Probabilità e Processi Stocastici (455AA)

Lezione 20

Dario Trevisan

5/12/2024

Section 1

Processi a stati continui ARIMA

Richiami

Dato un processo $(X_t)_{t \in \mathcal{T}}$ a stati continui ($\mathbb{R}$), abbiamo introdotto

- La funzione di media $t \mapsto \mathbb{E}[X_t]$,

Richiami

Dato un processo $(X_t)_{t\in\mathcal{T}}$ a stati continui ($\mathbb{R}$), abbiamo introdotto

- La funzione di media $t \mapsto \mathbb{E}[X_t]$,
- La funzione di autocovarianza $(s,t) \mapsto C(s,t) = \mathrm{Cov}(X_s, X_t)$,

$R(s,t)$

Richiami

Dato un processo $(X_t)_{t\in\mathcal{T}}$ a stati continui ($\mathbb{R}$), abbiamo introdotto

- La funzione di media $t \mapsto \mathbb{E}[X_t]$,
- La funzione di autocovarianza $(s,t) \mapsto C(s,t) = \mathrm{Cov}(X_s, X_t)$,
- La funzione di autocorrelazione $(s,t) \mapsto \mathrm{ACF}(s,t) = \rho_{X_s,X_t}$

Richiami

Dato un processo $(X_t)_{t\in\mathcal{T}}$ a stati continui ($\mathbb{R}$), abbiamo introdotto

- La funzione di media $t \mapsto \mathbb{E}[X_t]$,
- La funzione di autocovarianza $(s,t) \mapsto C(s,t) = \mathrm{Cov}(X_s, X_t)$,
- La funzione di autocorrelazione $(s,t) \mapsto \mathrm{ACF}(s,t) = \rho_{X_s, X_t}$
- Il concetto di *stazionarietà in senso lato*: media costante e $C(s,t) = C(0, |t-s|)$

Richiami

Dato un processo $(X_t)_{t\in\mathcal{T}}$ a stati continui ($\mathbb{R}$), abbiamo introdotto

- La funzione di media $t \mapsto \mathbb{E}[X_t]$,
- La funzione di autocovarianza $(s,t) \mapsto C(s,t) = \mathrm{Cov}(X_s, X_t)$,
- La funzione di autocorrelazione $(s,t) \mapsto \mathrm{ACF}(s,t) = \rho_{X_s,X_t}$
- Il concetto di *stazionarietà in senso lato*: media costante e $C(s,t) = C(0,|t-s|)$
- La classe dei *processi gaussiani*

Richiami

Dato un processo $(X_t)_{t \in \mathcal{T}}$ a stati continui ($\mathbb{R}$), abbiamo introdotto

- La funzione di media $t \mapsto \mathbb{E}[X_t]$,
- La funzione di autocovarianza $(s,t) \mapsto C(s,t) = \mathrm{Cov}(X_s, X_t)$,
- La funzione di autocorrelazione $(s,t) \mapsto \mathrm{ACF}(s,t) = \rho_{X_s, X_t}$
- Il concetto di *stazionarietà in senso lato*: media costante e $C(s,t) = C(0, |t-s|)$
- La classe dei *processi gaussiani*
- I modelli *ARIMA*(p,d,q): (W_t) rumore bianco di int. σ^2

$$p(L)(1-L)^d X = q(L) W$$

ARIMA

Section 2

Proprietà dei processi ARIMA

Modelli ARIMA: proprietà

Discutiamo tre proprietà fondamentali dei modelli ARIMA:

1. condizioni sulla stazionarietà,

Modelli ARIMA: proprietà

Discutiamo tre proprietà fondamentali dei modelli ARIMA:

1. condizioni sulla stazionarietà,
2. una equazione ricorsiva per la funzione di autocovarianza (nel caso stazionario)

Modelli ARIMA: proprietà

Discutiamo tre proprietà fondamentali dei modelli ARIMA:

1. condizioni sulla stazionarietà,
2. una equazione ricorsiva per la funzione di autocovarianza (nel caso stazionario)
3. come stimare i parametri sulla base delle osservazioni.

Modelli ARIMA: proprietà

Discutiamo tre proprietà fondamentali dei modelli ARIMA:

1. condizioni sulla stazionarietà,
2. una equazione ricorsiva per la funzione di autocovarianza (nel caso stazionario)
3. come stimare i parametri sulla base delle osservazioni.

- Consideriamo un processo $(X_t)_{t \in \mathcal{T}}$ ARIMA(p, d, q) con parametri $(\alpha_i)_{i=1}^p$ e $(\beta_j)_{j=1}^q$.

$\mathcal{T} = \mathbb{Z}$

Stazionarietà

Nel caso ARIMA$(1,0,0)$, la stazionarietà può avvenire se e solo se $|\alpha_1| < 1$.

- I processi a media mobile ARIMA$(0,0,q)$ possono sempre essere stazionari:

$$X_t = q(L)W_t = W_t + \beta_1 W_{t-1} + \beta_2 W_{t-2} + \ldots + \beta_q W_{t-q}$$

OSS Se Z è stazionario e g $\Rightarrow$ $Z * g$ è staz.

↑ NON aleatorio

$$\mathbb{E}[(Z*g)(t)] = \sum_i \mathbb{E}[Z_{t-i}\, g(i)] = \sum_i m \cdot g(i) \leftarrow \text{non dipende da } t$$

$$\mathbb{E}[(Z*g)(s)(Z*g)(t)] = \mathbb{E}\Big[\sum_{i,j} Z_{s-i}\, g(i)\, Z_{t-j}\, g(j)\Big] = \sum_{i,j} g(i) g(j)\, \overbrace{\mathbb{E}[Z_{s-i} Z_{t-j}]}^{t-s}$$

Stazionarietà

Nel caso ARIMA$(1,0,0)$, la stazionarietà può avvenire se e solo se $|\alpha_1| < 1$.

- I processi a media mobile ARIMA$(0,0,q)$ possono sempre essere stazionari:

$$X_t = q(L) W_t = W_t + \beta_1 W_{t-1} + \beta_2 W_{t-2} + \ldots + \beta_q W_{t-q}$$

- Nel caso generale, l'idea è "risolvere" l'equazione del modello

$$\frac{\cancel{p(L)(1-L)^d} X_t}{\cancel{p(L)(1-L)^d}} = \frac{q(L) W_t}{p(L)(1-L)^d}, \quad \text{da cui} \quad \boxed{X_t = \frac{q(L)}{p(L)(1-L)^d} W_t,}$$

??

Stazionarietà

Nel caso ARIMA$(1,0,0)$, la stazionarietà può avvenire se e solo se $|\alpha_1| < 1$.

- I processi a media mobile ARIMA$(0,0,q)$ possono sempre essere stazionari:

$$X_t = q(L)W_t = W_t + \beta_1 W_{t-1} + \beta_2 W_{t-2} + \ldots + \beta_q W_{t-q}$$

- Nel caso generale, l'idea è "risolvere" l'equazione del modello

$$p(L)(1-L)^d X_t = q(L)W_t, \quad \text{da cui} \quad X_t = \frac{q(L)}{p(L)(1-L)^d} W_t,$$

$X_t = \sum b_k L^k W_t$

- Per dare senso alla scrittura, sviluppiamo in serie

$$\frac{q(z)}{p(z)(1-z)^d} = \sum_{k=0}^{\infty} b_k z^k, \quad \text{così} \quad \boxed{X_t = \sum_{k=0}^{\infty} b_k L^k W_t.}$$

Criterio di stazionarietà

Il problema sta nella convergenza della serie, che dipende dalla crescita dei coefficienti b_k al tendere di $k \to \infty$ e in ultima analisi agli zeri (complessi) del denominatore $p(z)(1-z)^d$.

- Dati (p, d, q) e coefficienti $(\alpha_i)_{i=1}^p$, $(\beta_j)_{j=1}^q$, posto

$$p(z) = 1 - \sum_{i=1}^{p} \alpha_i z^i,$$

allora esiste un modello ARIMA(p, d, q) stazionario con tali coefficienti se e ~~solo se~~ $d = 0$ e tutte le radici complesse di $p(z)$ hanno modulo $|z| > 1$, ossia

se $z \in \mathbb{C}$ è tale che $p(z) = 0$, allora $|z| > 1$.

Esempio ARIMA(1,0,0) $q(z) = 1$ $d = 0$ $p(z) = 1 - \alpha_1 z = 0 \Rightarrow z = \frac{1}{\alpha_1}$

$|z| > 1 \Leftrightarrow |\alpha_1| < 1$

Criterio di stazionarietà

Il problema sta nella convergenza della serie, che dipende dalla crescita dei coefficienti b_k al tendere di $k \to \infty$ e in ultima analisi agli zeri (complessi) del denominatore $p(z)(1-z)^d$.

- Dati (p, d, q) e coefficienti $(\alpha_i)_{i=1}^p$, $(\beta_j)_{j=1}^q$, posto

$$p(z) = 1 - \sum_{i=1}^{p} \alpha_i z^i,$$

  allora esiste un modello ARIMA(p, d, q) stazionario con tali coefficienti se e solo se $d = 0$ e tutte le radici complesse di $p(z)$ hanno modulo $|z| > 1$, ossia

$$\text{se } z \in \mathbb{C} \text{ è tale che } p(z) = 0, \text{ allora } |z| > 1.$$

- Nel caso dell'equazione lineare con smorzamento si trova $|\alpha| < 1$.

Autocovarianza

ARIMA(1,0,0) $C(s,t) = \alpha_1^{t-s} C(s,s)$ se $t \geq s$

$C(s,t) = \alpha\, C(s,t-1)$

I processi ARIMA hanno media nulla e l'equazione permette anche di ottenere una formula ricorsiva (le *equazioni di Yule-Walker*) per la funzione di autocovarianza:

$$C(0,t) = \sum_{i=1}^{p} \alpha_i C(0,t-i), \quad \text{se } t > q.$$

$(d=0)$

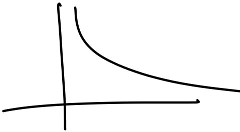
- Nel caso ARIMA$(1,0,0)$ si ha

$$C(0,t) = \alpha_1 C(0,t-1) = \ldots = \alpha_1^t C(0,0).$$

Autocovarianza

I processi ARIMA hanno media nulla e l'equazione permette anche di ottenere una formula ricorsiva (le *equazioni di Yule-Walker*) per la funzione di autocovarianza:

$$C(0,t) = \sum_{i=1}^{p} \alpha_i C(0, t-i), \quad \text{se } t > q.$$



- Nel caso ARIMA$(1,0,0)$ si ha

$$C(0,t) = \alpha_1 C(0, t-1) = \ldots = \alpha_1^t C(0,0).$$

- Nel caso ARIMA$(0,0,q)$ si ha

$$C(0,t) = 0 \quad \text{se } t > q.$$

Dimostrazione

$$X_t = \sum_{i=1}^{p} \alpha_i X_{t-i} + W_t + \sum_{j=1}^{q} \beta_j W_{t-j}$$

Hp: $\boxed{t > q}$

- possiamo supporre che X_t sia indipendente da W_r con $r > t$

$$X_t = f(X_0, X_1, \dots, X_{p-1}, W_0, W_1, \dots, W_t)$$

- $\mathbb{E}[X_t] = \sum_{i=1}^{p} \alpha_i \mathbb{E}[X_{t-i}] = 0$ possiamo supporre $\mathbb{E}[X_0] = \dots = \mathbb{E}[X_{p-1}] = 0$

$$\mathrm{Cov}(X_0, X_t) = \mathrm{Cov}\left(X_0, \sum_{i=1}^{p} \alpha_i X_{t-i} - W_t + \sum_{j=1}^{q} \beta_j W_{t-j}\right)$$

$$= \sum_{i=1}^{p} \alpha_i \mathrm{Cov}(X_0, X_{t-i}) - \cancel{\mathrm{Cov}(X_0, W_t)} + \cancel{\sum_{j=1}^{q} \mathrm{Cov}(X_0, \beta_j W_{t-j})}$$

se $t - q > 0$

~~Dimostrazione~~ Commento

- $X \in ARIMA(0,0,q)$ $\qquad X = W * g \qquad g = (1, \beta_1, \dots, \beta_q)$

$$Cov(X_0, X_t) = \sum_{i,j=0}^{q} \beta_i \cdot \beta_j \, \underbrace{Cov(W_{t-i}, W_{0-j})}_{\sigma^2 \delta_0(t-i+j)}$$

$$= \sigma^2 \sum_{\substack{i,j=0 \\ t-i+j=0}}^{q} \beta_i \beta_j$$

- OSS $X \in ARIMA(p,d,q) \Longrightarrow X \in ARIMA(p+d, 0, q)$

$$\underbrace{p(L)(1-L)^d}\, X = q(L)\, W$$

Stima dei parametri

time series

A partire dall'osservazione di una *serie storica* $(x_t)_{t=0}^n$, come stimare i parametri di un processo ARIMA che la modellizza?

- Abbiamo visto la stima di massima verosimiglianza negli esempi del rumore bianco gaussiano, della passeggiata aleatoria e dell'equazione lineare con smorzamento.

MLE $\Longleftrightarrow$ minimi quadrati

Stima dei parametri

A partire dall'osservazione di una *serie storica* $(x_t)_{t=0}^n$, come stimare i parametri di un processo ARIMA che la modellizza?

- Abbiamo visto la stima di massima verosimiglianza negli esempi del rumore bianco gaussiano, della passeggiata aleatoria e dell'equazione lineare con smorzamento.
- Il metodo si riduce alla minimizzazione dei residui quadratici.

Stima dei parametri

A partire dall'osservazione di una *serie storica* $(x_t)_{t=0}^n$, come stimare i parametri di un processo ARIMA che la modellizza?

- Abbiamo visto la stima di massima verosimiglianza negli esempi del rumore bianco gaussiano, della passeggiata aleatoria e dell'equazione lineare con smorzamento.
- Il metodo si riduce alla minimizzazione dei residui quadratici.
- Si può fare lo stesso per un ARIMA generale, ma la nozione di residuo va chiarita.

Partendo dall'equazione definente

$$p(L)(1-L)^d X_t = q(L) W_t$$

si ricava una equazione ricorsiva per il rumore bianco gaussiano W_t:

$$W_t = -\sum_{j=1}^{q} \beta_j W_{t-j} + p(L)(1-L)^d X_t$$

osservata

- Avendo osservato $X_t = x_t$, definiamo ricorsivamente i residui

$$\boxed{w_t = -\sum_{j=1}^{q} \beta_j w_{t-j} + p(L)(1-L)^d x_t,}$$

cosa succede se t è piccolo

Partendo dall'equazione definente

$$p(L)(1-L)^d X_t = q(L) W_t$$

si ricava una equazione ricorsiva per il rumore bianco gaussiano W_t:

$$W_t = -\sum_{j=1}^{q} \beta_j W_{t-j} + p(L)(1-L)^d X_t$$

- Avendo osservato $X_t = x_t$, definiamo ricorsivamente i residui

$$w_t = -\sum_{j=1}^{q} \beta_j w_{t-j} + p(L)(1-L)^d x_t,$$

α

- La stima di massima verosimiglianza si ottiene minimizzando la somma dei quadrati:

$$(\underline{\alpha_{\mathrm{MLE}}}, \underline{\beta_{\mathrm{MLE}}}) \in \boxed{\arg\min_{\alpha,\beta} \sum_t w_t^2}$$

Stima dei coefficienti in R

- In R vi sono diverse funzioni che permettono la stima (fit) di un ARIMA sui dati osservati, ad esempio `arima()` oppure `Arima()` dalla libreria `forecast`.

Stima dei coefficienti in R

- In R vi sono diverse funzioni che permettono la stima (fit) di un ARIMA sui dati osservati, ad esempio `arima()` oppure `Arima()` dalla libreria `forecast`.
- Oltre alle stime puntuali forniscono stime delle deviazoni standard associate ai parametri.

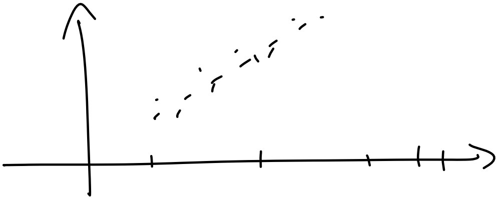
Stima dei coefficienti in R

- In R vi sono diverse funzioni che permettono la stima (fit) di un ARIMA sui dati osservati, ad esempio `arima()` oppure `Arima()` dalla libreria `forecast`.
- Oltre alle stime puntuali forniscono stime delle deviazoni standard associate ai parametri.
- Inoltre con la funzione `forecast()` si ottengono previsioni (basate sulla stima ottenuta) dei valori futuri.

Scelta del modello

Uno dei problemi principali oltre alla stima dei parametri è la scelta del modello (ossia i valori di (p, d, q)).

- come per la regressione, un maggior numero di parametri permette di aderire meglio alle osservazioni, ma non garantisce previsioni accurate.



Scelta del modello

Uno dei problemi principali oltre alla stima dei parametri è la scelta del modello (ossia i valori di (p, d, q)).

- come per la regressione, un maggior numero di parametri permette di aderire meglio alle osservazioni, ma non garantisce previsioni accurate.
- vi sono indicatori particolari (AIC, BIC) che "guidano" opportunamente nella scelta.

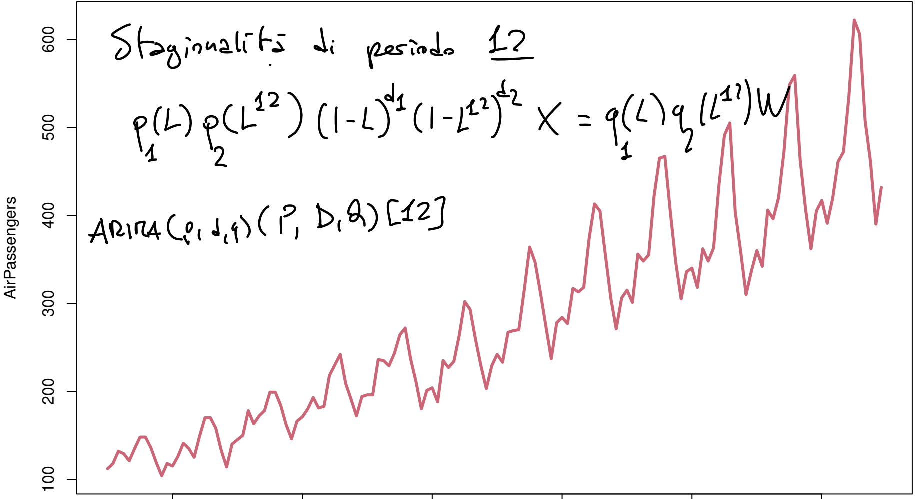
Scelta del modello

Uno dei problemi principali oltre alla stima dei parametri è la scelta del modello (ossia i valori di (p, d, q)).

- come per la regressione, un maggior numero di parametri permette di aderire meglio alle osservazioni, ma non garantisce previsioni accurate.
- vi sono indicatori particolari (AIC, BIC) che "guidano" opportunamente nella scelta.
- In R il comando `auto.arima()` valuta automaticamente il migliore modello che aderisce alle osservazioni sulla base di tali indicatori.

Un esempio

Consideriamo il dataset precaricato in R `AirPassengers`.



Stagionalità di periodo 12

$$\varphi_1(L)\,\varphi_2(L^{12})\,(1-L)^{d_1}(1-L^{12})^{d_2} X = \theta_1(L)\,\theta_2(L^{12})\,W$$

ARIMA(p, d, q)(P, D, Q)[12]

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

## Series: AirPassengers
## 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
##
## sigma^2 = 132.3:  log likelihood = -504.92
## AIC=1017.85   AICc=1018.17   BIC=1029.35
##
## Training set error measures:
##                    ME     RMSE     MAE       MPE     MAPE
## Training set 1.342299 10.84619 7.86754 0.4206976 2.800458 (
```

La funzione `auto.arima()` propone un modello ARIMA con stagionalità di periodo 12 (mesi) e ordine $(2, 1, 1)(0, 1, 0)$ (la seconda tripla si riferisce alla stagionalità).

- In realtà funzione `auto.arima()` non determina automaticamente il periodo 12 e questo va indicato prima di applicarla ai dati, nel momento in cui si definisce un oggetto di tipo *serie storica* (in inglese *time series*) in R.

La funzione `auto.arima()` propone un modello ARIMA con stagionalità di periodo 12 (mesi) e ordine $(2,1,1)(0,1,0)$ (la seconda tripla si riferisce alla stagionalità).

- In realtà funzione `auto.arima()` non determina automaticamente il periodo 12 e questo va indicato prima di applicarla ai dati, nel momento in cui si definisce un oggetto di tipo *serie storica* (in inglese *time series*) in R.
- Partendo da un vettore di dati osservati, il comando è `ts()`, che contiene l'opzione *frequency* (se non specificata è posta uguale ad 1).

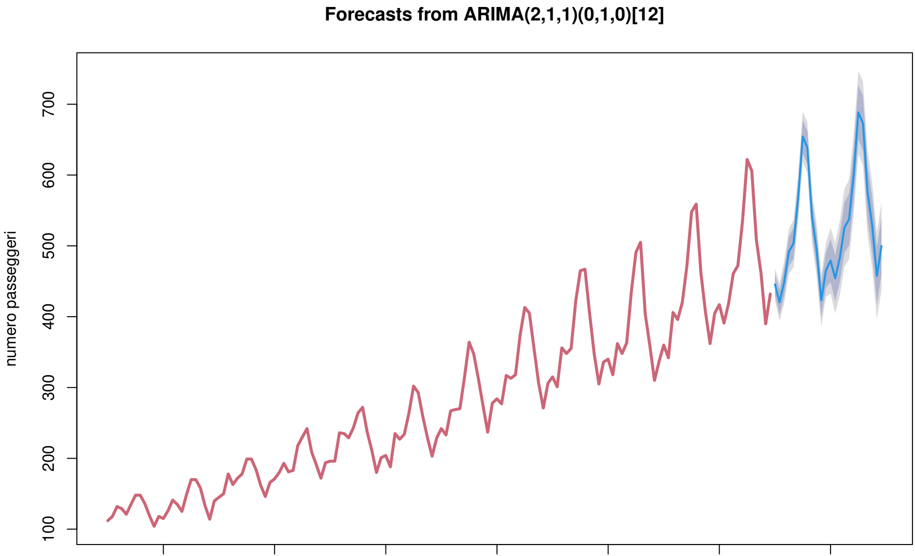
La funzione `auto.arima()` propone un modello ARIMA con stagionalità di periodo 12 (mesi) e ordine $(2, 1, 1)(0, 1, 0)$ (la seconda tripla si riferisce alla stagionalità).

- In realtà funzione `auto.arima()` non determina automaticamente il periodo 12 e questo va indicato prima di applicarla ai dati, nel momento in cui si definisce un oggetto di tipo *serie storica* (in inglese *time series*) in R.
- Partendo da un vettore di dati osservati, il comando è `ts()`, che contiene l'opzione *frequency* (se non specificata è posta uguale ad 1).
- In generale si può ricorrere alla funzione di autocorrelazione empirica o ad analisi spettrale per determinare eventuali stagionalità e il loro periodo.

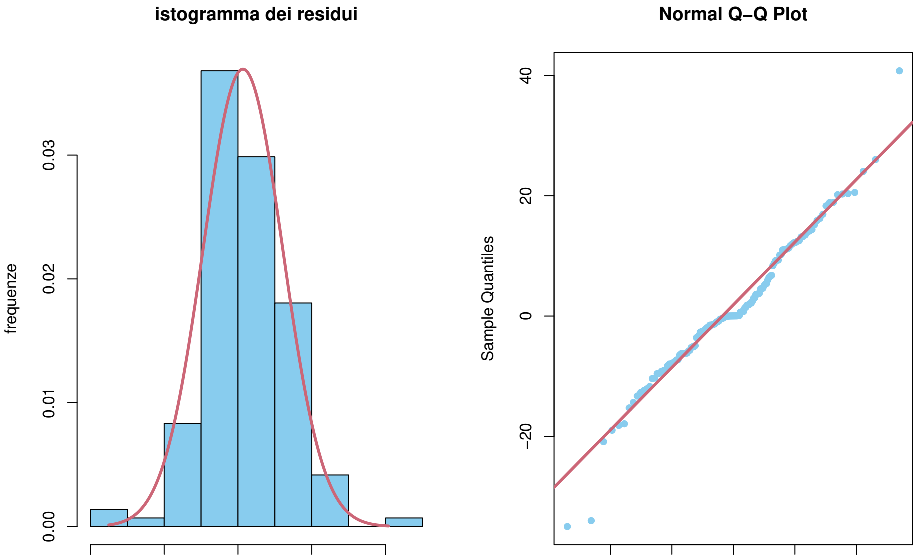
La funzione `auto.arima()` propone un modello ARIMA con stagionalità di periodo 12 (mesi) e ordine $(2, 1, 1)(0, 1, 0)$ (la seconda tripla si riferisce alla stagionalità).

- In realtà funzione `auto.arima()` non determina automaticamente il periodo 12 e questo va indicato prima di applicarla ai dati, nel momento in cui si definisce un oggetto di tipo *serie storica* (in inglese *time series*) in R.
- Partendo da un vettore di dati osservati, il comando è `ts()`, che contiene l'opzione *frequency* (se non specificata è posta uguale ad 1).
- In generale si può ricorrere alla funzione di autocorrelazione empirica o ad analisi spettrale per determinare eventuali stagionalità e il loro periodo.
- Un comando automatico è invece `findfrequency()`.

Previsione


Forecasts from ARIMA(2,1,1)(0,1,0)[12]
numero passeggeri
100
200
300
400
500
600
700

Residui


istogramma dei residui
frequenze
0.00
0.01
0.02
0.03
Normal Q−Q Plot
Sample Quantiles
−20
0
20
40

Section 3

Stima della funzione di autocovarianza

Stima della funzione di autocovarianza

Affrontiamo il problema generale di stimare la funzione di autocovarianza di un processo X a partire dall'osservazione dei valori $(X_t)_{t=0}^{n-1} = (x_t)_{t=0}^{n-1}$ (una *serie storica*).

- È una una generalizzazione del problema di stimare valor medio e varianza di una famiglia di variabili aleatorie indipendenti, tutte con la stessa legge.

Stima della funzione di autocovarianza

Affrontiamo il problema generale di stimare la funzione di autocovarianza di un processo X a partire dall'osservazione dei valori $(X_t)_{t=0}^{n-1} = (x_t)_{t=0}^{n-1}$ (una *serie storica*).

- È una una generalizzazione del problema di stimare valor medio e varianza di una famiglia di variabili aleatorie indipendenti, tutte con la stessa legge.
- Nel caso gaussiano avevamo ottenuto la media e la covarianza campionarie.

Verosimiglianza

Invece dell'indipendenza, supporremo la *stazionarietà del processo*, ossia la matrice di covarianza delle variabili $(X_t)_{t=0}^{n-1}$ è costante sulle diagonali:

$$(C(s,t))_{s,t=0}^{n-1} = (C(|t-s|))_{s,t=0}^{n-1}.$$

- Supponiamo pure che il processo X sia *gaussiano e centrato* (ossia la funzione di media è nota e costantemente nulla)

Verosimiglianza

Invece dell'indipendenza, supporremo la *stazionarietà del processo*, ossia la matrice di covarianza delle variabili $(X_t)_{t=0}^{n-1}$ è costante sulle diagonali:

$$(C(s,t))_{s,t=0}^{n-1} = (C(|t-s|))_{s,t=0}^{n-1}.$$

- Supponiamo pure che il processo X sia *gaussiano e centrato* (ossia la funzione di media è nota e costantemente nulla)
- La verosimiglianza per la matrice di covarianza è

$$\begin{aligned} L(C;x) &= p((X_t)_{t=0}^{n-1} = (x_t)_{t=0}^{n-1} | C) \\ &\propto \exp\left(-\frac{1}{2}x^T C^{-1} x\right) \frac{1}{\sqrt{\det C}}, \end{aligned}$$

dove abbiamo posto, per alleggerire la notazione, $x = (x_t)_{t=0}^{n-1}$.

Stima di massima verosimiglianza

Possiamo determinare la stima di massima verosimiglianza per C con i soliti passaggi: si tratta di minimizzare la funzione

$$C \mapsto x^T C^{-1} x + \log \det C.$$

- È difficile calcolare analiticamente C_{MLE}, ma si può ricorrere a metodi numerici.

Una ipotesi semplificativa

Per ottenere delle espressioni elementari per C_{MLE} introduciamo una ulteriore ipotesi nella struttura della matrice di covarianza.

- Oltre al fatto che $C(s,t)$ sia simmetrica, semidefinita positiva e costante sulle diagonali (per la stazionarietà), chiediamo che sia *circolante*:

$$C(k) = C(n-k) \quad \text{per ogni } k = 1, 2, \ldots, n-1.$$

Una ipotesi semplificativa

Per ottenere delle espressioni elementari per C_{MLE} introduciamo una ulteriore ipotesi nella struttura della matrice di covarianza.

- Oltre al fatto che $C(s,t)$ sia simmetrica, semidefinita positiva e costante sulle diagonali (per la stazionarietà), chiediamo che sia *circolante*:

$$C(k) = C(n-k) \quad \text{per ogni } k = 1, 2, \ldots, n-1.$$

- È solo una ipotesi di comodo, una volta ottenute delle stime esplicite cerchiamo di rimuoverla.

Passaggio alla base delle frequenze

Tramite trasformata di Fourier a tempi finiti, C diventa diagonale. Richiamiamo la notazione:

- Sia $F \in \mathbb{C}^{n \times n}$,

$$F_{\xi t} = e^{-2\pi i \xi t/n}, \quad \text{per } \xi, t = 0, 1, \ldots, (n-1).$$

Passaggio alla base delle frequenze

Tramite trasformata di Fourier a tempi finiti, C diventa diagonale. Richiamiamo la notazione:

- Sia $F \in \mathbb{C}^{n \times n}$,

$$F_{\xi t} = e^{-2\pi i \xi t/n}, \quad \text{per } \xi, t = 0, 1, \ldots, (n-1).$$

- La trasformata di Fourier a tempi finiti di $(x_t)_{t=0}^{n-1}$ è

$$\hat{x}(\xi) = \sum_{t=0}^{n-1} x_t e^{-2\pi i \xi t/n} = \sum_{t=0}^{n-1} F_{\xi t} x_t,$$

ossia $\hat{x} = Fx$.

Se il vettore x è l'osservazione del processo X, la matrice di covarianza del vettore aleatorio $\hat{X} = FX$ si calcola tramite la formula per le trasformazioni affini:

$$\Sigma_{FX} = F \Sigma_X \bar{F}^T = FC\bar{F}^T,$$

- Questo cambio di coordinate: dalla base dei "tempi" a quella delle "frequenze" diagonalizza la matrice delle covarianze.

Se il vettore x è l'osservazione del processo X, la matrice di covarianza del vettore aleatorio $\hat{X} = FX$ si calcola tramite la formula per le trasformazioni affini:

$$\Sigma_{FX} = F\Sigma_X \bar{F}^T = FC\bar{F}^T,$$

- Questo cambio di coordinate: dalla base dei "tempi" a quella delle "frequenze" diagonalizza la matrice delle covarianze.
- Introduciamo la notazione

$$\hat{C}(\xi) = \sum_{k=0}^{n-1} e^{-2\pi i \xi k/n} C(k).$$

Stima di massima verosimiglianza

I numeri $n\hat{C}(\xi)$ sono quindi multipli degli autovalori della matrice di covarianza C e quindi sono tutti positivi.

- In queste nuove coordinate, le componenti del vettore FX sono non correlate e quindi, essendo gaussiane, indipendenti.

Stima di massima verosimiglianza

I numeri $n\hat{C}(\xi)$ sono quindi multipli degli autovalori della matrice di covarianza C e quindi sono tutti positivi.

- In queste nuove coordinate, le componenti del vettore FX sono non correlate e quindi, essendo gaussiane, indipendenti.
- La verosimiglianza assume quindi una espressione molto più trattabile:

$$L(C;\hat{x}) = p(FX = \hat{x}|C) \propto \exp\left(-\frac{1}{2n}\sum_{\xi=0}^{n-1}\frac{|\hat{x}(\xi)|^2}{\hat{C}(\xi)}\right)\frac{1}{\sqrt{\prod_{\xi=0}^{n-1}\hat{C}(\xi)}}$$

Stima di massima verosimiglianza

I numeri $n\hat{C}(\xi)$ sono quindi multipli degli autovalori della matrice di covarianza C e quindi sono tutti positivi.

- In queste nuove coordinate, le componenti del vettore FX sono non correlate e quindi, essendo gaussiane, indipendenti.
- La verosimiglianza assume quindi una espressione molto più trattabile:

$$L(C;\hat{x}) = p(FX = \hat{x}|C) \propto \exp\left(-\frac{1}{2n}\sum_{\xi=0}^{n-1}\frac{|\hat{x}(\xi)|^2}{\hat{C}(\xi)}\right)\frac{1}{\sqrt{\prod_{\xi=0}^{n-1}\hat{C}(\xi)}}$$

- La stima di massima verosimiglianza si ottiene minimizzando la funzione

$$C \mapsto \sum_{\xi=1}^{n-1}\left[\frac{1}{n}\frac{|\hat{x}(\xi)|^2}{\hat{C}(\xi)} + \log\hat{C}(\xi)\right]$$

Discutiamo l'espressione

$$\frac{1}{n^2}\sum_{s,t=0}^{n-1} x_t x_s \sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n}$$

- Se $t = s + k$, allora il termine

$$\sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n} = \sum_{\xi=0}^{n-1} 1 = n,$$

Discutiamo l'espressione

$$\frac{1}{n^2}\sum_{s,t=0}^{n-1} x_t x_s \sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n}$$

- Se $t = s + k$, allora il termine
$$\sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n} = \sum_{\xi=0}^{n-1} 1 = n,$$
- Se $t = s + k - n$ ugualmente si avrebbe
$$\sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n} = \sum_{\xi=0}^{n-1} e^{2\pi i \xi} = \sum_{\xi=0}^{n-1} 1 = n.$$

Discutiamo l'espressione

$$\frac{1}{n^2}\sum_{s,t=0}^{n-1} x_t x_s \sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n}$$

- Se $t = s + k$, allora il termine
$$\sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n} = \sum_{\xi=0}^{n-1} 1 = n,$$
- Se $t = s + k - n$ ugualmente si avrebbe
$$\sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n} = \sum_{\xi=0}^{n-1} e^{2\pi i\xi} = \sum_{\xi=0}^{n-1} 1 = n.$$
- In tutti gli altri casi possibili, si trova invece
$$\sum_{\xi=0}^{n-1} e^{-2\pi i(t-s-k)\xi/n} = 0,$$

Concludiamo che

$$C_{\mathrm{MLE}}(k) = \frac{1}{n} \sum_{s=0}^{n-1-k} x_s x_{s+k} + \frac{1}{n} \sum_{s=n-k+1}^{n-1} x_s x_{s+k-n}.$$

- È la somma due contributi:

Concludiamo che

$$C_{\mathrm{MLE}}(k) = \frac{1}{n} \sum_{s=0}^{n-1-k} x_s x_{s+k} + \frac{1}{n} \sum_{s=n-k+1}^{n-1} x_s x_{s+k-n}.$$

- È la somma due contributi:
  1. la covarianza campionaria tra X_s e il processo "traslato" avanti nel tempo di k istanti, X_{s+k}

Concludiamo che

$$C_{\mathrm{MLE}}(k) = \frac{1}{n} \sum_{s=0}^{n-1-k} x_s x_{s+k} + \frac{1}{n} \sum_{s=n-k+1}^{n-1} x_s x_{s+k-n}.$$

- È la somma due contributi:
  1. la covarianza campionaria tra X_s e il processo “traslato” avanti nel tempo di k istanti, X_{s+k}
  2. la covarianza campionaria tra X_s e il traslato indietro di $n-k$ istanti, X_{s+k-n}.

Concludiamo che

$$C_{\mathrm{MLE}}(k) = \frac{1}{n}\sum_{s=0}^{n-1-k} x_s x_{s+k} + \frac{1}{n}\sum_{s=n-k+1}^{n-1} x_s x_{s+k-n}.$$

- È la somma due contributi:
  1. la covarianza campionaria tra X_s e il processo "traslato" avanti nel tempo di k istanti, X_{s+k}
  2. la covarianza campionaria tra X_s e il traslato indietro di $n-k$ istanti, X_{s+k-n}.
- Se $k << n$, il primo è rilevante.

Autocovarianza campionaria

I calcoli svolti suggeriscono di proporre come stima generale per C solo il primo contributo: la covarianza campionaria tra X_s e il traslato X_{s+k}.

- Definiamo la funzione di **autocovarianza campionaria** (o empirica) come

$$c(k) = \frac{1}{n-k} \sum_{s=0}^{n-1-k} (x_s - \bar{x}_0)(x_{s+k} - \bar{x}_k),$$

  dove le medie campionarie $\bar{x}_0$, $\bar{x}_k$ sono rispettivamente sui primi $n-k$ e sugli ultimi $n-k$ valori.

Autocovarianza campionaria

I calcoli svolti suggeriscono di proporre come stima generale per C solo il primo contributo: la covarianza campionaria tra X_s e il traslato X_{s+k}.

- Definiamo la funzione di **autocovarianza campionaria** (o empirica) come

$$c(k) = \frac{1}{n-k} \sum_{s=0}^{n-1-k} (x_s - \bar{x}_0)(x_{s+k} - \bar{x}_k),$$

  dove le medie campionarie $\bar{x}_0$, $\bar{x}_k$ sono rispettivamente sui primi $n-k$ e sugli ultimi $n-k$ valori.

- $c(k)$ è la covarianza tra due variabili aleatorie così definite: si sceglie $S \in \{0, 1, \ldots, n-1-k\}$ casuale uniforme e si considerano i valori x_S (prima variabile) e x_{S+k} (seconda variabile).

Funzione di autocorrelazione campionaria

Possiamo definire anche le varianze campionarie

$$\sigma_0^2 = \frac{1}{n-k} \sum_{s=0}^{n-1-k} (x_s - \bar{x}_0)^2$$

e

$$\sigma_k^2 = \frac{1}{n-k} \sum_{s=0}^{n-1-k} (x_{s+k} - \bar{x}_k)^2$$

e quindi la funzione di **autocorrelazione campionaria**, data da

$$\mathrm{acf}(k) = \frac{c(k)}{\sigma_0 \sigma_k},$$

che assume sempre valori tra $[-1, 1]$

ACF in R

In R è l'autocorrelazione campionaria si calcola tramite la funzione `acf()`.

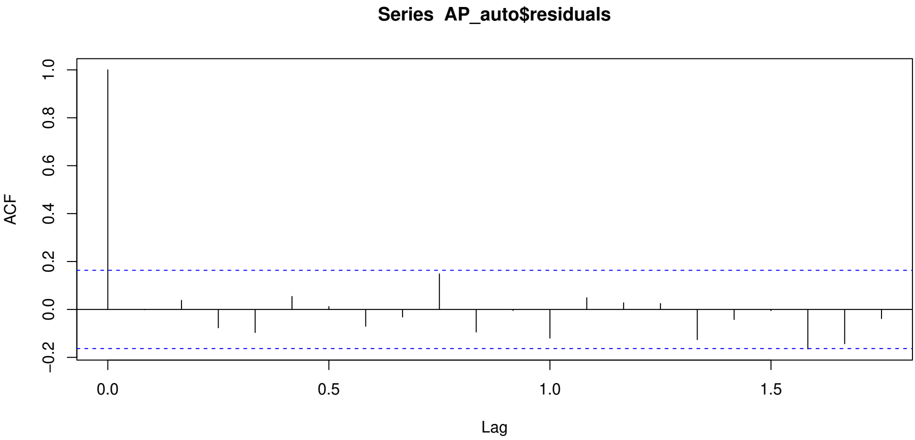


Figure 4: autocorrelazione empirica dei residui

Attenzione: la presenza di un "trend" può fare apparire la ACF uniformemente grande.

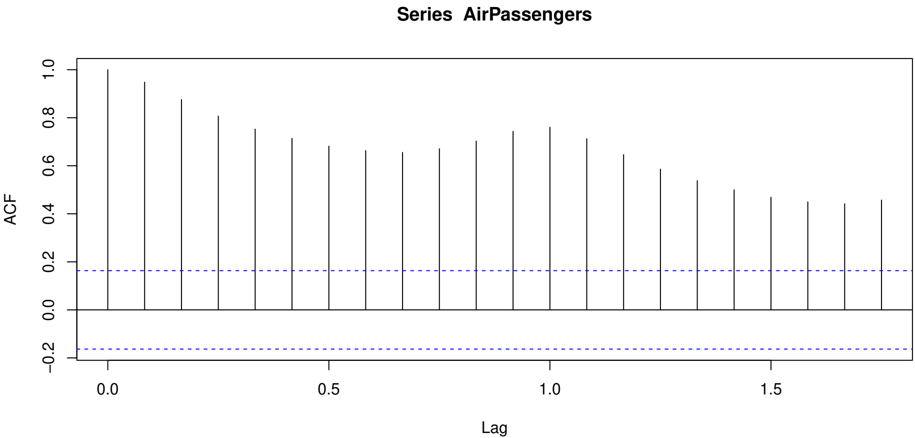


Figure 5: la funzione di autocorrelazione è uniformemente grande per via del trend

Per rimuovere questo effetto basta passare ad una derivata discreta tramite la funzione `diff()`.

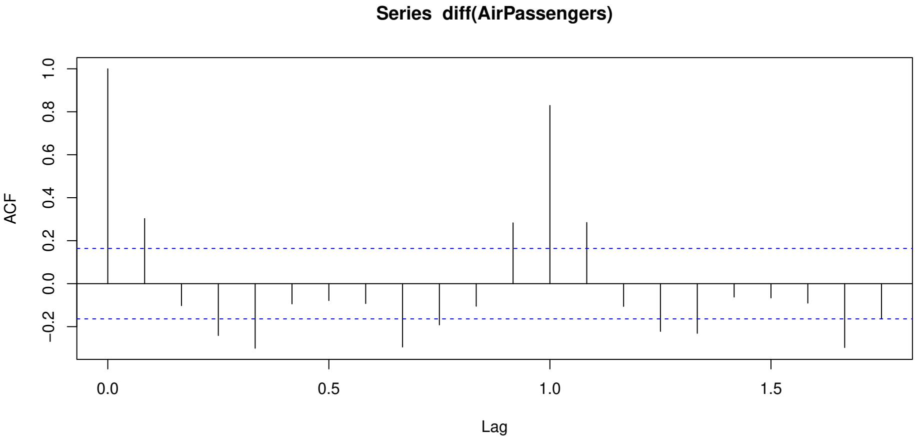


Figure 6: la derivata discreta rimuove il trend e la funzione di autocorrelazione evidenzia la stagionalità a 12 mesi

Densità spettrale di potenza

Torniamo alla base delle frequenze: cosa accade ai calcoli se non vale l'ipotesi semplificativa $C(k) = C(n-k)$?

- Se n è molto grande la stima

$$\hat{C}(\xi) \approx \frac{|\hat{x}(\xi)|^2}{n} = \frac{1}{n}\left|\sum_{t=0}^{n-1} x_t e^{-2\pi i t\xi/n}\right|^2$$

  è comunque una buona approssimazione.

Densità spettrale di potenza

Torniamo alla base delle frequenze: cosa accade ai calcoli se non vale l'ipotesi semplificativa $C(k) = C(n-k)$?

- Se n è molto grande la stima

$$\hat{C}(\xi) \approx \frac{|\hat{x}(\xi)|^2}{n} = \frac{1}{n}\left|\sum_{t=0}^{n-1} x_t e^{-2\pi i t\xi/n}\right|^2$$

  è comunque una buona approssimazione.

- Il membro a destra è l'energia associata alla frequenza ξ diviso il tempo n, quindi si intepreta come *potenza*.

Il Teorema di Wiener-Khinchin

Sia $(X_t)_{t=0}^{\infty}$ un processo a valori reali, stazionario in senso lato, con media nulla $\mathbb{E}[X_t] = 0$, e tale che

$$\sum_{k=0}^{\infty} |C(0,k)| < \infty.$$

- Per ogni $\xi \in [0,1]$, si definisce *densità spettrale di potenza* associata alla frequenza ξ la quantità
$$\hat{C}(\xi) = \sum_{k \in \mathbb{Z}} C(|k|) e^{-2\pi i k \xi}.$$

Il Teorema di Wiener-Khinchin

Sia $(X_t)_{t=0}^{\infty}$ un processo a valori reali, stazionario in senso lato, con media nulla $\mathbb{E}[X_t] = 0$, e tale che

$$\sum_{k=0}^{\infty} |C(0,k)| < \infty.$$

- Per ogni $\xi \in [0,1]$, si definisce *densità spettrale di potenza* associata alla frequenza ξ la quantità

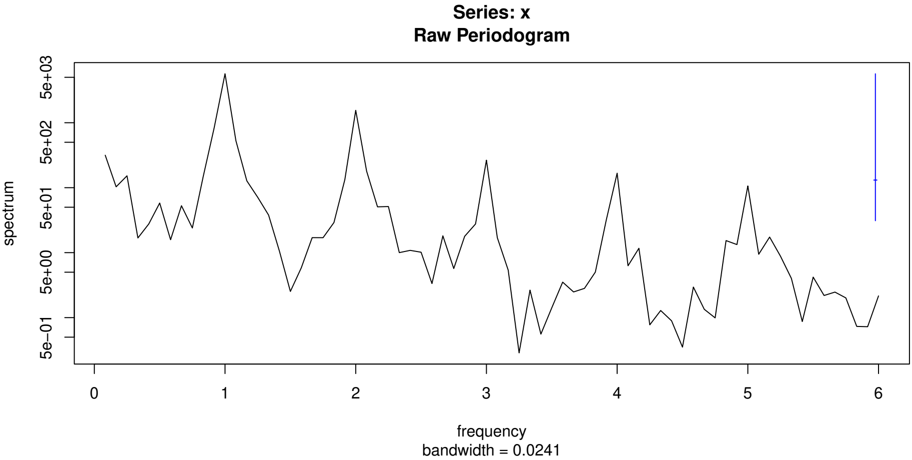
$$\hat{C}(\xi) = \sum_{k \in \mathbb{Z}} C(|k|) e^{-2\pi i k \xi}.$$

- Vale il limite

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left[\left| \sum_{t=0}^{n-1} X_t e^{-2\pi i t \xi} \right|^2 \right] = \hat{C}(\xi).$$

Spettrogramma in R

In R si può utilizzare direttamente il comando `spectrum()`, che fonisce una stima della densità spettrale di potenza da una serie storica osservata.



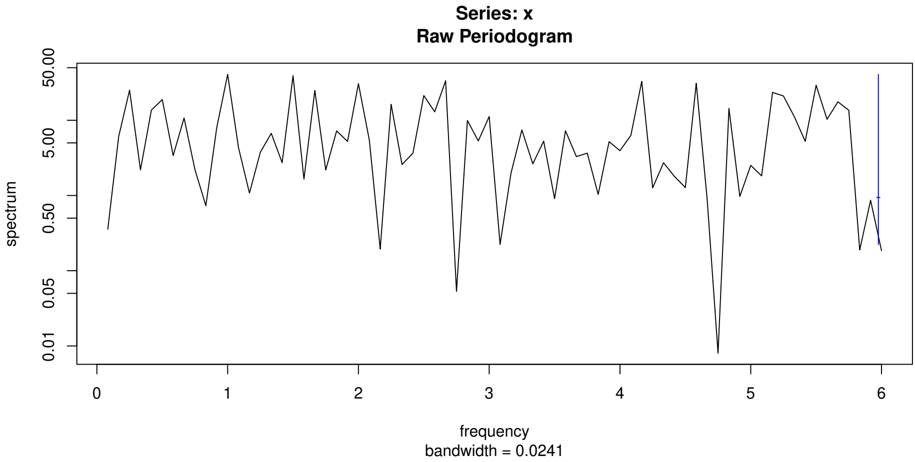


Figure 8: Stima della densità spettrale di potenza dei residui della serie AirPassengers dopo un fit con un modello ARIMA