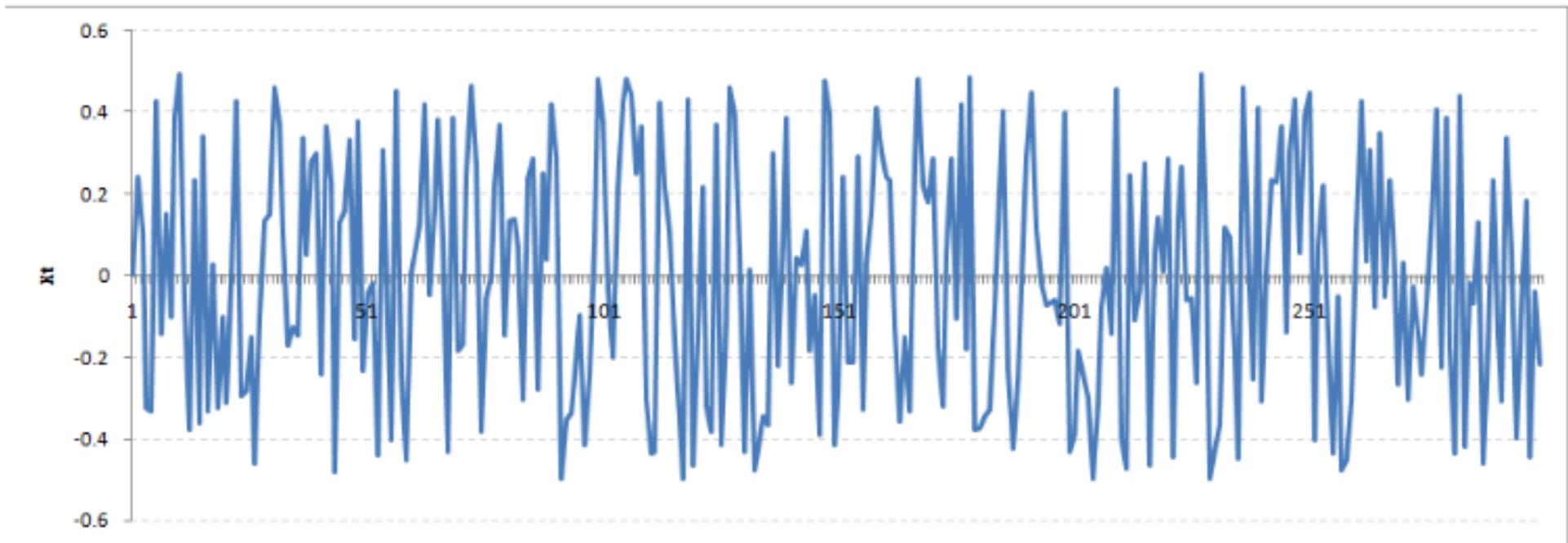
RWD:

$$Y_t = \mu + Y_{t-1} + a_t$$

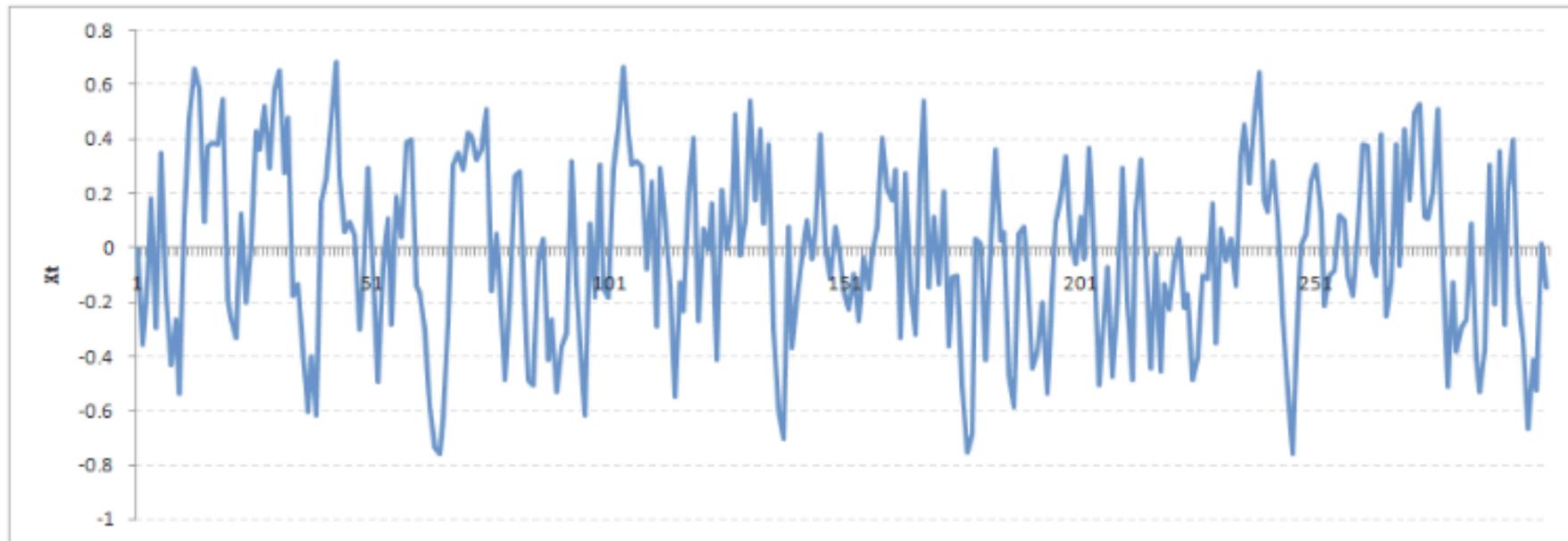
Comportamento dei modelli

MODELLO	$Z_t = \varphi_1 Z_{t-1} + a_t$ AR(1)	$Z_t = Z_{t-1} + a_t$ RW	$Z_t = \mu + Z_{t-1} + a_t$ RWD
Stationarietà	SI	NO	NO
Media $E(Z_t)$	0	0	μ_t
Varianza $VAR(Z_t)$	$\frac{1}{1 - \varphi_1^2}$	t	t
Autocov $Cov(Z_t, Z_{t-k})$	$\frac{\varphi_1^k}{1 - \varphi_1^2}$	(t-k)	(t-k)

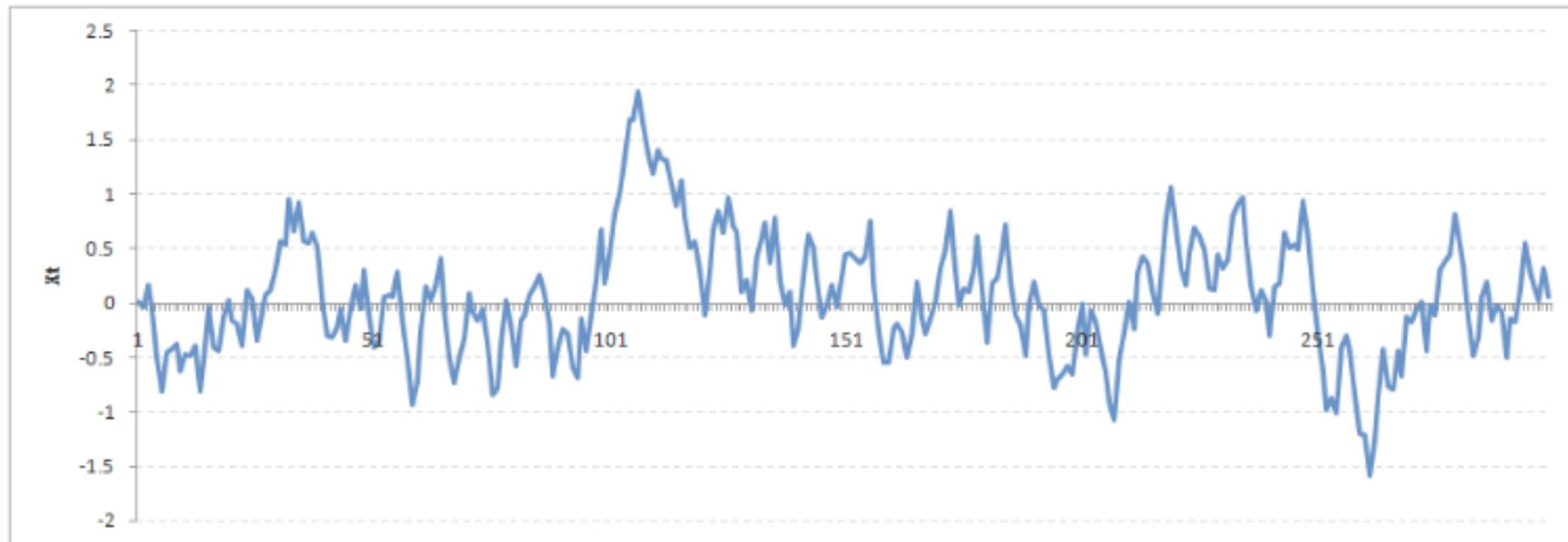
AR(1) with $\phi=0$



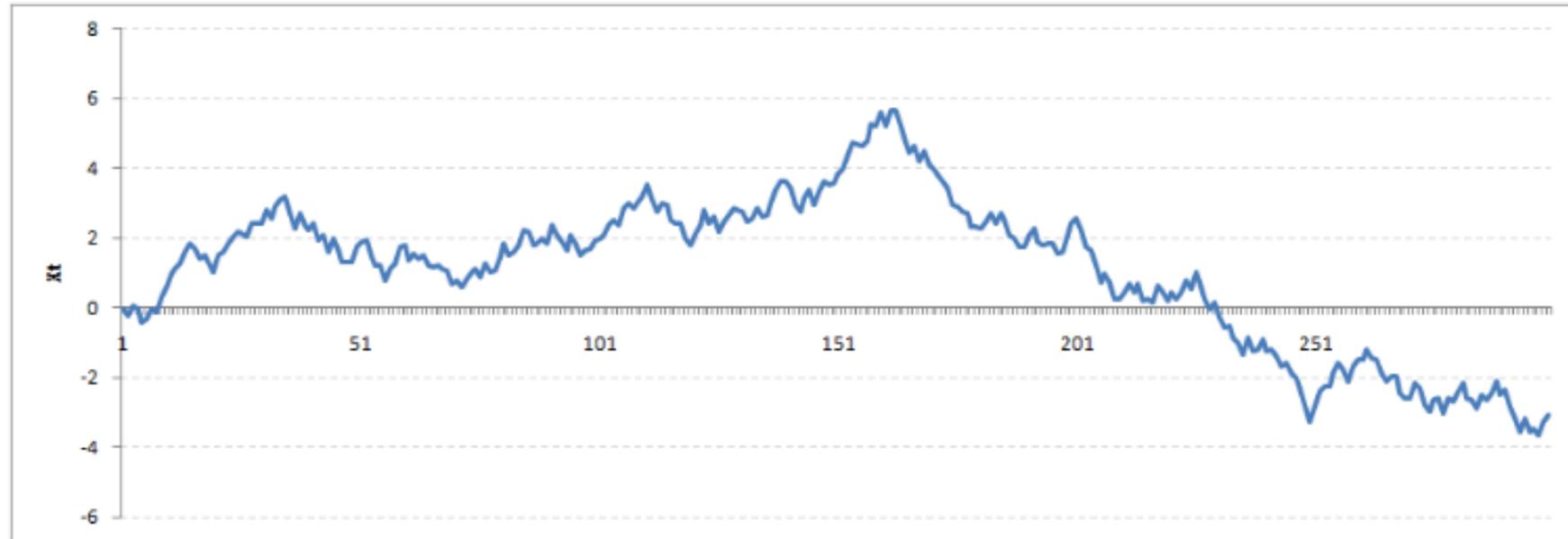
AR(1) with $\phi=0.5$



AR(1) with $\phi=0.8$



AR(1) with $\phi=-1.0$



Nonstazionarietà e shocks

A seconda che il processo sia stazionario AR oppure RW, RWD o TD, gli effetti di uno shock C al tempo t hanno effetti differenti sulla dinamica futura della variabile.

$$\text{AR}(1): Y_t = \varphi_1 Y_{t-1} + a_t \quad |\varphi_1| < 1$$

Al tempo $t+1$ si registra un effetto $C \phi$,
al tempo $t+2$ si avrà un effetto $C \phi^2$

....

al tempo $t+k$ si avrà un effetto $C \phi^k$

L'effetto si smorza nel tempo

Nonstazionarietà e shocks

Nel caso di processo RW e RWD l'effetto dello shock è permanente:

$$\text{RW: } Y_t = Y_{t-1} + a_t$$

Al tempo t+1 si registra un effetto C,
al tempo t+2 si avrà un ulteriore effetto C

....

al tempo t+k si avrà un effetto C

L'effetto si cumula nel tempo e permane in ogni

→ shock monetari

Nel caso di trend deterministico, gli effetti dello shock si comportano come per un AR(1).

Lo shock agisce solo sulla parte stocastica, che si aggiunge a quella deterministica.

Ispezione preliminare della non stazionarietà

Modelli MA sono sempre stazionari.

Modelli AR(p) per essere stazionari "*devono avere le soluzioni dell'equazione caratteristica fuori dal cerchio unitario*"

devono essere soddisfatte delle condizioni sui coefficienti autoregressivi, tra cui:

$$\rightarrow \phi_1 + \phi_2 + \dots + \phi_p < 1$$

Trend

- TREND DETERMINISTICO:

$$X_t = \mu + \beta t + a_t$$

- TREND STOCASTICO:

$$X_t = \mu + X_{t-1} + a_t \quad (\text{RWD})$$

- che viene a

$$X_t = X_{t-1} + a_t \quad (\text{RW})$$

Trend stocastico

- Se si sostituisce iterativamente il valore di X_t nel RW, si ha:

$$\begin{aligned} X_t &= X_{t-1} + a_t = (X_{t-2} + a_{t-1}) + a_t = \dots = \\ &= X_0 + a_1 + \dots + a_t = X_0 + \sum_i^t a_i \end{aligned}$$

- Mentre per RWD:

$$\begin{aligned} X_t &= \mu + X_{t-1} + a_t = \mu + (\mu + X_{t-2} + a_{t-1}) + a_t = \dots = \\ &= X_0 + \mu t + a_1 + \dots + a_t = X_0 + \mu t + \sum_i^t a_i \end{aligned}$$

Trend stocastico e deterministico

Se interpretiamo con una retta deterministica un trend stocastico:

$$X_t - \beta t = X_0 + \sum a_i$$

quindi la parte di breve periodo è ancora non stazionaria (perchè la varianza cresce con t)

Trend stocastico e deterministico

Se interpretiamo con un RW un trend deterministico:

$$\begin{aligned}\nabla X_t &= X_t - X_{t-1} = (\beta t + a_t) - [\beta(t-1) + a_{t-1}] = \\ &= \beta + a_t - a_{t-1}\end{aligned}$$

quindi la parte di breve periodo verrebbe spiegata tramite un MA(1) con coefficiente 1. Il processo è stazionario ma non invertibile (problemi in fase di previsione).

Trend stocastico e deterministico

Quindi, al fine di avere una corretta specificazione del processo generatore della serie, nelle sue dinamiche di breve e di lungo termine, è necessario capire correttamente quale sia il meccanismo sottostante la dinamica.

LE due tipologie di non stazionarietà devono essere correttamente individuate per avere un modello corretto che spieghi il fenomeno nel suo complesso nel tempo

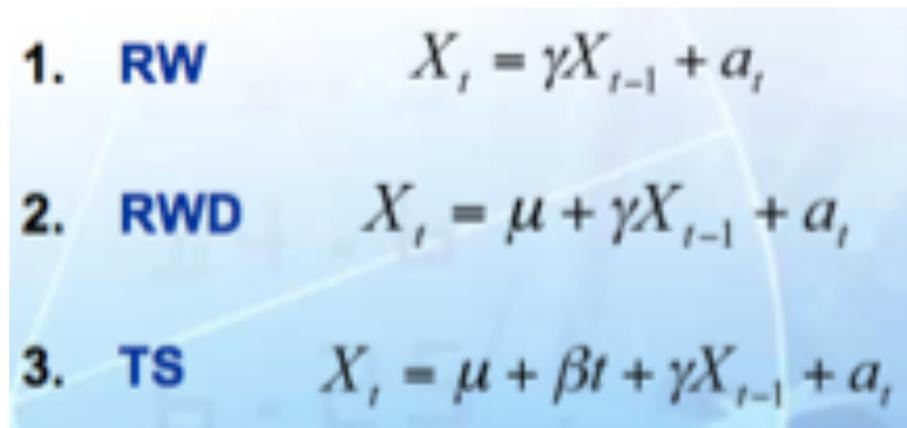


Test di non stazionarietà

Test non stazionarietà

In presenza di radici unitarie e non stazionarietà i classici test sui coefficienti non sono validi e non si applicano le distribuzioni standard (t, F, etc.).

Dickey e Fuller (1981) hanno specificato le distribuzioni per i test sui seguenti tre modelli di non stazionarietà:



The image shows three regression models for non-stationarity tests, each with a label and a corresponding equation. The labels are in blue and bold, and the equations are in black. The background is a light blue grid.

- 1. RW** $X_t = \gamma X_{t-1} + a_t$
- 2. RWD** $X_t = \mu + \gamma X_{t-1} + a_t$
- 3. TS** $X_t = \mu + \beta t + \gamma X_{t-1} + a_t$

Non stazionarietà

A convenient reformulation of the nonstationary (1)-(3) models is:

1. **RW** $\Delta X_t = (\gamma - 1)X_{t-1} + a_t = \gamma^* X_{t-1} + a_t$

2. **RWD** $\Delta X_t = \mu + (\gamma - 1)X_{t-1} + a_t = \mu + \gamma^* X_{t-1} + a_t$

3. **TS** $\Delta X_t = \mu + \beta t + (\gamma - 1)X_{t-1} + a_t = \mu + \beta t + \gamma^* X_{t-1} + a_t$

For (1) model: $H_0 : \gamma^* = 0$

For (2) model: $H_0 : \gamma^* = 0 \quad \mu = 0$

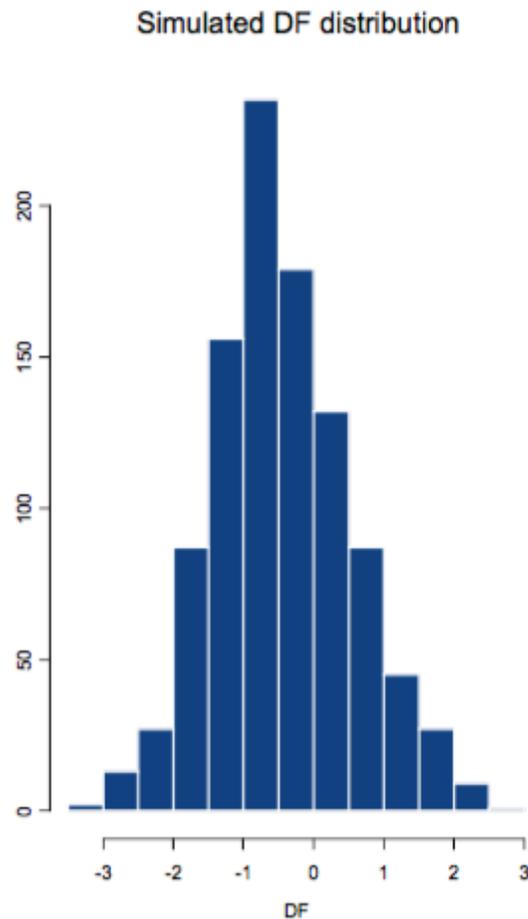
For (3) model: $H_0 : \beta = 0 = \gamma^*$

Test di stazionarietà

- Nei test DF, ADF e PP, l'ipotesi nulla è la NON STAZIONARIETA'
- Quindi se il p-value > 0.05 e accetto l'ipotesi nulla, significa che la mia serie è non stazionaria
- In quel caso devo "trasformare" la serie, in modo da forzarla ad essere stazionaria.

Test di stazionarietà

- In realtà il test non ha una distribuzione "standard" e i valori critici e i p-value devono essere determinati per simulazione.



Test di stazionarietà

- Per introdurre i test **UNIT ROOTS** riprendiamo il modello AR(1):

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

- L'ipotesi che vogliamo verificare, per la stazionarietà, è:

$$H_0 : \phi = 1 \text{ (unit root in } \phi(z) = 0) \Rightarrow y_t \sim I(1)$$

$$H_1 : |\phi| < 1 \Rightarrow y_t \sim I(0)$$

Test di stazionarietà

- In realtà il modello viene riscritto:

$$Y_t - Y_{t-1} = (1 - L)Y_t = \phi Y_{t-1} - Y_{t-1} + \varepsilon_t$$
$$\Delta Y_t = (\phi - 1)Y_{t-1} + \varepsilon_t \quad \varepsilon_t \approx WN$$

- e quindi le ipotesi da verificare:

$$H_0: (\phi - 1) = 0$$

$$H_1: (\phi - 1) < 0$$

Test di stazionarietà

- Nel caso di RWD, per introdurre il rispettivo test **UNIT ROOTS** riprendiamo il modello AR(1):

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t \quad \varepsilon_t \approx WN$$

- L'ipotesi che vogliamo verificare, per la stazionarietà, è:

$$H_0 : \phi = 1 \Rightarrow y_t \sim I(1) \text{ without drift}$$

$$H_1 : |\phi| < 1 \Rightarrow y_t \sim I(0) \text{ with nonzero mean}$$

Test di stazionarietà

- E infine, nel caso di presenza di TD, il modello AR(1):

$$Y_t = c + \beta t + \phi Y_{t-1} + \varepsilon_t \quad \varepsilon_t \approx WN$$

- L'ipotesi che vogliamo verificare, per la stazionarietà, è:

H_0 : $\phi = 1 \Rightarrow y_t \sim I(1)$ with drift

H_1 : $|\phi| < 1 \Rightarrow y_t \sim I(0)$ with deterministic time trend

Test unit roots

Riformuliamo i tre modelli nel seguente modo:

1. RW $\Delta X_t = (\phi - 1)X_{t-1} + a_t$
2. RWD $\Delta X_t = \mu + (\phi - 1)X_{t-1} + a_t$
3. TS $\Delta X_t = \mu + \beta t + (\phi - 1)X_{t-1} + a_t$

Per (1): $H_0: (\phi - 1) = 0$

Per (2): $H_0: (\phi - 1) = 0 \quad \mu = 0$

Per (3): $H_0: \beta = 0 = (\phi - 1)$

Test di stazionarietà

TEST1 e TEST2:

Ipotesi nulla → La serie deve essere differenziata (ΔX) per diventare stazionaria

Ipotesi alternativa → La serie è stazionaria e non deve essere trasformata

TEST3:

Ipotesi nulla → La serie deve essere differenziata (ΔX) per diventare stazionaria

Ipotesi alternativa → La serie è stazionaria, **MA SI DEVE TOGLIERE UN TREND DETERMINISTICO**

Test unit roots

R Resources

- Function `ur.df` in package `urca`.
- Function `ADF.test` in package `uroot`.
- Function `adf.test` in package `tseries`.
- Function `urdfTest` in package `fUnitRoots`.

Literature

- Dickey, D. and W. Fuller, Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Society*, 74 (1979), 427–341.
- Dickey, D. and W. Fuller, Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica*, 49, 1057–1072.
- Fuller, W., *Introduction to Statistical Time Series*, 2nd Edition, 1996, New York: John Wiley.
- MacKinnon, J., Numerical Distribution Functions for Unit Root and Cointegration Tests, *Journal of Applied Econometrics*, 11 (1996), 601–618.

Test ADF

I test introdotti da Dickey e Fuller nel 1981 vennero poi estesi, sempre dagli stessi autori, introducendo nel modello da sottoporre a verifica oltre alla variabile ritardata di 1 tempo, altre ritardate



Test ADF

(stessa logica di base dei test DF)

ARIMA

- Un processo stazionario viene detto anche "integrato di ordine 0" e indicato con:

$$X_t \approx I(0)$$

Definizione

- Un processo è detto "integrato di ordine d" se è necessario applicare d volte l'operatore differenza per ottenere la stazionarietà:

$$\nabla^d X_t \approx I(0) \rightarrow X_t \approx I(d)$$

ARIMA

- Se una serie è stazionaria, non è necessario spiegare nel modello la non stazionarietà (il trend) e possiamo descrivere la dinamica con i modelli ARMA.
- Se la serie presenta trend (stocastico) dobbiamo includerlo nel modello:



ARIMA

- Il modello ARIMA riesce quindi a spiegare processi non stazionari.
- ARIMA(p,d,q):

$$\begin{aligned} \nabla X_t &= \phi_1 \nabla X_{t-1} + \phi_2 \nabla X_{t-2} + \dots + \phi_p \nabla X_{t-p} + a_t \\ &\quad - \theta_1 a_{t-1} + \dots - \theta_q a_{t-q} \end{aligned}$$

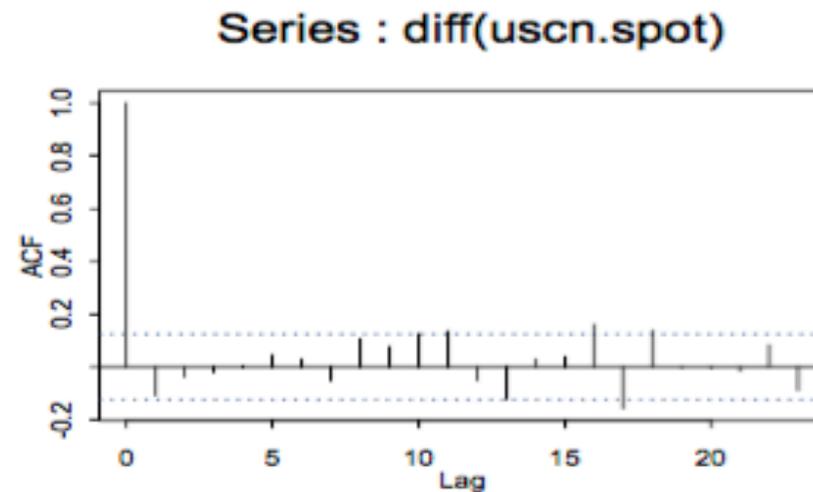
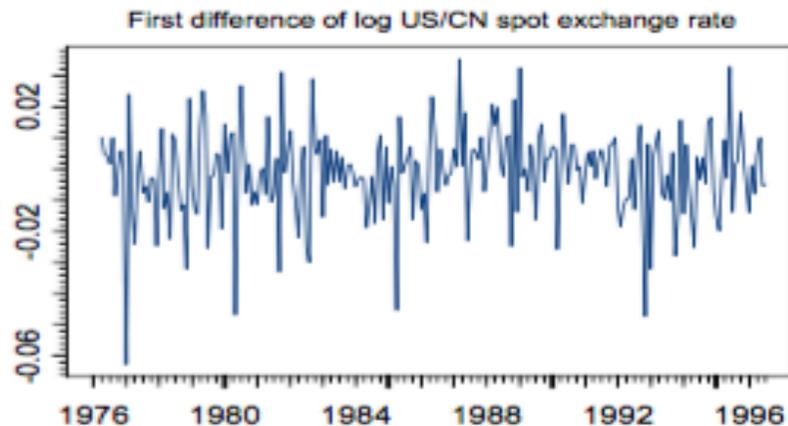
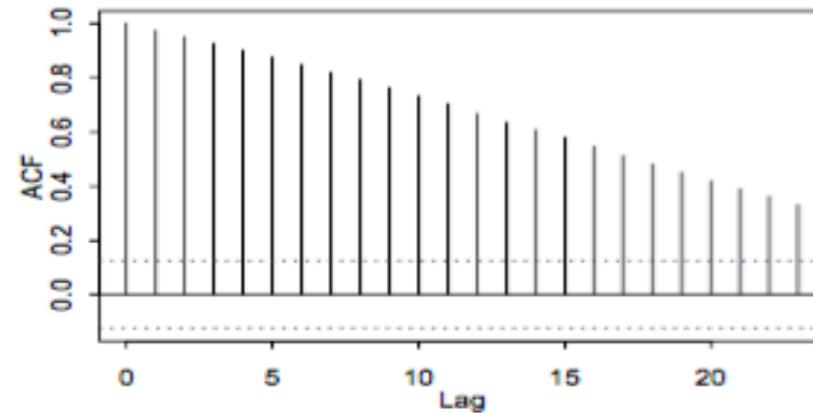
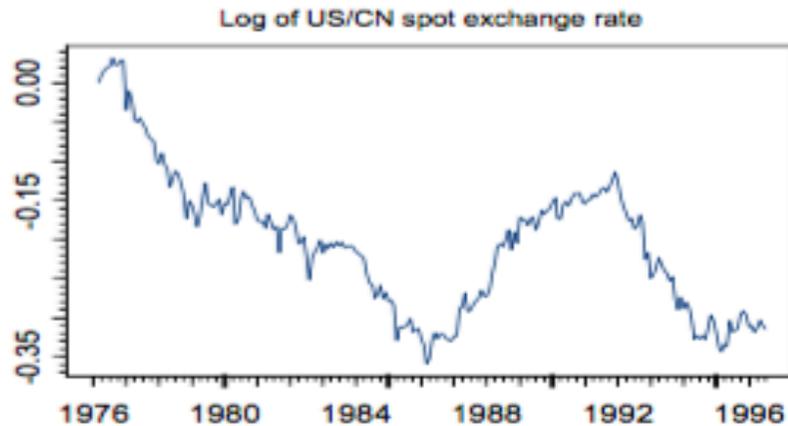
$$\text{ARIMA}(1,1,1): \nabla X_t = \phi_1 \nabla X_{t-1} + a_t - \theta_1 a_{t-1}$$

$$\text{ARI}(1,1,0): \nabla X_t = \phi_1 \nabla X_{t-1} + a_t$$

$$\text{IMA}(0,1,1): \nabla X_t = a_t - \theta_1 a_{t-1}$$

Esempio: tassi di cambio

- Il logaritmo del US/tasso di cambio mensile in un periodo di 30 anni 1976 - 1996



Esempio: tassi di cambio

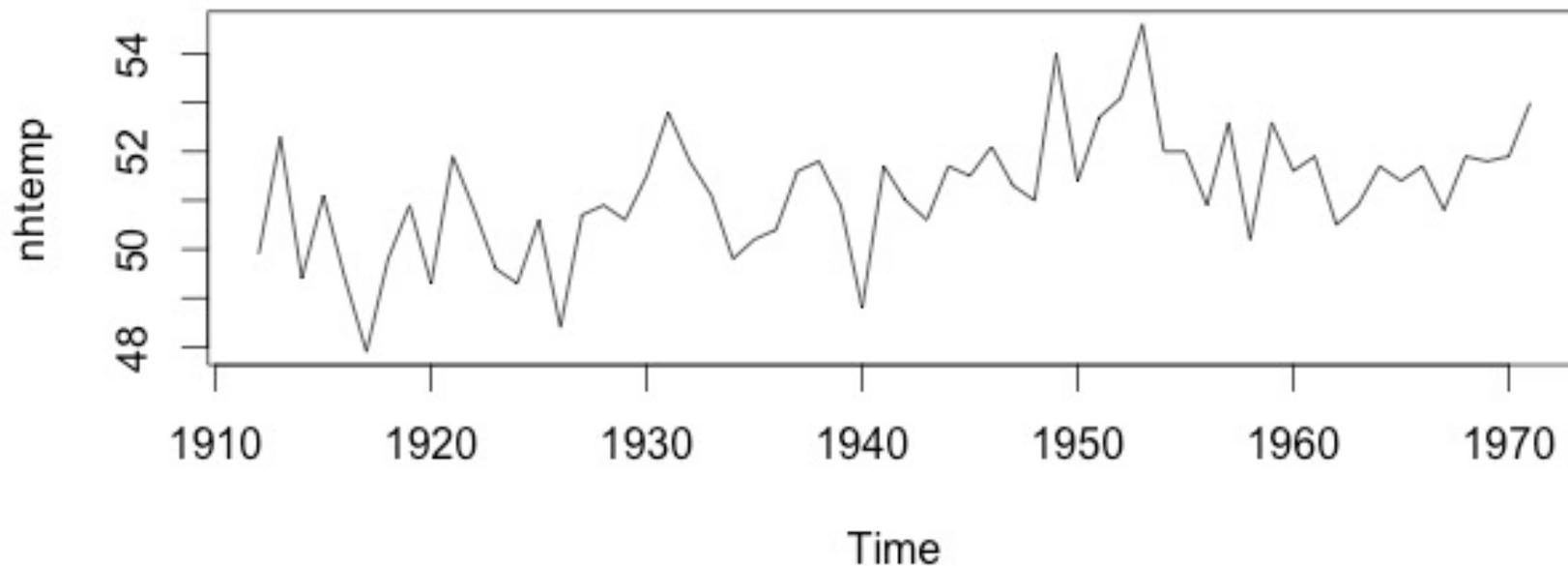
- Calcolo del test ADF

```
> summary(adft.out)
Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root
Type of Test: t test
Test Statistic: -2.6
P-value: 0.09427
```

ESEMPIO 2

- Sia data la seguente serie "Average Yearly Temperatures in New Haven 1912 to 1971"



Esempio: tassi di cambio

- Applico il test ADF con R:

```
> adf.test(nhtemp)
```

```
Augmented Dickey-Fuller Test
```

```
data: nhtemp
```

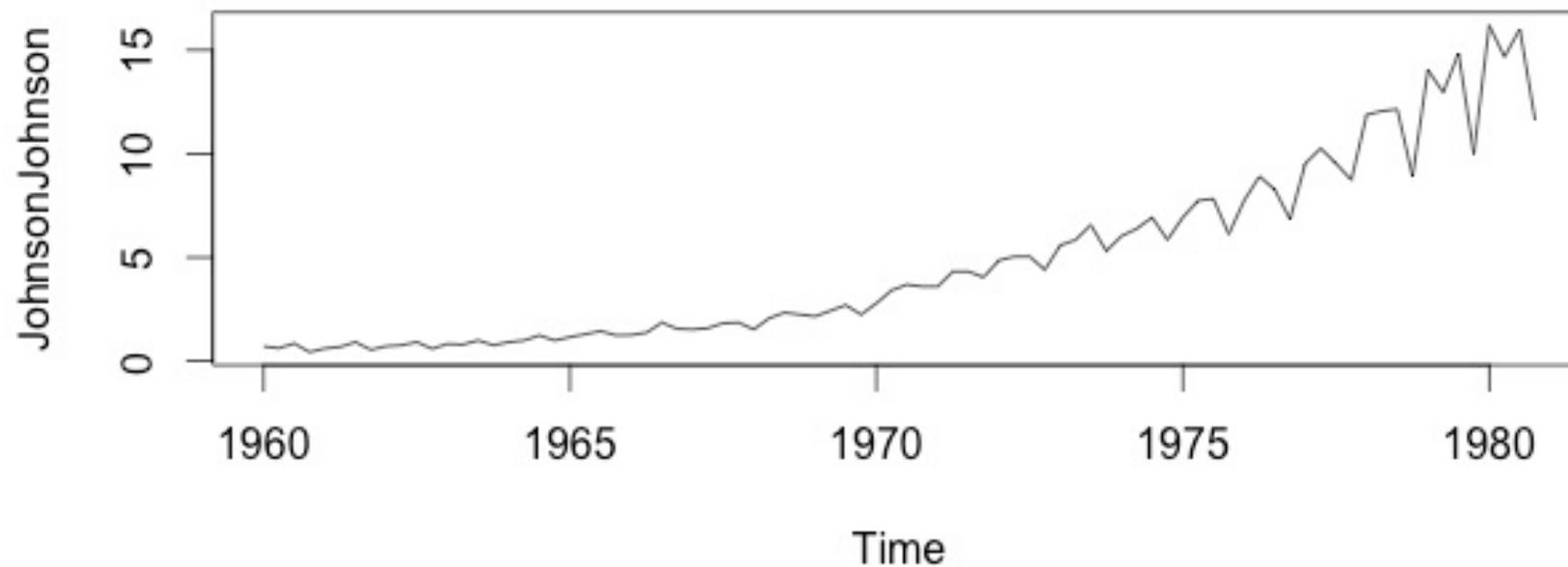
```
Dickey-Fuller = -3.2773, Lag order = 3, p-value = 0.08376
```

```
alternative hypothesis: stationary
```

```
■
```

Esempio: utili azionari

- Utili trimestrali (dollari) per azione Johnson & Johnson 1960-80



Esempio: utili azionari

- Applico il test ADF con R:

```
> adf.test(nhtemp)
```

Augmented Dickey-Fuller Test

data: JohnsonJohnson

Dickey-Fuller = -2.4267, Lag order = 4, p-value = 0.56321

alternative hypothesis: stationary

Test Unit Roots

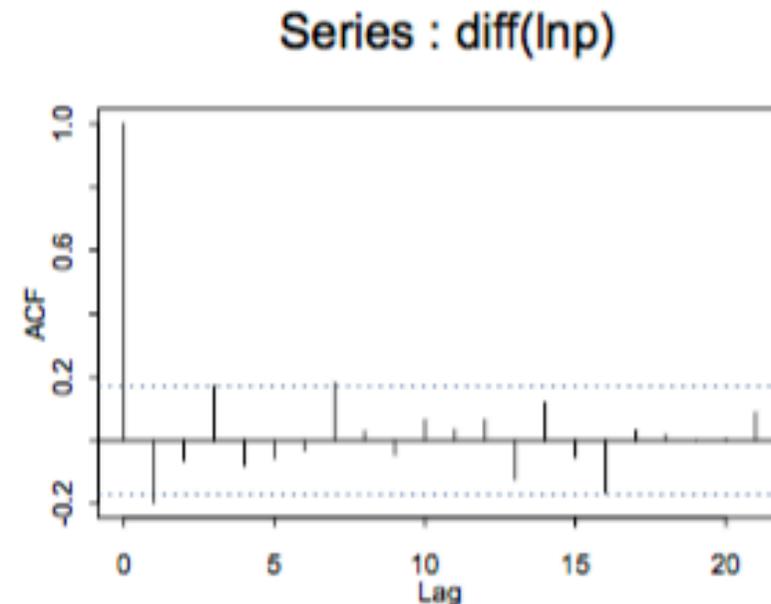
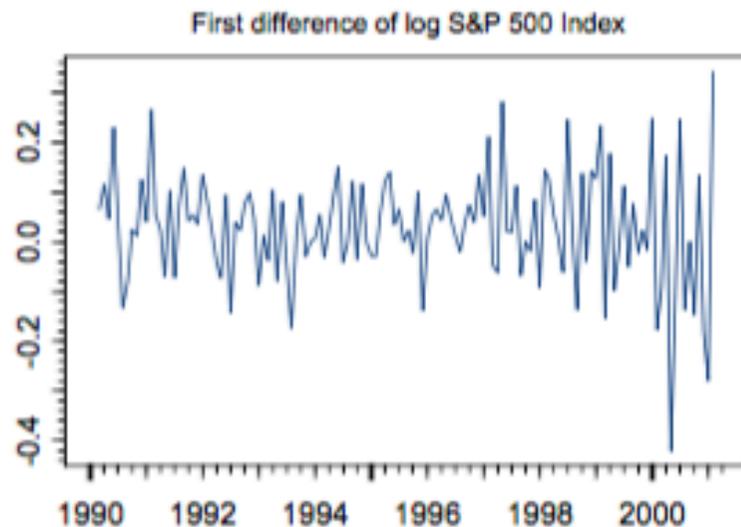
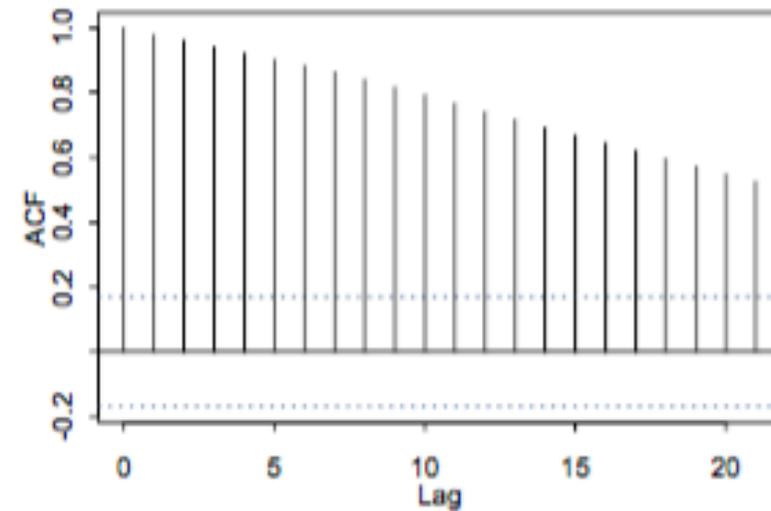
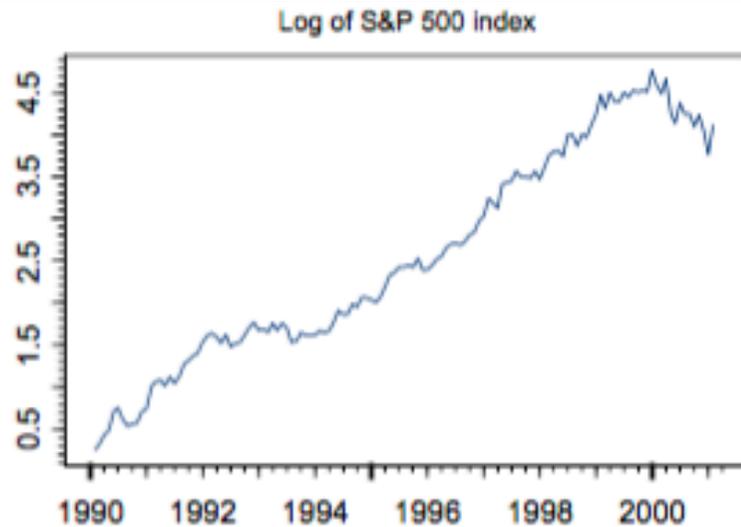
- Applico il test ADF con R:

Nel pacchetto **tseries** ritroviamo la funzione `adf.test(...)`

Tale funzione trova in automatico i ritardi da inserire (per passare da DF a ADF)

"The general regression equation which incorporates a constant and a linear trend is used and the t-statistic for a first order autoregressive coefficient equals one is computed. The number of lags used in the regression is k."

Esempio: Log indice S&P 500 Gennaio 1990 Gennaio 2001



Esempio: Log indice S&P 500 Gennaio 1990 Gennaio 2001

```
> adft.out = unitroot(lnp, trend="ct", lags=4)
> summary(adft.out)
```

Test for Unit Root: Augmented DF Test

Null Hypothesis: there is a unit root

Type of Test: t test

Test Statistic: -1.315

P-value: 0.8798

Test Unit Roots

- Applico il test ADF con R:

Nel pacchetto **urca** ritroviamo la funzione `ur.df(...)`

Tale funzione permette di eseguire distintamente tutti e tre i test precedenti, sia nella versione DF che ADF (lag = 0 corrisponde al test DF).

```
library(urca)
df1 <- ur.df(y, type = "none", lags = 0)
summary(df1)
```

Test Unit Roots

- INTERPRETAZIONE OUTPUT

Test regression none

Call: `lm(formula = z.diff ~ z.lag.1 - 1)`

	Estimate	Std. Error	t value	Pr(> t)
z.lag.1	0.003321	0.002405	1.381	0.168

Value of test-statistic is: 1.3808

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.58	-1.95	-1.62

Test Unit Roots

- INTERPRETAZIONE OUTPUT:

Si deve confrontare il valore della statistica test con i valori critici riportati in tabella.

Nel nostro caso statistica test ADF = 1.3808

Valori critici: $vc_{95\%} = -1.95$

Se $ADF < cv$ si rifiuta H_0

Se $ADF > cv$ si accetta H_0

Nel nostro caso si accetta H_0 → serie non stazionaria

Test Unit Roots

- INTERPRETAZIONE OUTPUT:

Nella funzione `ur.df` specificare **type**

1. `type = "none"` : RW
2. `type = ""drift"` : RWD
3. `type = "trend"` : Deterministic trend

Test Unit Roots

```
library(urca)
```

```
df2 <- ur.df(y2, type = "drift", lags = 0)
```

Value of test-statistic is: -1.8096 48.2151

Critical values for test statistics:

	1pct	5pct	10pct
tau2	-3.44	-2.87	-2.57
phi1	6.47	4.61	3.79

stat 1 = -1.8096 per verificare $\phi = 1$

stat1 > cv: si accetta H_0 ($\phi = 1$)

stat 2 = 48.2151 per verificare la costante diversa da zero

stat2 > cv: si accetta che costante diversa da zero

RWD

Test Unit Roots

```
library(urca)
```

```
df3 <- ur.df(y3, type = "trend", lags = 0)
```

Value of test-statistic is: -17.8381 144.9083 159.099

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-3.98	-3.42	-3.13
phi2	6.15	4.71	4.05
phi3	8.34	6.30	5.36

tau serve per la radice unitari

phi1 e phi2 per la significatività della costante e del trend bt

Qui si rifiuta per tau e si accetta per phi1 e phi2

= no radice unitaria, si trend deterministico

ACF and PACF for Z_t and ΔZ_t

Z_t

k:	1	2	3	4	5	6	7	8	9	10	11	12	13	14
ACF:	0.972	0.952	0.932	0.911	0.896	0.877	0.858	0.843	0.831	0.817	0.805	0.791	0.775	0.763
PACF:	0.972	0.138	0.010	-0.022	0.091	-0.049	-0.028	0.049	0.087	-0.033	0.034	-0.040	-0.040	0.035

ΔZ_t

k:	1	2	3	4	5	6	7	8	9	10	11	12	13	14
ACF:	-0.347	0.004	-0.040	-0.064	0.078	-0.034	-0.018	-0.049	0.135	-0.059	0.031	0.046	-0.127	0.087
PACF:	-0.347	-0.132	-0.097	-0.132	-0.002	-0.024	-0.049	-0.093	0.097	0.011	0.026	0.090	-0.065	0.013

Significance = ± 0.126

ARIMA(0,1,1)

structural break

ATTENZIONE

DF e ADF (unit root test in generale) possono suggerire la presenza di non stazionarietà, mentre si è in presenza di structural break

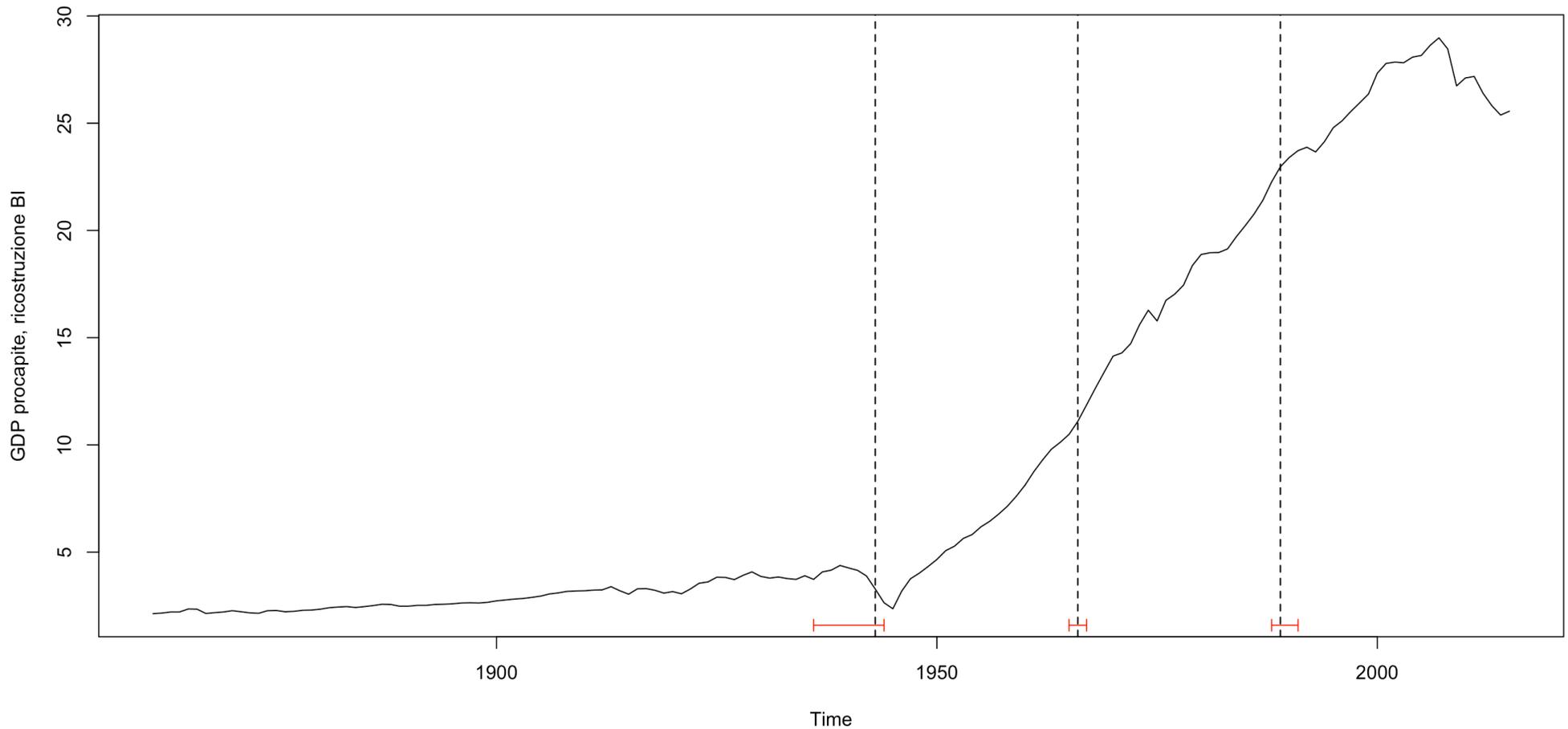
structural break: La ricostruzione Banca d'Italia-Istat 1861-2015 (da Statistiche storiche della Banca d'Italia)

```
library(strucchange)
bptsgdp_it_bi<-breakpoints(tsgdp_it_bi[,3]~1)
summary(bptsgdp_it_bi)
ci_bptsgdp_it_bi <- confint(bptsgdp_it_bi)
par(mfrow=c(1,1),mar=c(4.5,4.5,1,1))
plot(tsgdp_it_bi[,3],main="",ylab="GDP
procapite, ricostruzione BI")
lines(bptsgdp_it_bi)
lines(ci_bptsgdp_it_bi)
```

structural break: La ricostruzione Banca d'Italia-Istat 1861-2015 (da Statistiche storiche della Banca d'Italia)

Optimal (m+1)-segment partition:	CI for breakpoints
Breakpoints at observation number:	of optimal 4-segment
m = 1	partition:
m = 2	Breakpoints at obs. number:
m = 3	2.5 % breakpoints 97.5 %
m = 4	1 76 83 84
m = 5	2 105 106 107
Corresponding to breakdates:	3 128 129 131
m = 1	Corresponding to breakdates:
m = 2	2.5 % breakpoints 97.5 %
m = 3	1 1936 1943 1944
m = 4	2 1965 1966 1967
m = 5	3 1988 1989 1991
Fit:	
m	0 1 2 3 4 5
RSS	13028.6 1579.2 602.2 464.7 438.5 434.7
BIC	1136.8 819.8 680.5 650.4 651.5 660.2

structural break: La ricostruzione Banca d'Italia-Istat 1861-2015 (da Statistiche storiche della Banca d'Italia)



Stabilità

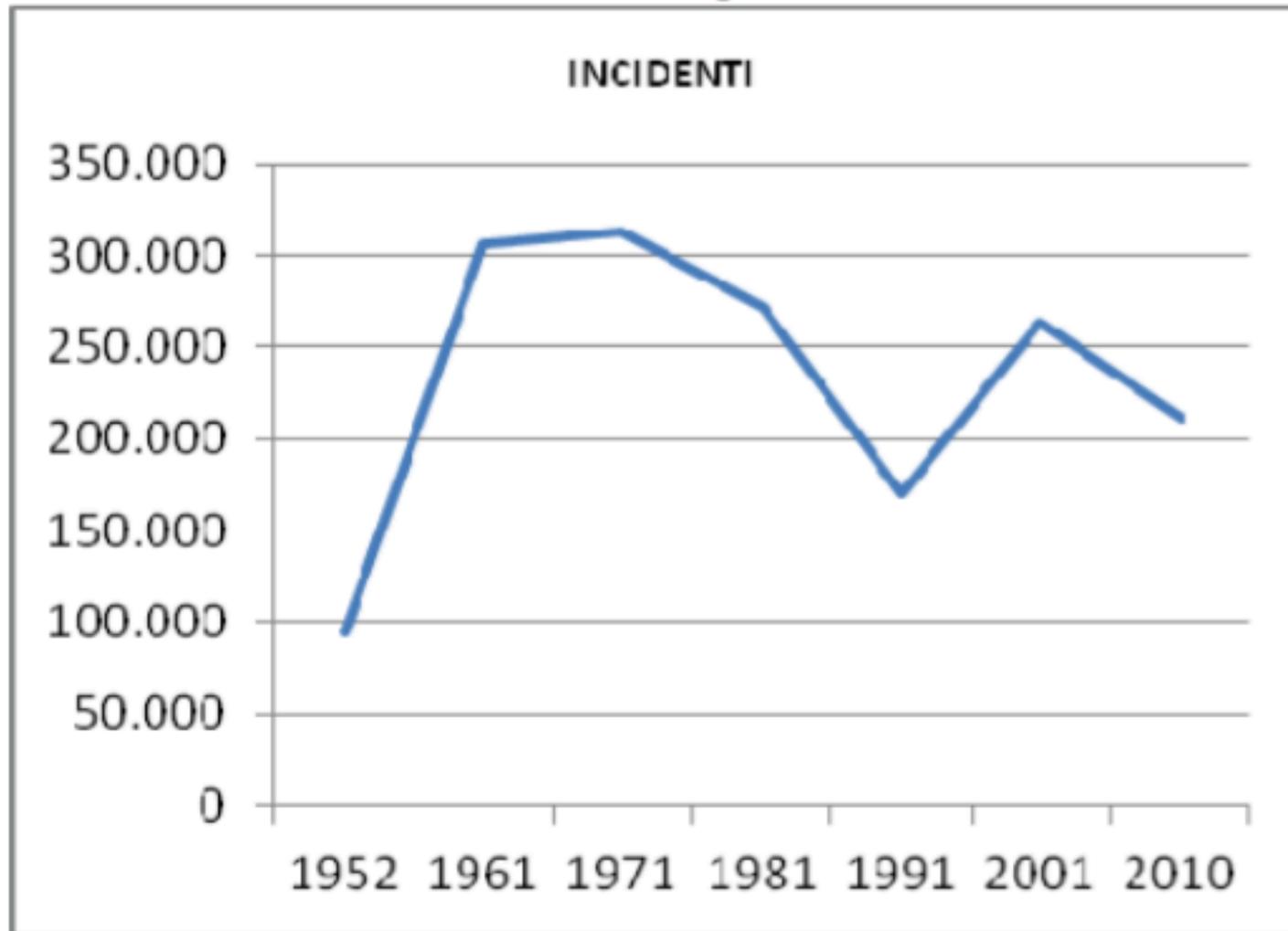
Per verificare la stabilità dei coefficienti lungo il campione osservato si può applicare il test CUSUM

$$CUSUM_{\tau} = \frac{\sum_{t=M}^{\tau} \hat{a}_t^2}{\hat{\sigma}_a} \quad \tau = M + 1, \dots, T$$

La statistica test viene normalmente rappresentata graficamente, in maniera iterata.

Stabilità

Graf. 4 Andamento del numero degli incidenti -Anni 1952-2010 -



Dati anomali in serie storica

- Cause che posso generare outliers
 - Scioperi
 - Guerra
 - Improvvise crisi economiche
 - Errori di digitazione
 - Improvvisi avvenimenti finanziari
- Se non considerati in modo opportuno nella stima, si possono avere residui non WN
- Gli effetti possono essere puntuali o permanenti

Dati anomali in serie storica

L'approccio classico ai valori anomali basato sui residui stimati è:

- Guarda i residui
- Elimina le osservazioni con residui grandi (± 2 s.e.)
- Stimare di nuovo l'equazione
- Vedere il cambiamento nel valore stimato di y se l'osservazione i_{th} viene eliminata

La previsione

***“Prediction is very difficult,
especially if it's about the future.”***

Nils Bohr

Cosa è ? E' il processo di previsione di un evento futuro

Alla base di molte decisioni aziendali

- Produzione
- Inventario
- Personale
- Strutture

Una previsione è valida solo come le informazioni incluse nella previsione (dati passati)

La storia non è un perfetto predittore del futuro (cioè: non esiste una previsione perfetta) quindi le previsioni raramente sono imperfette per loro stessa natura

La maggior parte delle tecniche presuppone una stabilità di base nel sistema

Le previsioni aggregate sono più accurate delle previsioni su singoli individui o prodotti

Importante definire l'orizzonte temporale della previsione (short, medium, long range)



Accuratezza della previsione

Errore di previsione = Valore Osservato(A) – Valore previsto(F)

$$MAD = \frac{\sum_{t=1}^n |A_t - F_t|}{n}$$
$$MSE = \frac{\sum_{t=1}^n (A_t - F_t)^2}{n-1} \approx \frac{\sum_{t=1}^n (A_t - F_t)^2}{n} = B^2 + \sigma^2$$
$$Tracking\ Signal = \frac{\sum_{t=1}^n A_t - F_t}{MAD}$$
$$RMSE = \sqrt{MSE}$$

La previsione : Metodi Qualitativi

Metodi Qualitativi

- Opinioni di un pool di esperti di alto livello, talvolta rafforzate da modelli statistici
- Metodo Delphi (gruppo di esperti, interrogato ripetutamente fino al raggiungimento del consenso)
- Consumer Market Survey (Chiedi al cliente)

Metodi Quantitativi

1. Naive approach
2. Medie Mobili
3. Livellamento Esponenziale
4. Estrapolazione del Trend
5. ARIMA
6. Regressione Lineare

Modelli di serie storiche

Modelli associativi

La previsione

I modelli ARIMA devono la loro principale diffusione alla ottima capacità di prevedere i fenomeni osservati in serie storiche.

I modelli deterministici sono troppo rigidi nelle loro formulazioni e facilmente "crollano" in fase previsiva

Le previsioni basate solo su valori passati, nessuna altra variabile è importante

Presuppongono che i fattori che influenzano il passato e il presente continueranno ad influenzare in future

Spesso si preferisce usare metodi semplice ed automatici in quanto possono risultare più "robusti" (e di conseguenza anche più conservativi), ad esempio i metodi deterministici sono molto diffusi in ambito aziendale

La previsione

Approccio Naive: Suppone che la variabile nel periodo successivo sia la stessa che nel periodo più recente. Ad esempio, se le vendite di gennaio fossero 68, le vendite di febbraio saranno 68. A volte economico ed efficiente. Può essere un buon punto di partenza

Medie mobili: sono una serie di medie aritmetiche (pesate e non). Usato se di tendenza scarsa o nulla. Utilizzato spesso per levigare, lisciare i dati (smoothing). Fornisce un quadro generale dei dati nel tempo. L'aumento di n uniforma la previsione ma la rende meno sensibile alle modifiche, non prevedere bene le tendenze e richiede molti dati.

Livellamento esponenziale (exponential smoothing) : Forma di media mobile ponderata. I pesi diminuiscono in modo esponenziale quindi i dati più recenti pesano di più. Richiede il calcolo di parametri di livellamento che possono essere fissati soggettivamente. Utilizza pochi dati passati

La previsione: Exponential Smoothing

**Nuova previsione = previsione dell'ultimo periodo
+ α (valore osservato dell'ultimo
periodo – previsione dell'ultimo
periodo)**

$$\ell_T = \ell_{T-1} + \alpha(y_{T-1} - \ell_{T-1})$$

dove ℓ_T = nuova previsione

ℓ_{T-1} = previsione precedente

**α = parametro di smoothing (or weighting)
($0 \leq \alpha \leq 1$)**

La previsione: Exponential Smoothing

- **Step 1:** Calcolare la stima iniziale della media della serie al tempo $t = 0$

$$\ell_0 = \bar{y} = \frac{\sum_{t=1}^n y_t}{n}$$

- **Step 2:** Calcolare la stima aggiornata secondo l'equazione di smoothing

$$\ell_T = \alpha y_T + (1 - \alpha)\ell_{T-1}$$

dove α è la costante di smoothing tra 0 ed 1.

Notare che

$$\begin{aligned}\ell_T &= \alpha y_T + (1 - \alpha)\ell_{T-1} \\ &= \alpha y_T + (1 - \alpha)\alpha y_{T-1} \\ &\quad + (1 - \alpha)^2 \alpha y_{T-2} + \dots + (1 - \alpha)^{T-1} \alpha y_1 + (1 - \alpha)^T \ell_0\end{aligned}$$

La previsione: Exponential Smoothing con Trend

Se assumiamo che esiste un trend nei dati dobbiamo modificare l'ES

Previsione con il trend (FIT_T) = **Previsione exponential smoothed** (ℓ_T) + **Trend exponential smoothed** (b_T)

$$\ell_T = \alpha y_T + (1 - \alpha)(\ell_{T-1} + b_{T-1})$$

$$b_T = \gamma(\ell_T - \ell_{T-1}) + (1 - \gamma)b_{T-1}$$

Step 1: calcolare ℓ_T

Step 2: calcolare b_T

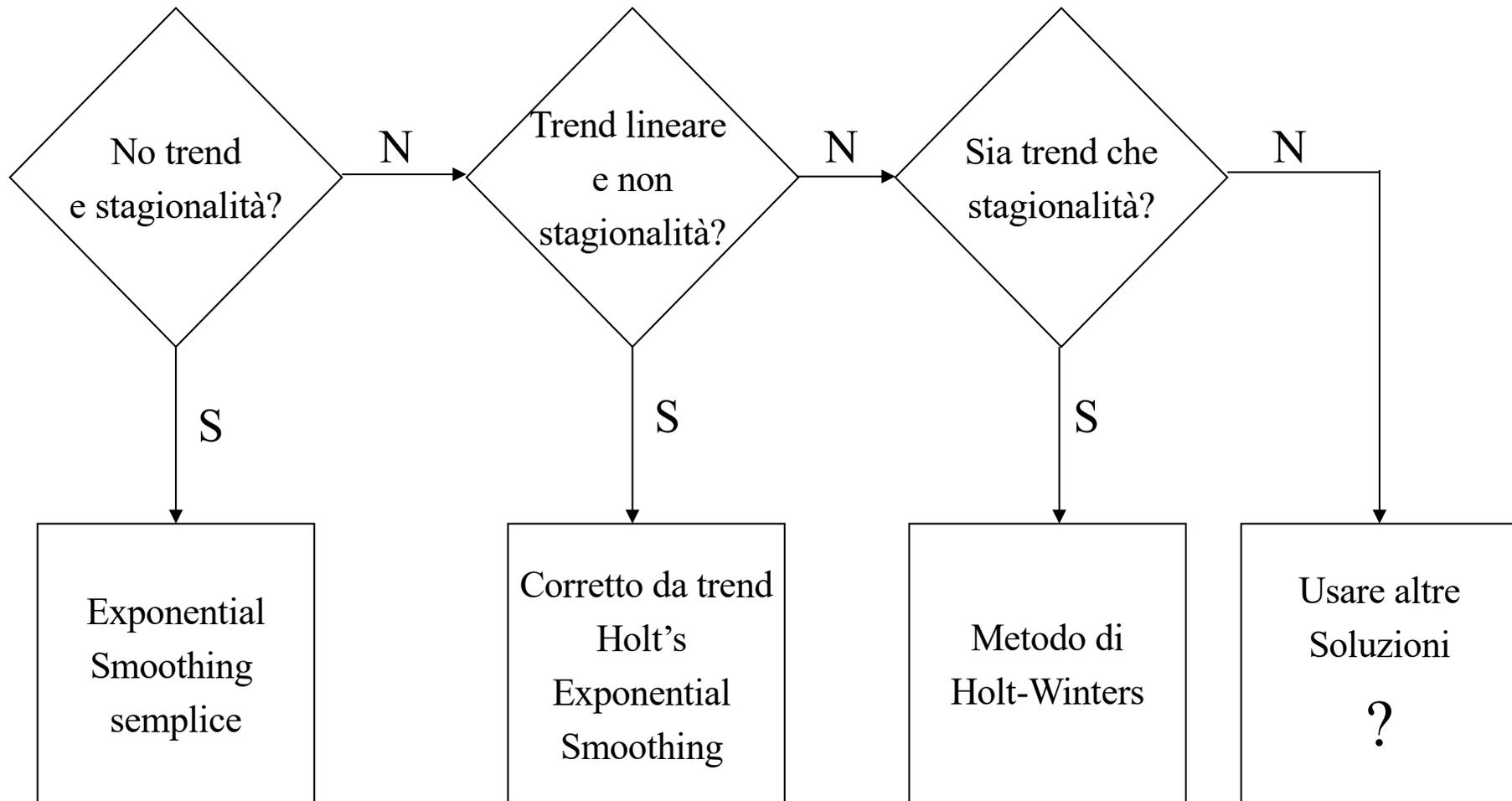
Step 3: calcolare la previsione $FIT_T = \ell_T + b_T$

NB : I parametri diventano due e $0 < \beta < 1$ è il tasso di crescita della serie

La previsione: Exponential Smoothing con Trend

- I metodi di livellamento esponenziale danno pesi maggiori alle osservazioni più recenti e i pesi diminuiscono esponenzialmente man mano che le osservazioni diventano più distanti.
- Questi metodi sono più efficaci quando i parametri che descrivono le serie temporali cambiano LENTAMENTE nel tempo.

La previsione: Estensione dell'Exponential Smoothing (Holt-Winters)



La previsione: Estensione dell'Exponential Smoothing (Holt-Winters)

Due metodi Holt-Winters sono progettati per le serie temporali che mostrano un trend lineare

- Metodo additivo Holt-Winters: utilizzato per le serie storiche con variazioni stagionali (additive) costanti
- Metodo moltiplicativo di Holt-Winters: utilizzato per serie storiche con variazioni stagionali (moltiplicative) crescenti

Il metodo Holt-Winters è un approccio di livellamento esponenziale per la gestione dei dati STAGIONALI.

Il metodo moltiplicativo di Holt-Winters è il più noto dei due metodi.

Holt-Winters Metodo moltiplicativo

- È generalmente considerato più adatto alla previsione di serie temporali che possono essere descritte dall'equazione:

$$y_t = (\beta_0 + \beta_1 t) \times SN_t \times IR_t$$

- SN_t : componente stagionale
- IR_t : componente irregolare (random)
- Questo metodo è appropriato quando una serie temporale ha un andamento lineare con un modello stagionale moltiplicativo per cui il livello $(\beta_0 + \beta_1 t)$, il tasso di crescita (β_1) e il modello stagionale (SN_t) possono cambiare lentamente nel tempo

Holt-Winters Metodo moltiplicativo

- Stima del livello

$$\ell_T = \alpha(y_T / sn_{T-L}) + (1 - \alpha)(\ell_{T-1} + b_{T-1})$$

- Stima del tasso di crescita (o trend)

$$b_T = \gamma(\ell_T - \ell_{T-1}) + (1 - \gamma)b_{T-1}$$

- Stima della componente stagionale

$$sn_T = \delta(y_T / \ell_T) + (1 - \delta)sn_{T-L}$$

dove α , γ , e δ sono parametri di smoothing tra 0 ed 1,

L = numero di Stagioni nell'anno ($L = 12$ per dati mensili, ed $L = 4$ per dati trimestrali)

Holt-Winters Metodo additivo

- È generalmente considerato più adatto alla previsione di serie temporali che possono essere descritte dall'equazione:

$$y_t = (\beta_0 + \beta_1 t) + SN_t + IR_t$$

- SN_t : componente stagionale
- IR_t : componente irregolare (random)
- Questo metodo è appropriato quando una serie temporale ha un andamento lineare con un modello stagionale additivo per cui il livello ($\beta_0 + \beta_1 t$), il tasso di crescita (β_1) e il modello stagionale (SN_t) possono cambiare lentamente nel tempo

Holt-Winters Metodo additivo

- Stima del livello

$$\ell_T = \alpha(y_T - sn_{T-L}) + (1 - \alpha)(\ell_{T-1} + b_{T-1})$$

- Stima del tasso di crescita (o trend)

$$b_T = \gamma(\ell_T - \ell_{T-1}) + (1 - \gamma)b_{T-1}$$

- Stima della componente stagionale

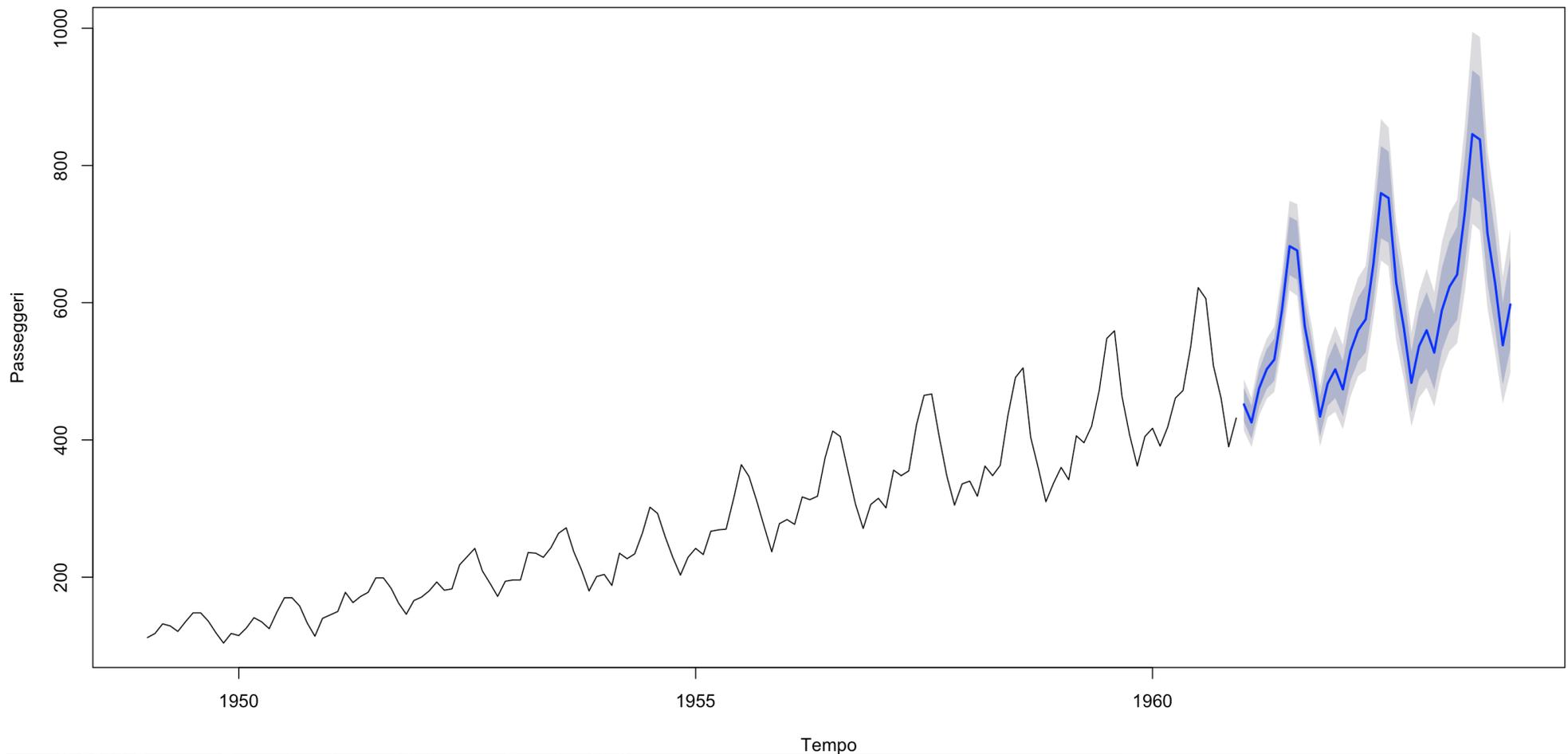
$$sn_T = \delta(y_T - \ell_T) + (1 - \delta)sn_{T-L}$$

dove α , γ , e δ sono parametri di smoothing tra 0 ed 1,

L = numero di Stagioni nell'anno ($L = 12$ per dati mensili, ed $L = 4$ per dati trimestrali)

Esempio Holt-Winters

```
library(forecast)
par(mar=c(4,4,1,1))
fcast <- hw(AP,h=36,seasonal = "multiplicative",exponential = T)
plot(fcast,main="",xlab="Tempo",ylab="Passeggeri",cex.lab=2)
```



Esempio Holt-Winters

Forecast method: Holt-Winters' multiplicative method with exponential trend

Model Information:

Holt-Winters' multiplicative method with exponential trend

Smoothing parameters:

alpha = 0.2436

beta = 1e-04

gamma = 0.5801

Initial states:

l = 120.7142

b = 1.0089

s = 0.8967 0.8420 0.9107 1.0208 1.2017 1.1931

1.0616 0.9573 1.0225 1.0638 0.9663 0.8636

sigma: 0.0416

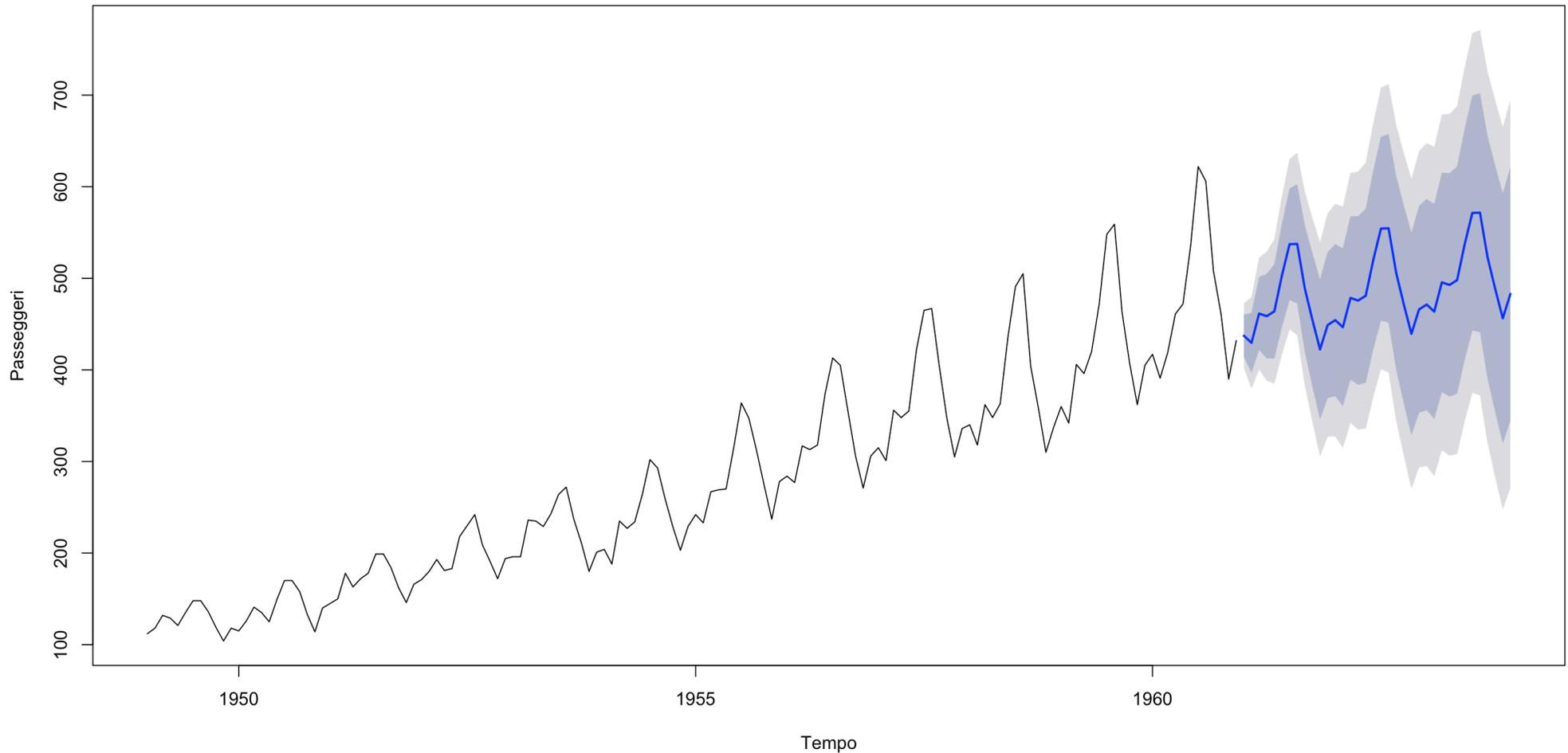
AIC	AICc	BIC
1412.498	1417.356	1462.985

Error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.4623683	10.60822	7.761685	0.1616139	2.937329	0.2423232	0.2717026

Esempio Holt-Winters

```
fcast <- hw(AP,h=36,seasonal = "additive",dumped = T)
plot(fcast,main="",xlab="Tempo",ylab="Passeggeri",cex.lab=2)
```



La previsione ARIMA

Formualzione del problema:

Data una serie storica osservata fino al tempo n , si vuole prevedere il suo valore $h=1, 2, \dots, k$ tempi in avanti

Il set informativo si ferma al tempo $t = n$

$$X_0, X_1, \dots, X_{n-1}, X_n$$

$$a_0, a_1, \dots, a_{n-1}, a_n$$

La previsione ARIMA

Indichiamo con:

X_{n+1} : il valore futuro al tempo $n+1$

$f_n(1)$: la previsione ad un tempo $h = 1$ di X , fatta sulla base delle informazioni al tempo n

.....

X_{n+k} : il valore futuro al tempo $n+k$

$f_n(k)$: la previsione ad un tempo $h = k$ di X , fatta sulla base delle informazioni al tempo n

Errore e MSE

L'errore che si commette con la previsione è dato da:

$$e_n(h) = X_{n+h} - f_n(h)$$

che per la previsione ad un tempo $h=1$ sarà:

$$e_n(1) = X_{n+1} - f_n(1)$$

Una misura sulla bontà della previsione su tutto l'intervallo di previsione è dato da MSE (*Mean Square Error*):

$$MSE = \sum_{h=1}^k e_n(h)^2 = \sum_{h=1}^k [X_{n+h} - f_n(h)]^2$$

Questo si può calcolare solo dopo che è noto X_{n+h}

Errore e MSE

Previsioni nel campione:

- di norma si lasciano fuori dal campione su cui stimare il modello gli ultimi dati, per verificare la bontà del modello stesso in fase di previsione.

Previsioni fuori del campione:

- Se si fanno le previsioni fuori dal campione, necessario attendere il verificarsi dell'evento per controllarne la bontà.

Previsione con ARIMA

Il valore $f_n(\mathbf{1})$ viene determinato sulla base del modello ARMA (o ARIMA) stimato sulla serie osservata fino al tempo n .

Supponiamo che su X si sia stimato un ARMA(p, q):

Allora:

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=1}^q \theta_i a_{t-i} \quad \theta_0 = 1$$

$$f_{n,h} = \sum_{i=1}^p \phi_i x_{n+h-i} + \sum_{i=1}^q \theta_i \hat{a}_{t-i} \quad \theta_0 = 1$$

Previsione con ARIMA

Dove si assume che:

$$f_n(h) = x_{n+h}$$
$$\hat{a}_{n+h} = \begin{cases} 0 & k > 0 \\ \hat{a}_{n+h} & k \leq 0 \end{cases}$$

Ossia i valori oltre il campo di osservazione n sono sostituiti con le loro previsioni e gli shock da n in poi sono posti pari a zero.

Previsione con ARIMA

Supponiamo che su X si sia stimato il seguente modello ARMA(2,0):

$$X_t = 61.895 + 1.351X_{t-1} - 0.461X_{t-2} + a_t$$

Allora i valori futuri di X (oltre il tempo n) sono dati da:

$$f_{n,h} = 61.895 + 1.351X_{n+h-1} - 0.461X_{n+h-2} + a_{n+h}$$

Previsione con ARIMA

Per fare previsioni oltre il tempo n basta conoscere i valori di X ai tempi $(n-1)$ e $(n-2)$.

Supponiamo di avere:

$$X_{n-1} = 17.87 \quad \text{e} \quad X_{n-2} = 22.95$$

Allora la previsione ad un tempo $h = 1$, $f_n(1)$ è data da:

$$f_n(1) = 61.895 + 1.351 \cdot 17.87 - 0.461 \cdot 22.95 = 75.457$$

Previsione con ARIMA

Per fare previsioni all'orizzonte temporale $h = 2$ usiamo il valore previsto $f_n(1)$:

$$f_n(2) = 61.895 + 1.351 \cdot 75.457 - 0.461 \cdot 17.87 = 155.589$$

E così a seguire, per le previsioni ad orizzonti temporali più in avanti ($h=2, 3, \dots$)

Aggiornamenti delle previsioni

Supponiamo che dopo aver fatto le previsioni per $h=1,2,3,\dots,k$ tempi avanti sia noto il vero valore di X_{n+1} . Come aggiusto tutte le previsioni già calcolate fino ad $h = k$?

Innanzitutto calcolo l'errore commesso per $h=1$:

E poi aggiusto le previsioni successive in base all'errore commesso.

$$e_n(1) = X_{n+1} - f_n(1)$$

Aggiornamenti delle previsioni

La previsione a due tempi ($h = 2$) sarà aggiustata:

$$f_{n+1}(1) = f_n(2) + c_1(X_{n+1} - f_n(1)) = f_n(2) + c_1 \cdot e_n(1)$$

dove il coefficiente c_1 è dato da in generale dai coefficienti stimati per l'ARMA(p,q):

$$c_1 = \phi_1 - \theta_1$$

Aggiornamenti delle previsioni

La previsione a tre tempi ($h = 3$) sarà anche essa aggiustata in base alle informazioni nuovi disponibili al tempo $n+1$::

$$f_{n+1}(2) = f_n(3) + c_2(X_{n+1} - f_n(1)) = f_n(3) + c_2 \cdot e_n(1)$$

dove il coefficiente c_2 è dato da:

$$c_2 = \phi_1 c_1 + \phi_2 - \theta_2$$

Aggiornamenti delle previsioni

In generale avremo quindi che le previsioni nuove, sulla base del dato $n+1$ sono ottenute dalle precedenti:

$$f_{n+1}(h) = f_n(h + 1) + c_h(X_{n+1} - f_n(1))$$

dove il coefficiente c_1 è dato da:

$$\begin{aligned} c_1 &= \phi_1 - \theta_1 \\ c_2 &= \phi_1 c_1 + \phi_2 - \theta_2 \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ c_j &= \phi_1 c_{j-1} + \phi_2 c_{j-2} + \cdots + \phi_p c_{j-p} - \theta_j \end{aligned}$$

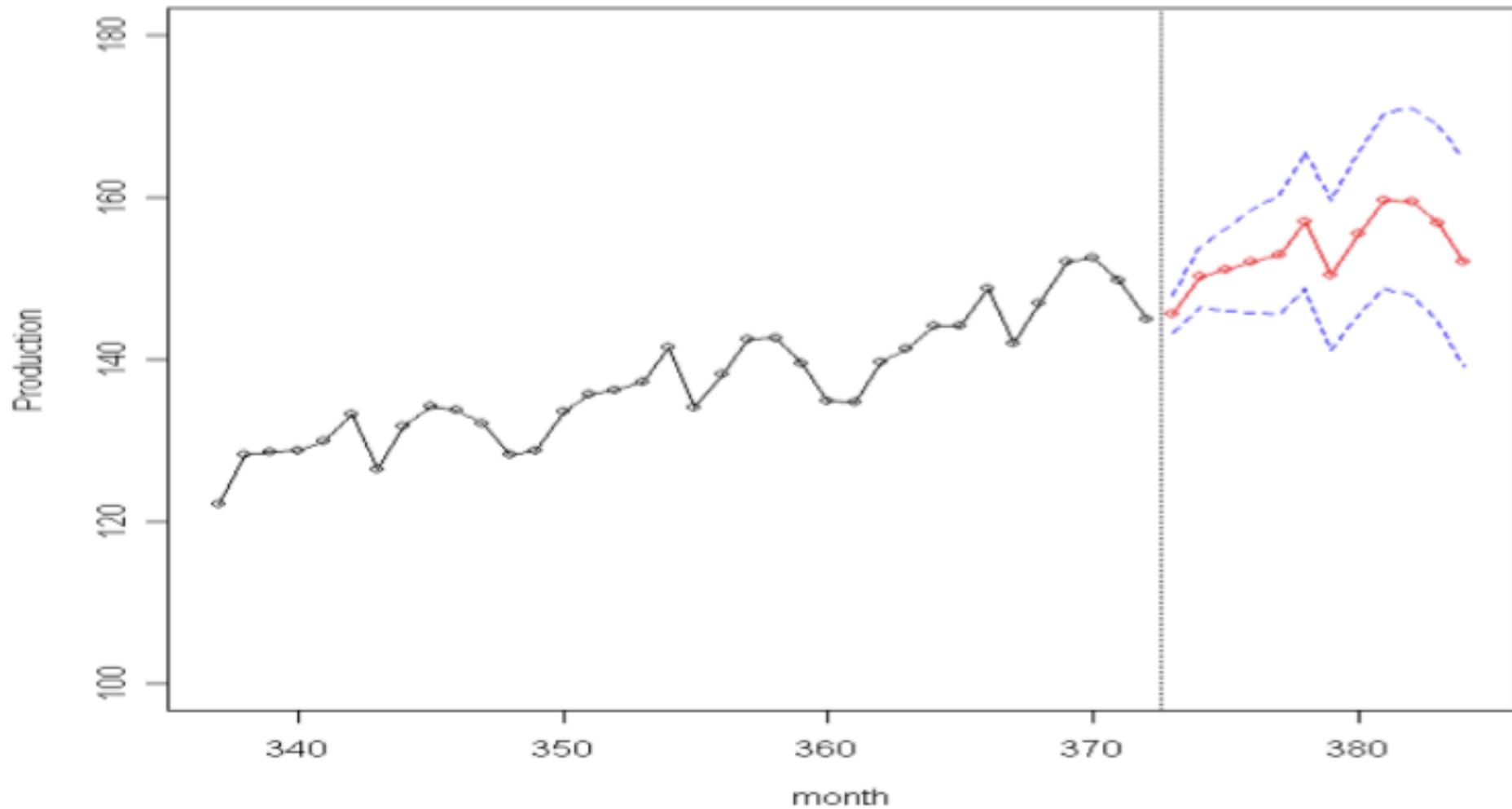
Intervalli di stima

Ai valori puntuali delle previsioni vanno associati gli intervalli di stima.

L'ampiezza dell'intervallo di stima cresce con l'aumentare dell'orizzonte temporale.

Al crescere di h , le previsioni diventano meno precise.

Intervalli di stima



Previsioni con R

1. Installare il pacchetto "forecast"
2. Stimare il modello arima
3. Salvare output della stima
4. usare il comando

```
forecast(serie, h=k)
```

5. plottare la previsione plot(...)
6. Vedere i valori con associati gli intervalli

Previsioni con R

```
> arima_lc <- arima(lc, order=c(1,1,0))  
> arima_lc
```

```
Call:  
arima(x = lc, order = c(1, 1, 0))
```

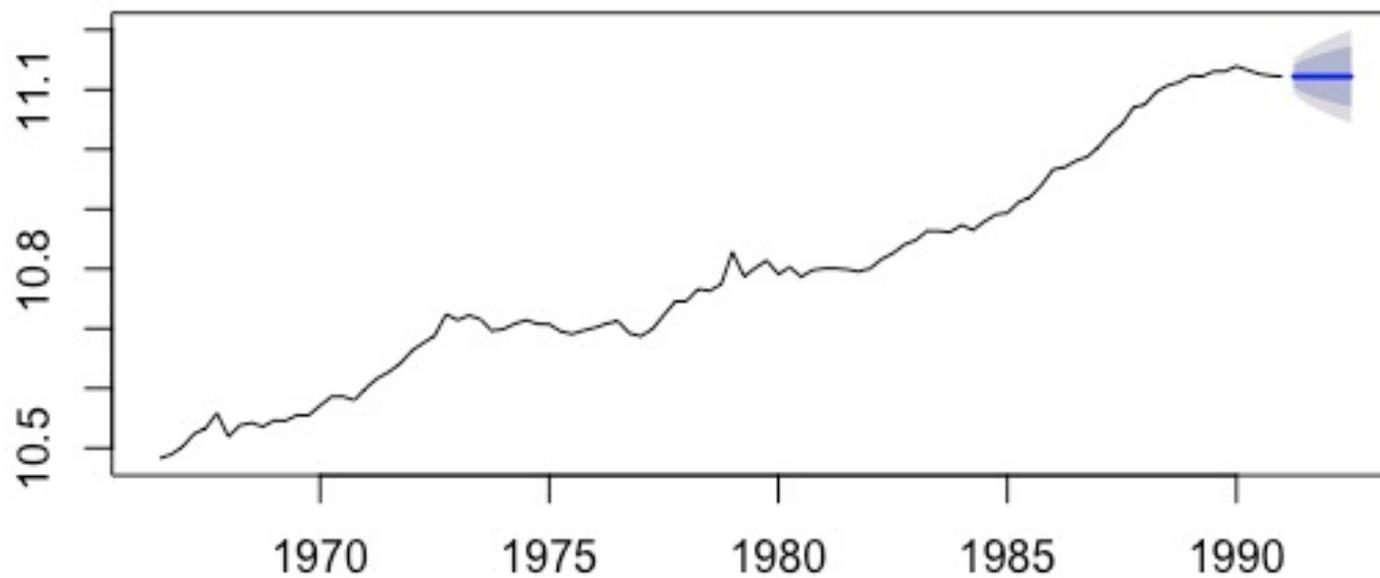
```
Coefficients:  
      ar1  
    0.0864  
s.e. 0.1002
```

```
sigma^2 estimated as 0.0002256:  log likelihood = 272.38,  aic = -540.76
```

Previsioni con R

```
> flc <- forecast(arima_lc,h=6)  
> plot(flc)
```

Forecasts from ARIMA(1,1,0)



Previsioni con R

```
> flc
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
1991 Q2	11.12190	11.10265	11.14115	11.09246	11.15134
1991 Q3	11.12189	11.09346	11.15031	11.07842	11.16536
1991 Q4	11.12189	11.08651	11.15726	11.06779	11.17599
1992 Q1	11.12189	11.08071	11.16306	11.05892	11.18485
1992 Q2	11.12189	11.07564	11.16814	11.05115	11.19262
1992 Q3	11.12189	11.07106	11.17271	11.04416	11.19962

Procedura completa ARIMA

1. grafico serie
2. ACF
3. PACF
4. ADF test
5. ur.df
6. Stazionario?

Procedura completa ARIMA

13. Analisi dei residui:

- normalità JB
- omoschedasticità ARCH
- ACF e PACF sui residui
- Test Ljung-Box o Box-Pierce su autocorrelazione
- analisi dati anomali
- test stazionarietà sui residui

14. Previsione

15. Grafico e valori della previsione

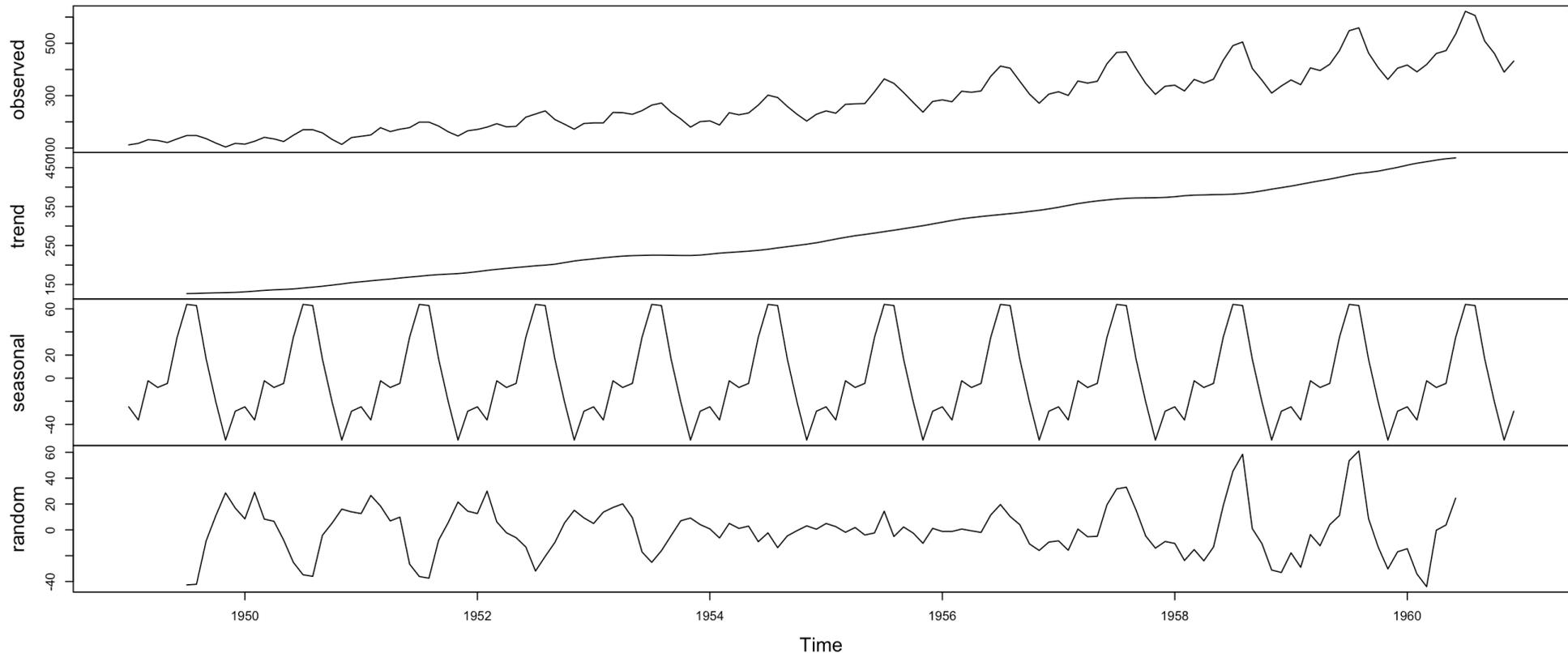
16. Se disponibili dati al tempo $n+1$, calcolo dell'errore e del MSE

Esempio: passeggeri aerei

```
library(tseries)
decAP <- decompose(AP, type = "additive")
plot(decAP)
adf.test(AP, alternative = "stationary", k=12)
findbest <- auto.arima(AP)
par(mfrow=c(1,1),mar=c(4.5,4.5,1,1))
plot(forecast(findbest,h=20))
fitAP <- arima(AP, order=c(0, 1, 1), list(order=c(0, 1, 0),
period = 12))
fore <- predict(fitAP, n.ahead=24)
U <- fore$pred + 2*fore$se
L <- fore$pred - 2*fore$se
ts.plot(AP, fore$pred, U, L, col=c(1, 2, 4, 4), lty=c(1,1,2,2))
legend("topleft", c("Actual", "Forecast", "Error Bounds (95%
prediction interval)"),col=c(1, 2, 4),lty=c(1, 1, 2))
```

Esempio: passeggeri aerei

Decomposition of additive time series



Augmented Dickey-Fuller Test

data: AP

Dickey-Fuller = -1.5094, Lag order = 12, p-value = 0.7807

alternative hypothesis: stationary

Esempio: passeggeri aerei

Series: AP

ARIMA(2,1,1)(0,1,0)[12]

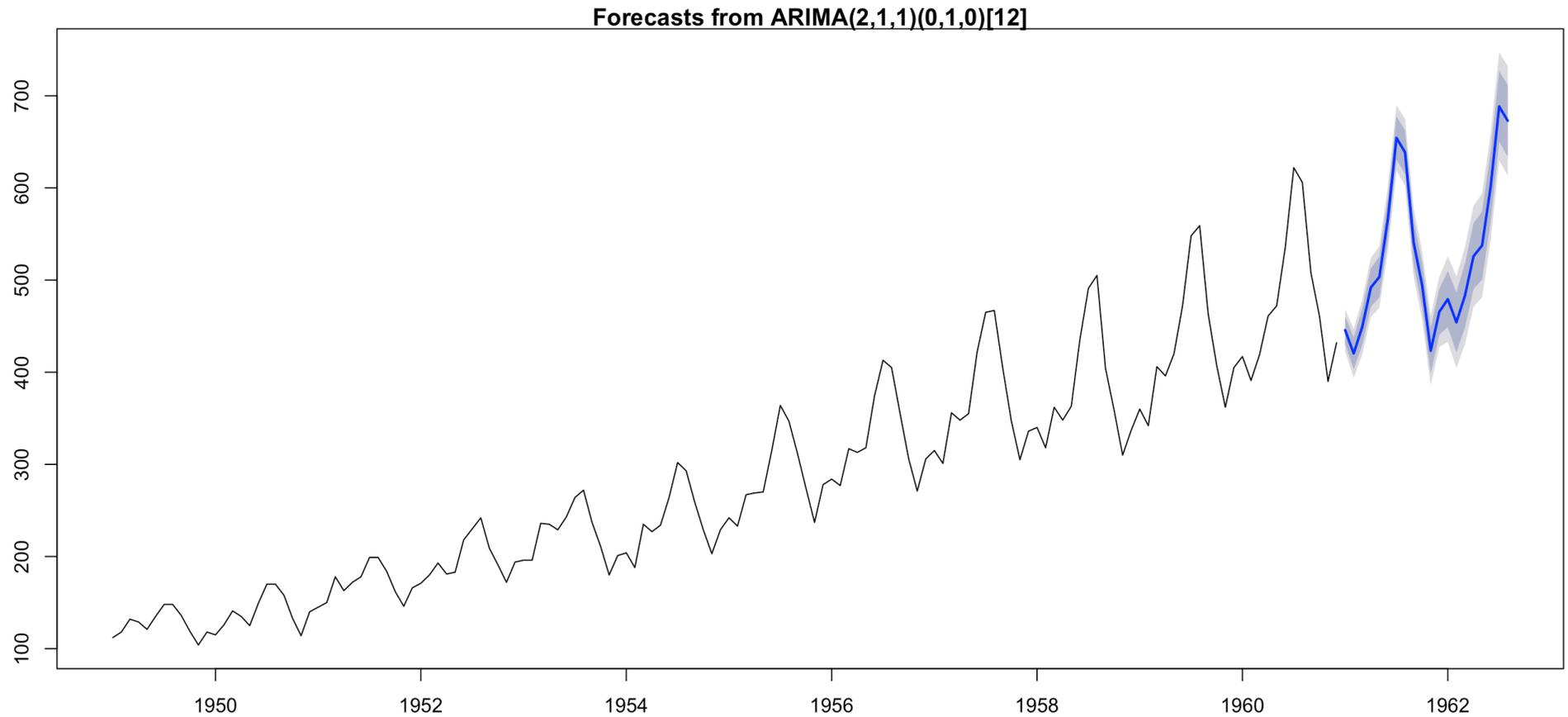
Coefficients:

	ar1	ar2	ma1
	0.5960	0.2143	-0.9819
s.e.	0.0888	0.0880	0.0292

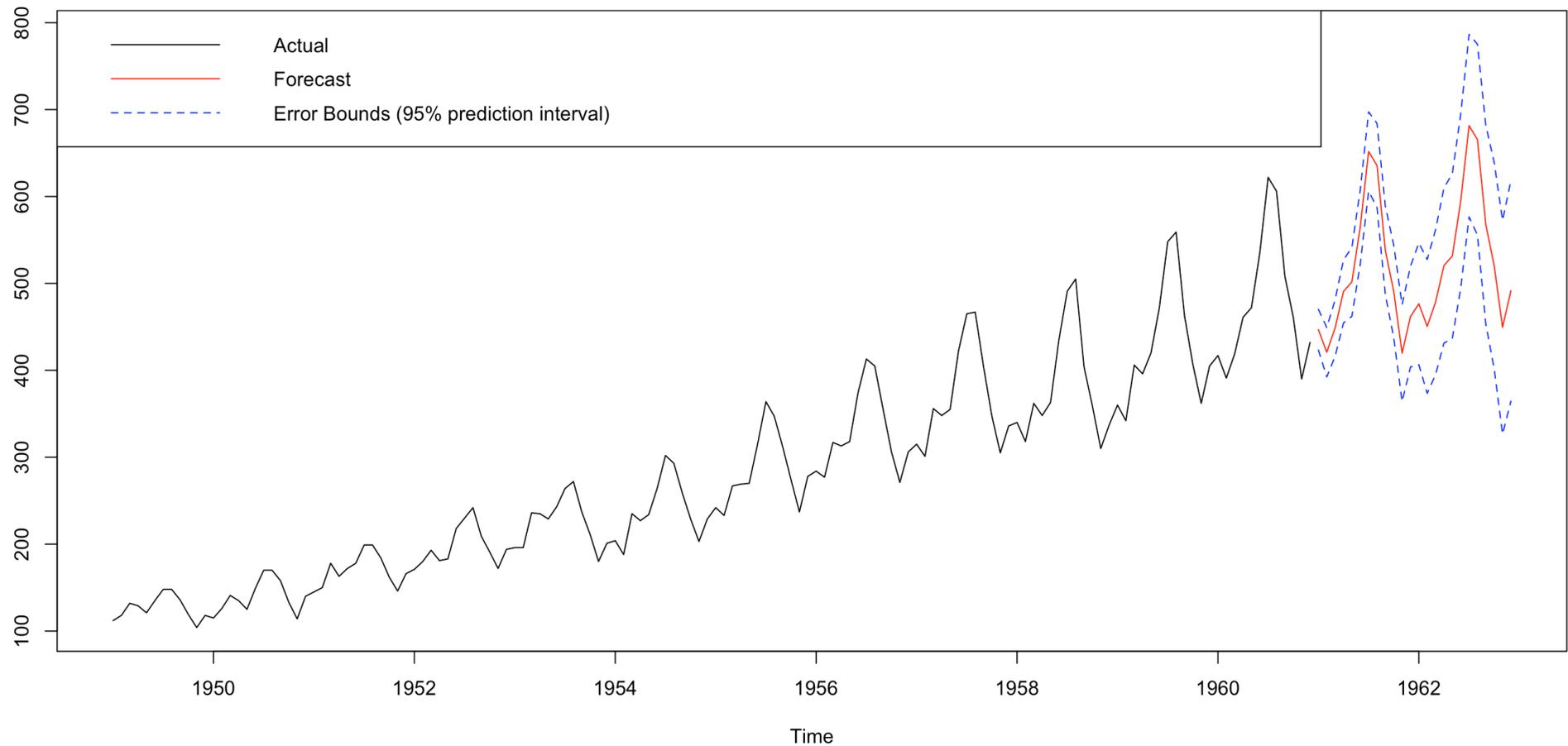
sigma² estimated as 132.3: log likelihood=-504.92

AIC=1017.85 AICc=1018.17 BIC=1029.35

Esempio: passeggeri aerei



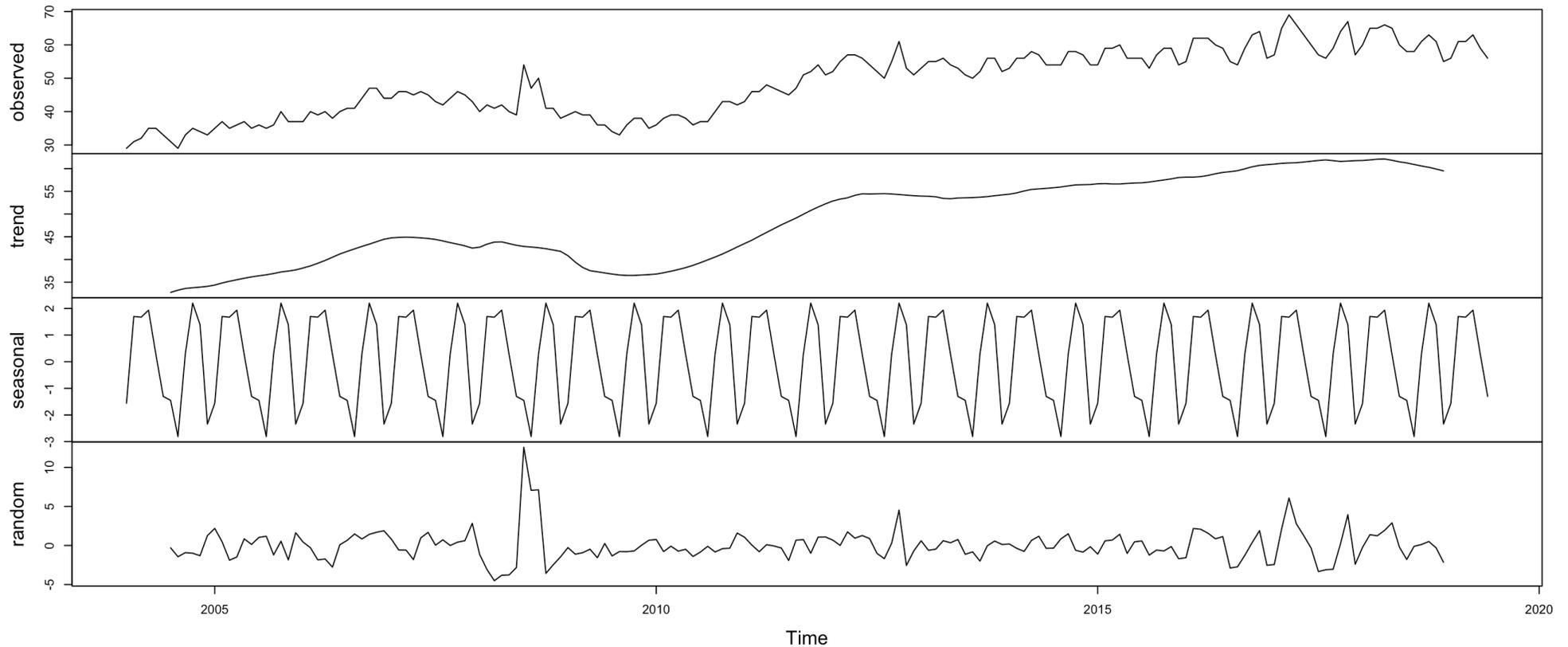
Esempio: passeggeri aerei



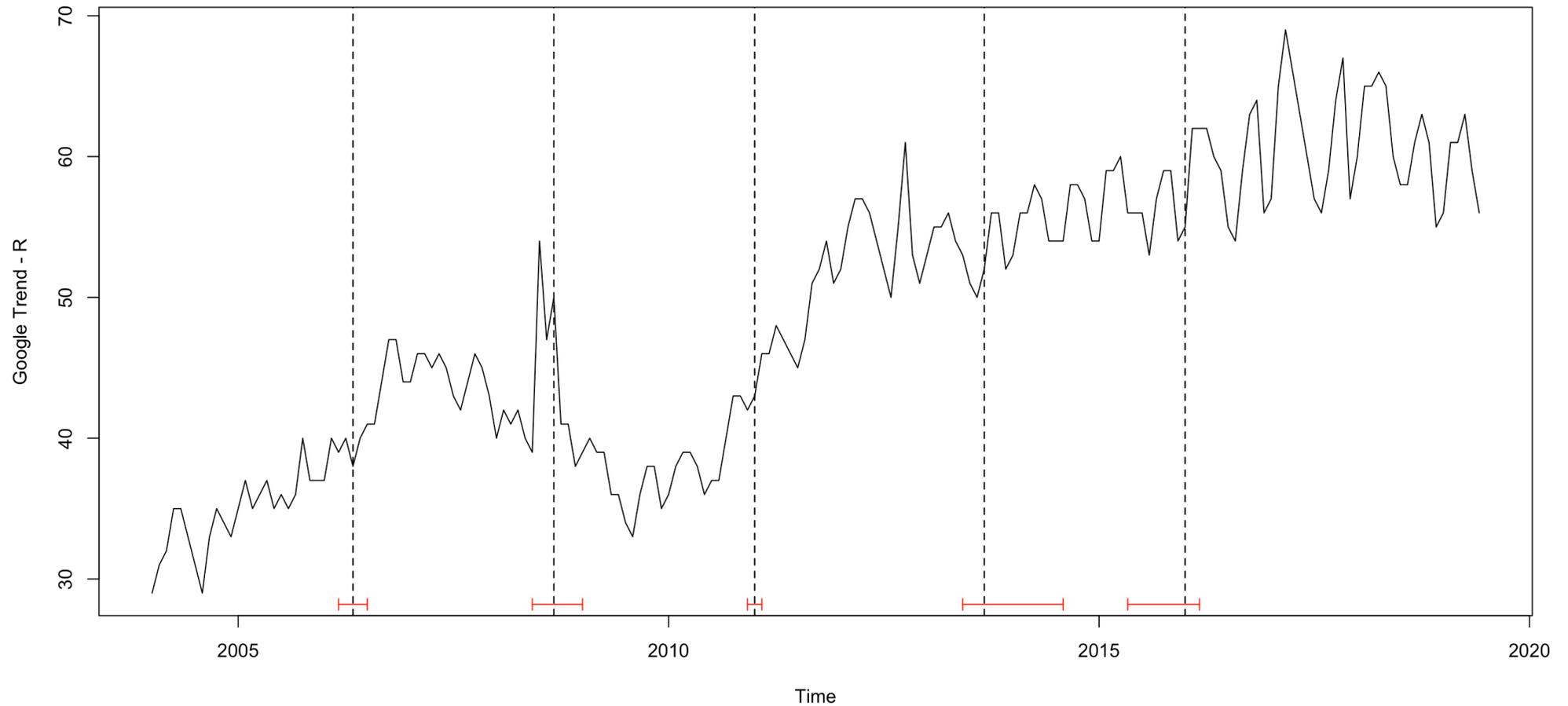
Esempio di serie storica: Google Trend

Mensile - R

Decomposition of additive time series



Esempio di serie storica: Google Trend Mensile - R



Esempio di serie storica: Google Trend Mensile - R

Augmented Dickey-Fuller Test

```
data:  tsgtm[, 1]
Dickey-Fuller = -2.1966, Lag order = 12, p-value = 0.494
alternative hypothesis: stationary
```

```
Series:  tsgtm[, 1]
ARIMA(1,1,1)(2,0,0)[12]
```

Coefficients:

	ar1	ma1	sar1	sar2
	0.4056	-0.6745	0.3440	0.297
s.e.	0.1419	0.1095	0.0685	0.074

```
sigma^2 estimated as 5.819:  log likelihood=-426.21
AIC=862.42  AICc=862.76  BIC=878.52
```

Esempio di serie storica: Google Trend Mensile - R

