

Les sèries temporals

Frederic Utzet
Departament d'Estadística.
Facultat de Matemàtiques.
Universitat de Barcelona

Prefaci

Aquestes notes corresponen (amb l'afegit de la darrera secció) a un curs sobre sèries temporals impartit a l'Institut d'Estudis Catalans durant el mes de maig de 1992. Els participants eren matemàtics amb coneixements d'estadística, molt interessats per aquest tema, i varen ser unes sessions molt agradables.

L'objectiu va ser donar una visió general sobre la teoria i les aplicacions de les sèries temporals (metodologia Box-Jenkins i anàlisi espectral), sense entrar gaire en els detalls tècnics —no hi havia prou temps. Aquest plantejament sempre és difícil; no tant per la falta de rigor en l'exposició, sinó perquè en estudiar temes de matemàtiques, *entrar en els detalls*, treballar a fons les demostracions, fer exercicis i problemes, dóna temps a la intel·ligència per assimilar els conceptes explicats i relacionar-los amb altres conceptes. Les idees acostumen a ser ràpides de dir, però més lentes d'entendre. En saltar-se els detalls sempre hi ha el perill d'anar massa de pressa. Vaig intentar, en la mesura del possible, explicar les intuïcions que hi havia al darrera de la teoria i treballar amb exemples il·lustratius de les idees que exposava.

Vull agrair a la Societat Catalana de Matemàtiques, i especialment als professors Josep Pla i Carrera i Antoni Gomà i Nassarre, per haver-me convidat a donar aquestes lliçons, i publicar aquestes notes al Butlletí de la Societat.

Frederic Utzet
Barcelona, octubre de 1992

Índex

1. Introducció	25
2. Anàlisi de Fourier de funcions reals.....	30
3. Processos estocàstics.....	36
4. Representació espectral de processos estacionaris	60
5. Estimació de la funció d'autocorrelació d'un procés estacionari	66
6. Estimació de paràmetres d'un procés ARMA	74
7. Sèries temporals no estacionàries: models ARIMA	78
8. Previsió	87
9. Anàlisi espectral	93
Apèndix: conjunts de dades	105
Bibliografia	109

Les sèries temporals

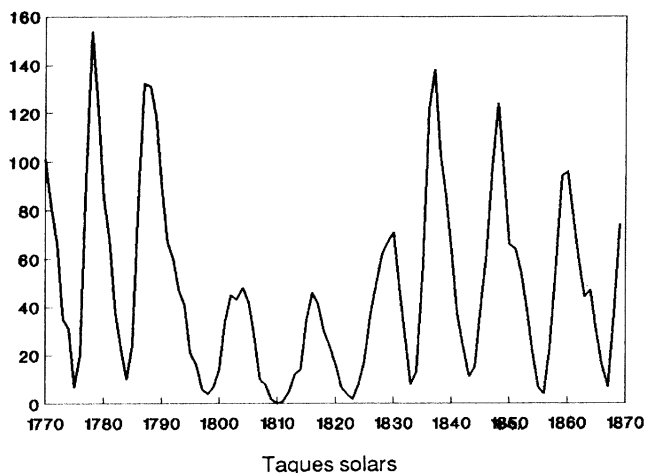
1. Introducció

A grans trets, una *sèrie temporal*, és una successió d'observacions al llarg del temps: x_1, \dots, x_n . Per exemple, les dades següents corresponen a un índex que mesura l'activitat del sol al llarg de l'any (*Wölfer sunspot numbers*) de l'any 1770 a l'any 1896, i a la qual ens referim com a la sèrie de les taques solars (Box i Jenkins, 1976). Vegeu també Morris (1977).

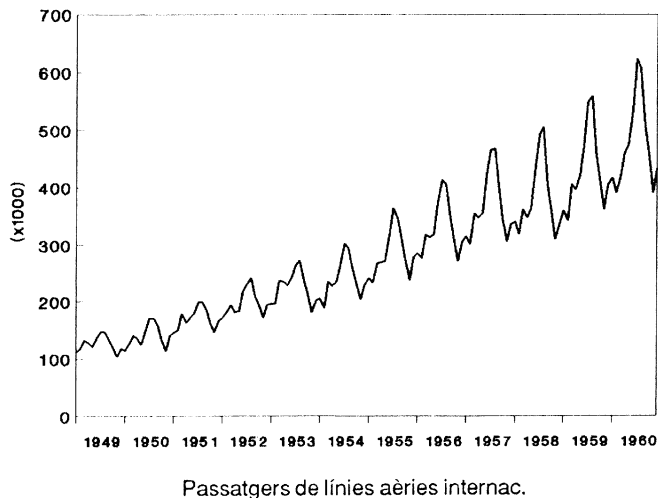
Taques solars (1770-1869)

Any	Taques	Any	Taques	Any	Taques	Any	Taques	Any	Taques
1770	101	1790	90	1810	0	1830	71	1850	66
1771	82	1791	67	1811	1	1831	48	1851	64
1772	66	1792	60	1812	5	1832	28	1852	54
1773	35	1793	47	1813	12	1833	8	1853	39
1774	31	1794	41	1841	14	1834	13	1854	21
1775	7	1795	21	1815	35	1835	57	1855	7
1776	20	1796	16	1816	46	1836	122	1856	4
1777	92	1797	6	1817	41	1837	138	1857	23
1778	154	1798	4	1818	30	1838	103	1858	55
1779	125	1799	7	1819	24	1839	86	1859	94
1780	85	1800	14	1820	16	1840	63	1860	96
1781	68	1801	34	1821	7	1841	37	1861	77
1782	38	1802	45	1822	4	1842	24	1862	59
1783	23	1803	43	1823	2	1843	11	1863	44
1784	10	1804	48	1824	8	1844	15	1864	47
1785	24	1805	42	1825	17	1845	40	1865	30
1786	83	1806	28	1826	36	1846	62	1866	16
1787	132	1807	10	1827	50	1847	98	1867	7
1788	131	1808	8	1828	62	1848	124	1868	37
1789	118	1809	2	1829	67	1849	96	1869	74

Una representació gràfica d'aquestes dades és la següent:



Un altre exemple interessant és el nombre de passatgers de línies aèries internacionals del gener de 1949 al desembre de 1960 (Box i Jenkins, 1976). (Vegeu l'apèndix per a les dades.)



Naturalment poden considerar-se moltes altres situacions: per exemple, les observacions poden fer-se en temps continu (un registre continu de la temperatura d'un lloc); o les observacions poden ser vectorials (per exemple, una sèrie temporal bidimensional seria la formada per l'índex de preus i el nombre de conflictes laborals; es tractarà no sols d'estudiar cada sèrie separatament sinó també de buscar les relacions entre ambdues).

Nosaltres, però, ens limitarem al cas més habitual (i senzill): considerarem sèries en temps discret, observacions escalars, i les observacions estaran igualment separades (cada hora, cada mes, etc.).

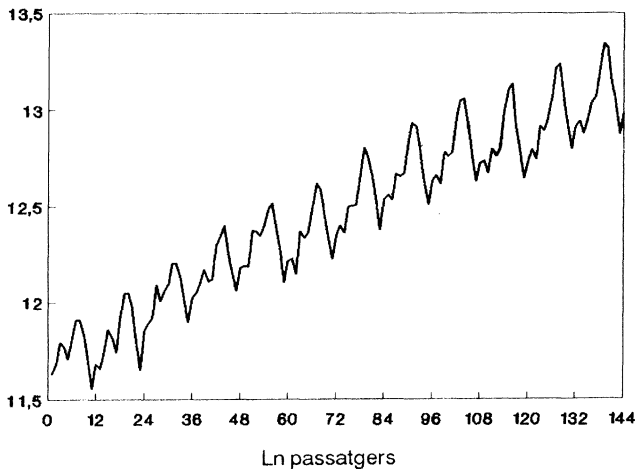
El nostre objectiu és construir models matemàtics que es puguin aplicar al màxim nombre de situacions, que ens permetin descriure en termes d'uns pocs paràmetres els fenòmens (i a ser possible entendre'ls), comparar diferents casos, i fer prediccions (previsions) sobre l'evolució futura.

En general, la millor manera de mirar una sèrie temporal és des del punt de vista estadístic. En efecte, en tots els exemples que hem considerat la component aleatòria és força evident: l'existència de nombrosos factors que no controlem (errors en els aparells de mesura, condicions ambientals, etc.) fa que encara que poguéssim repetir la sèrie temporal en idèntiques condicions, els resultats no serien exactament iguals. Sí que en canvi és raonable que, com en el llançament d'un dau, tinguéssim en la repetició d'una sèrie temporal una *regularitat estadística* global (és a dir, no només tirada a tirada, sinó en tot el recorregut). És, per tant, lògic proposar un model on l'evolució de la sèrie estigui regida per les lleis de la probabilitat, és a dir, un procés estocàstic.

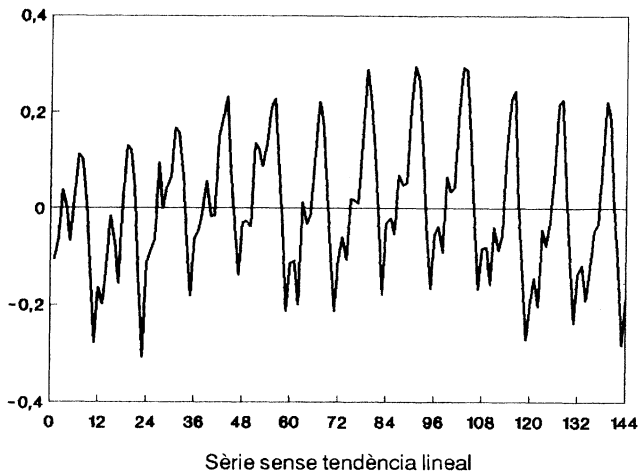
Molt sovint en una sèrie temporal (com la del nombre de passatgers) pot observar-se una *tendència* (en aquest cas creixent, més o menys lineal), unes fluctuacions degudes al mes (al mes d'agost hi ha un màxim local), que s'anomenen component estacional. Finalment, si treiem la tendència i la component *estacional*, queda un residu força irregular, que serà la component pròpiament aleatòria. El que hem fet, doncs, ha estat descompondre la sèrie observada x_t en suma:

$$x_t = T_t + S_t + z_t.$$

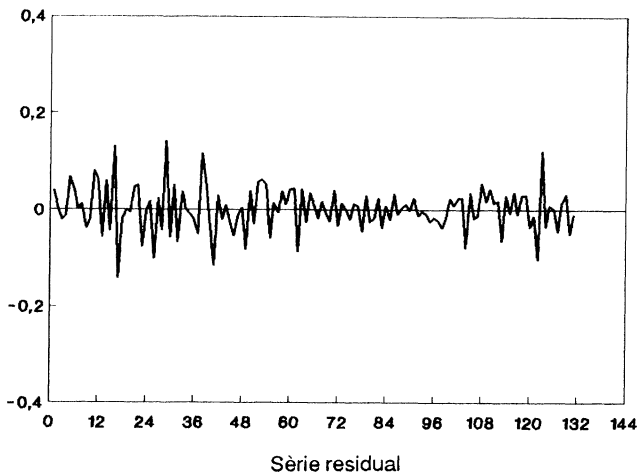
Concretament, en la sèrie del nombre de passatgers, atès que les oscil·lacions són cada cop més fortes, començarem transformant les dades aplicant logaritmes (nepèrians). La sèrie obtinguda és:



A continuació hi treiem la tendència lineal, i s'obté (noteu el canvi d'escala):



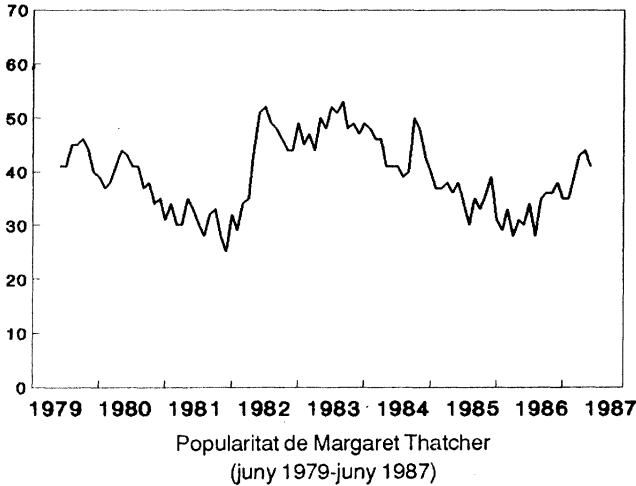
Finalment, si hi treiem la component estacional queda:



D. S. Moore explica una anècdota (força estranya, per cert) sobre la *desestacionalització* d'una sèrie temporal. La sèrie temporal del nombre d'aturats té una component estacional molt important. Als Estats Units de febrer a març hi ha creació de llocs de treball. Explica Mooer que el president Reagan l'abril de 1982 va anunciar que l'atur havia disminuït en 88.000 persones. Però una oficina d'estadística (Bureau of Labor Statistics) acabava d'anunciar un augment en el nombre d'aturats de 98.000. Naturalment la diferència esta en el fet que un utilitzava la sèrie normal, i els

altres, la sèrie ajustada sense la component estacional. Reagan va comentar: «Els estadístics de Washington tenen unes divertides maneres de comptar».

Una altra sèrie temporal interessant és la que dona la popularitat (% d'acceptació) de Margaret Thatcher des de juny de 1979 a juny de 1987. Noteu l'impacte de la guerra de les Malvines l'any 1982.



A l'estudi de les sèries temporals hi ha dos grans enfocaments: l'anomenada anàlisi en el *domini del temps*, que tracta d'utilitzar les relacions de dependència entre les distintes observacions, i dona lloc a la *metodologia Box-Jenkins*; i l'anàlisi (*en el domini de les freqüències o anàlisi espectral*) que —parlant informalment— és una versió estocàstica de l'anàlisi de Fourier.

L'ordre que seguirem serà el següent: primer repassarem breument els conceptes fonamentals de l'anàlisi de Fourier. Després estudiarem les sèries temporals com a processos estocàstics, i en veurem algunes propietats teòriques, tant en el domini del temps com en el de les freqüències. Aleshores començarem l'estudi estadístic (primer en el temps, al final en les freqüències): donades unes observacions (les taques solars, per exemple), quin model teòric dels estudiats és escaiant? Com es calculen els paràmetres que hi intervenen?

Tots els càlculs —excepte les suavitzacions de les sèries temporals i les bandes de confiança per a la densitat espectral de la secció 9— estan fets amb el programa d'ordinador Statgraphics. Les suavitzacions del periodograma estan fetes amb un full de càlcul a partir del periodograma calculat per l'Statgraphics, i les bandes de confiança estan calculades amb el programa BMDP (també es poden fer amb l'Statgraphics i el full de càlcul). Totes les sèries temporals simulades estan construïdes amb el full de càlcul a partir d'un *soroll blanc* generat per l'Statgraphics. Aquesta és una pràctica que recomanaria a tothom que vulgui aprofundir en l'estudi de les sèries temporals (o, en general de l'estadística): abans d'analitzar dades reals —exercici de

circ sense xarxa— convé educar la intuïció simulant unes dades, analitzant-les i comparant allò que obtenim amb el que havíem d'obtenir —exercici amb xarxa.

2. Anàlisi de Fourier de funcions reals

Per desenvolupar la intuïció del que després es farà començarem amb un estudi (no probabilístic) de les funcions periòdiques. Després iniciarem l'estudi de les sèries temporals com a processos estocàstics.

2.1 Funcions periòdiques

Tot i que, com hem dit, tractarem sempre amb sèries temporals amb temps discret, començarem estudiant unes propietats que s'interpreten més bé considerant funcions de variable de tipus continu; després ho particularitzarem al cas de successions de nombres. Considerem una funció real de variable real.

$$\begin{aligned}x: \mathbb{R} &\rightarrow \mathbb{R} \\ t &\rightarrow x(t)\end{aligned}$$

que per analogia amb les notacions de les sèries temporals escriurem $\{x(t), t \in \mathbb{R}\}$. Direm que aquesta funció és **periòdica** de *període* p si

$$x(t+p) = x(t), \quad \forall t \in \mathbb{R},$$

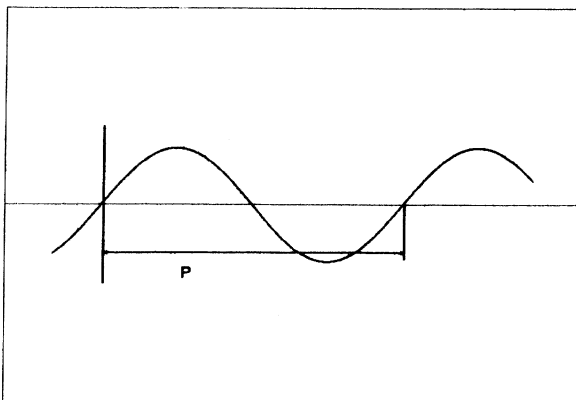
i p és el menor nombre que compleix aquesta condició. És a dir, cada p unitats de temps, la funció repeteix exactament la mateixa estructura. Per exemple, donat $\lambda > 0$, la funció

$$x(t) = \sin(\lambda t)$$

és periòdica de període $p = \frac{\lambda}{2\pi}$.

El valor de λ es $\sin(\lambda t)$ (o $\cos(\lambda t)$) s'anomena la *frequència*, i es mesura en radians per unitat de temps¹.

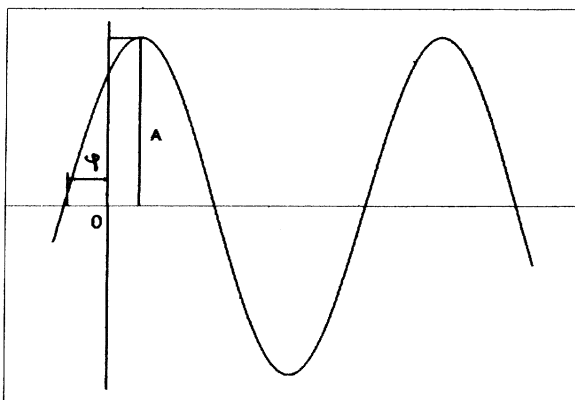
1. Habitualment la paraula *frequència* s'utilitza per designar l'invers del període, i es mesura en cicles per unitat de temps (si el temps està en segons, la unitat és l'hertz = 1 cicle/seg). La quantitat que nosaltres hem anomenat *frequència*, s'acostuma a anomenar *frequència angular*. Evidentment es passa d'una definició a l'altra multiplicant (o dividint) per 2π .



$$x(t) = \sin(\lambda t)$$

Introduint dos paràmetres més, la *fase* φ i l'*amplitud* A , s'obtenen una infinitat de funcions periòdiques amb el mateix període:

$$x(t) = A \sin(\lambda t + \varphi).$$



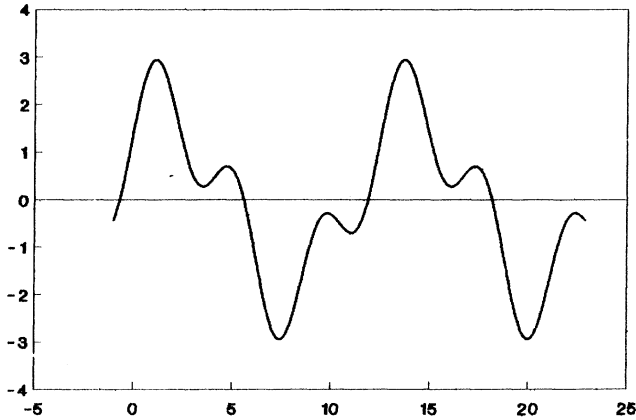
Així, la funció

$$x(t) = 2 \sin\left(\frac{t}{2} + \frac{\pi}{4}\right),$$

té període 4π i freqüència $1/2$. Si ara considerem la funció

$$x(t) = 2 \sin\left(\frac{t}{2} + \frac{\pi}{4}\right) + \sin\left(\frac{3}{2}t\right),$$

s'observa que té període 4π , però ara es diu que té freqüències $1/2$ i $3/2$.



El resultat fonamental de l'anàlisi de Fourier estableix que tota funció periòdica de període p prou bona (per exemple, tal que $\int_{-p}^p x^2(t) dt < \infty$) es pot desenvolupar en una suma (possiblement infinita), que s'anomena representació en sèrie de Fourier o representació espectral:

$$x(t) = \sum_{j=0}^{\infty} A_j \sin(\lambda_j t + \varphi_j),$$

on $\lambda_j = \frac{2\pi j}{p}$.

La freqüència $\lambda_1 = 2\pi/p$ s'anomena *freqüència fonamental*, mentre que les altres freqüències λ_j , $j \geq 2$ s'anomenen els *harmònics*. Aquesta representació és la versió matemàtica de la descomposició de la llum per un prisma en les seves components monocromàtiques: cada color correspon a una freqüència; o bé a la descomposició en harmònics d'un so.

Si utilitzem ara que

$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta,$$

la representació anterior es podrà escriure

$$x(t) = \sum_{j=0}^{\infty} (a_j \sin(\lambda_j t) + b_j \cos(\lambda_j t))$$

on

$$a_j = A_j \cos \varphi_j, \quad b_j = A_j \sin \varphi_j$$

o encara, utilitzant l'exponencial complexa, i el fet que

$$e^{i\theta} = \cos \theta + i \sin \theta$$

s'obté

$$x(t) = \sum_{j=-\infty}^{\infty} c_j \exp\{i\lambda_j t\} \quad (1)$$

on

$$c_0 = A_0,$$

$$c_j = \frac{1}{2i} A_j \exp\{i\lambda_j\}, \text{ si } j \geq 1$$

i

$$c_{-j} = \bar{c}_j.$$

Observeu que aquesta representació fa intervenir *freqüències negatives*. Noteu també que la condició $c_{-j} = \bar{c}_j$ garanteix que la suma de la sèrie serà un nombre real.

Si ara, amb analogia amb certs fenòmens físics, introduïm l'*energia total* sobre l'interval $(-p, p)$ amb $\int_{-p}^p x^2(t) dt$, utilitzant les propietats de les funcions trigonomètriques es dedueix que

$$\int_{-p}^p x^2(t) dt = 2p \sum_{j=-\infty}^{\infty} |c_j|^2,$$

d'on l'energia per unitat de temps, que habitualment s'anomena *potència*, valdrà

$$\text{Potència} = \frac{1}{2p} \int_{-p}^p x^2(t) dt = 2p \sum_{j=-\infty}^{\infty} |c_j|^2. \quad (2)$$

Si ara es calcula la potència de la component $c_j \exp\{i\lambda_j t\}$ s'obté $|c_j|^2$, d'on l'expressió (2) indica que la potència total de la funció $x(t)$ s'obté sumant les potències corresponents a cada component, o en altres paraules, les potències de les diferents components no es barregen entre elles.

La representació de la funció

$$g(\lambda) = \begin{cases} |c_j|^2, & \text{si } \lambda = \lambda_j, j \in \mathbf{Z} \\ 0, & \text{en cas contrari.} \end{cases}$$

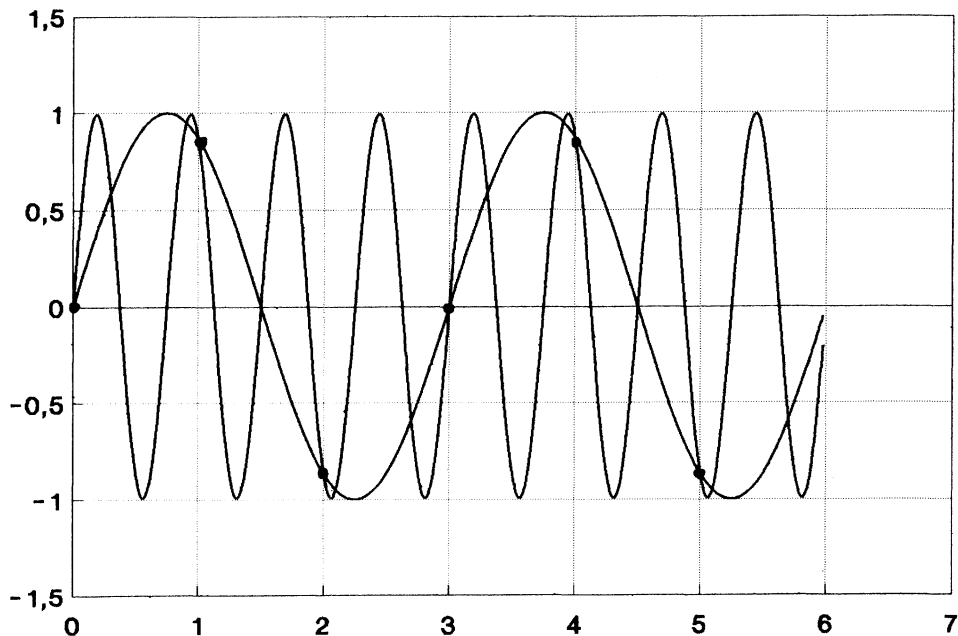
s'anomena l'*espectrograma* de $x(t)$. A causa de la condició $c_{-j} = \bar{c}_j$, es tracta d'una funció simètrica respecte de l'eix de les y . L'espectrograma és un gràfic realment interessant per estudiar una funció periòdica: ens diu les diferents freqüències que intervenen i la importància de cadascuna.

2.2 Anàlisi de Fourier d'una funció periòdica amb temps discret

Tota l'anàlisi anterior pot repetir-se per a una funció $x: \mathbf{Z} \rightarrow \mathbf{R}$. Però imaginem ara que la nostra funció és la restricció a \mathbf{Z} d'una funció definida en \mathbf{R} (aquesta és una situació molt habitual: el fenomen esdevé en temps continu, però nosaltres l'obssem

vem en temps discret). Aleshores es dóna una situació molt interessant que estudiarem mitjançant un exemple. Considerem les funcions següents:

$$x(t) = \sin\left(\frac{2\pi}{3}t\right) \quad \text{i} \quad x(t) = \sin\left(\frac{8\pi}{3}t\right).$$



S'observa que coincideixen en els punts enters.

Tornant a la situació general, partint de les observacions d'una funció en els enters, no es poden determinar les freqüències $\lambda > \pi$. Equivalentment, tota freqüència $> \pi$ té una freqüència en $[0, \pi]$ (que s'anomena *l'alias*) que dóna els mateixos valors en els punts $t \in \mathbf{Z}$. D'aquest fenomen se'n diu *aliasització*. Per aquest motiu, la representació espectral d'una funció periòdica de període π , $\{x(n), n \in \mathbf{Z}\}$, és de la forma

$$x(n) = \sum_{j=-\infty}^{\infty} c_j \exp\{i\tilde{\lambda}_j n\},$$

on $\lambda_j = \frac{2\pi j}{p} \in (-\pi, \pi]$. Llavors, l'espectrograma d'una funció amb temps discret inclou solament les freqüències $[0, \pi]$ (recordeu la simetria per estar treballant amb funcions reals).

2.3 Funcions no periòdiques

Algunes funcions no periòdiques prou bones poden considerar-se com una funció periòdica de període infinit, i aleshores pot intentar-se una representació en sèrie de Fourier. Donada la funció $\{x(t), t \in \mathbb{R}\}$, la idea consisteix a considerar, per a cada $p > 0$, la restricció a l'interval $(-p, p]$, i estendre-la per periodicitat a tot \mathbb{R} :

$$x_p(t) = x(t), \quad \text{si } t \in (-p + k, p + k), \quad k \in \mathbf{Z}.$$

Cada funció $x_p(t)$ es representa en sèrie de Fourier, i es fa un pas a límit quan $p \rightarrow \infty$. Suposant que, per exemple, $\int_{-\infty}^{\infty} x^2(t) dt < \infty$, s'obté una representació:

$$x(t) = \int_{-\infty}^{\infty} c(\lambda) e^{i\lambda t} d\lambda. \quad (3)$$

Cal notar que aquesta representació indica que a la funció $x(t)$ hi intervenen una infinitat no numerable de freqüències. També, utilitzant les propietats de l'exponencial complexa (i amb les bones condicions que estem suposant), l'energia total sobre \mathbb{R} és

$$\int_{-\infty}^{\infty} x^2(t) dt = \int_{-\infty}^{\infty} |c(\lambda)|^2 d\lambda.$$

Ara, però, la potència,

$$\lim_{p \rightarrow \infty} \int_{-p}^p x^2(t) dt$$

serà, evidentment, zero. Però podem interpretar el factor $|c(\lambda)|^2 d\lambda$ com la contribució a l'energia total de les freqüències entre λ i $\lambda + d\lambda$. En altres paraules, la funció

$$g(\lambda) = |c(\lambda)|^2$$

pot considerar-se com una funció de densitat que dóna la importància de la contribució de cada freqüència a l'energia total. La representació d'aquesta funció també s'anomena espectrograma. També en aquest cas que estem considerant es tracta d'una funció simètrica.

En el cas d'una bona funció no periòdica amb temps discret $\{x(n), n \in \mathbf{Z}\}$ el comentari de l'apartat 2.2 també és veritat: el rang de freqüències estarà en $[0, \pi]$.

2.4 Observacions

1. Com hem vist, la diferència essencial entre una funció periòdica i una no periòdica consisteix en el fet que la periòdica té un nombre finit o numerable de freqüències (totes múltiples de la freqüència fonamental), mentre que la no periòdica en té una infinitat no numerable. L'anàlisi harmònica generalitzada de Wiener considera les anomenades **funcions quasi-periòdiques**, que són les que tenen un nombre finit o numerable de freqüències, però que no són múltiples d'una d'aquestes. Per exemple, la funció

$$x(t) = \sin t + \sin(\sqrt{2}t)$$

2. Ambdues representacions, la del cas periòdic i la del cas no periòdic (i també la de les quasi-periòdiques), poden unificar-se utilitzant integrals de Lebesgue. Podem escriure:

$$x(t) = \int_{-\infty}^{\infty} e^{i\lambda t} C(d\lambda), \quad (4)$$

on C representa una mesura —discreta en un cas, contínua en l'altre— en \mathbb{R} a valors complexos. Integrar pot interpretar-se com amitjanar una funció d'acord amb una altra funció (la mesura) que assigna diferents pesos.

3. Les representacions espectrals (2), (3) o (4) tenen a l'esquerra una funció del temps, i a la dreta una funció de les freqüències. Aquestes són les dues grans òptiques de les sèries temporals: una, que estudia les propietats en el temps (*domini del temps*), i l'altra, les relacions amb les freqüències (*domini de les freqüències*). Tot i que sembla que estudiant un domini n'hi hauria d'haver prou, a la pràctica es veu que cal complementar els dos.

3. Processos estocàstics

3.1 Conceptes generals

Un **procés estocàstic** és una família de variables aleatòries $\{X_t, t \in T\}$ definides en un espai de probabilitat (Ω, \mathcal{F}, P) . El conjunt T s'anomena conjunt d'índexs, i generalment és \mathbb{R}^+ o \mathbb{N} , encara que nosaltres utilitzarem \mathbb{Z} .

Potser l'exemple més senzill és una successió de tirades d'una moneda: cada minut $t = 1, 2, \dots$ una persona tira una moneda. Si surt cara apunta un 1, si surt creu -1 . Sigui X_n el resultat de la tirada n -èsima. El procés estocàstic $\{X_n, n \in \mathbb{N}\}$ quedarà determinat (des de el punt de vista de les probabilitats) coneixent les lleis de totes les variables X_n , i totes les lleis conjuntes d'un nombre finit de variables X_{n_1}, \dots, X_{n_k} (això no és pas del tot trivial de demostrar), que s'anomenen les *distribucions en dimensió finita del procés*. En aquest cas les distribucions en dimensió finita són trivials de calcular perquè hi ha independència entre les variables.

Però així com a l'estadística clàssica se suposa sempre independència entre les distintes observacions, en els processos estocàstics es tracta sempre de treballar en situacions on hi hagi dependència entre les variables aleatòries (l'exemple anterior és, en aquest sentit, una excepció). Considerem, per exemple, l'anomenada passejada aleatòria: una persona decideix cada minut tirant una moneda si fer un pas endavant o enrera. Sigui Y_n la seva posició a l'instant n . Evidentment, la seva posició en cada moment dependrà de la seva posició en l'instant anterior. Si surt del zero, tindrem:

$$\begin{aligned} Y_0 &= 0 \\ Y_1 &= X_1 \\ Y_2 &= X_1 + X_2 \\ &\vdots \\ Y_n &= \sum_{j=1}^n X_j \end{aligned}$$

on $\{X_n, n \in \mathbb{N}\}$ és el procés estocàstic de l'exemple anterior. Per al procés $\{Y_n, n \geq 0\}$, també és senzill de calcular les distribucions en dimensió finita, però ara caldrà tenir en compte les posicions en els instants anteriors.

Malgrat aquesta idea d'una persona que cada minut tira una moneda per decidir cap on fa el pas següent, hi ha una altra manera més fèrtil (i profunda) de mirar un procés estocàstic: es tracta de considerar totes les possibles trajectòries d'aquesta persona, i escollir-ne una a l'atzar.

De forma general, donat un procés estocàstic $\{X_t, t \in T\}$, cada X_t és una variable aleatòria, és a dir, una aplicació mesurable

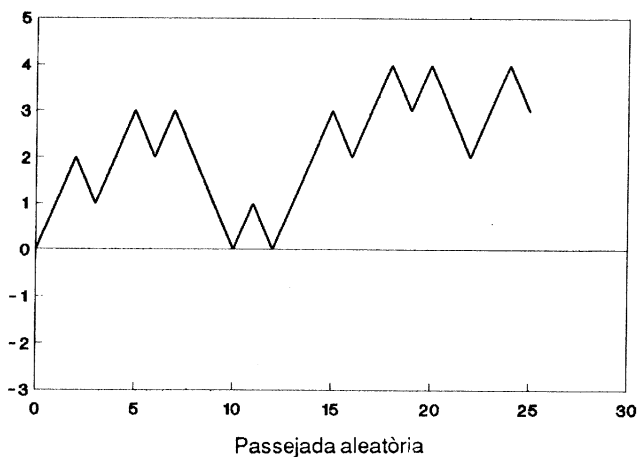
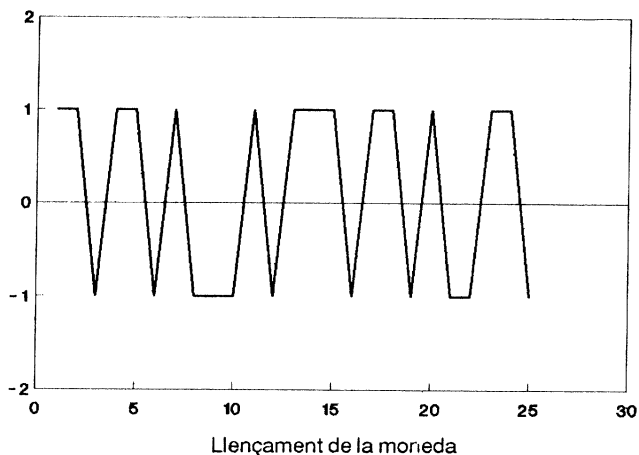
$$X_t : \Omega \rightarrow \mathbb{R}$$

Fixat ω , l'aplicació

$$X_t(\omega) : T \rightarrow \mathbb{R}$$

s'anomena la *trajectòria* associada a ω , o una *realització* del procés estocàstic. El procés estocàstic —les distribucions en dimensió finita— determinen una probabilitat a l'espai de les trajectòries $\{f : T \rightarrow \mathbb{R}\}$.

Una trajectòria dels exemples anteriors seria



3.2 Processos estocàstics estacionaris

Habitualment l'estudi de les sèries temporals comporta la realització de prediccions sobre el futur. Interessa llavors assegurar que el comportament (probabilístic) de la sèrie en el futur serà anàleg al comportament del passat. Aquesta idea és la d'**estacionarietat**. Nosaltres treballarem sempre amb processos de *segon ordre*, és a dir, tals que totes les variables aleatòries tenen variància finita (i, per tant, esperança). En aquesta situació l'estacionarietat que utilitzarem fa referència al fet que totes les variables tinguin la mateixa esperança i que la covariància entre dues variables només depengui de la distància entre ambdues. Concretament, direm que un procés $\{X_n, n \in \mathbb{Z}\}$

de segon ordre és *estacionari* si totes les variables tenen la mateixa esperança: per a tot $n \in \mathbf{Z}$,

$$E[X_n] = m,$$

i la covariància entre X_n i X_{n+k} només depèn de k :

$$\begin{aligned} Cov(X_1, X_2) &= Cov(X_2, X_3) = Cov(X_3, X_4) = \dots, \\ Cov(X_1, X_3) &= Cov(X_2, X_4) = \dots, \\ &\vdots \end{aligned}$$

Vegeu la nota² per a una altra noció relacionada amb l'estacionarietat.

Anomenarem **funció d'autocovariància** del procés a

$$C(k) = Cov(X_n, X_{n+k}) = E[(X_n - m)(X_{n+k} - m)], \quad k = 0, 1, \dots,$$

i **funció d'autocorrelació** a

$$\rho(k) = \frac{C(k)}{\sqrt{var(X_n)var(X_{n+k})}} = \frac{C(k)}{C(0)}, \quad k = 0, 1, \dots$$

El nombre k en $C(k)$ o $\rho(k)$ s'anomena *lag (retard)*. La representació de $\{\rho(k), k = 0, 1, \dots\}$, s'anomena **correlograma**.

A l'exemple primer, de la successió de tirades d'una moneda, donada la independència tenim

$$\rho(0) = 1, \quad \rho(k) = 0, \quad \forall k \geq 1.$$

La passejada aleatòria no és estacionària

$$Cov(Y_1, Y_{1+k}) = 1, \quad Cov(Y_2, Y_3) = 2.$$

2. Direm que un procés estocàstic és *estrictament estacionari* si per a qualsevol $k \in \mathbf{Z}$, la distribució de X_{r_1}, \dots, X_{r_n} i la de $X_{r_1+k}, \dots, X_{r_n+k}$ és la mateixa, és a dir, el comportament probabilístic del procés és invariant per translacions. En particular, totes les variables han de tenir la mateixa llei. En els exemples anteriors, el procés de les tirades de monedes és estrictament estacionari, mentre que la passejada aleatòria no ho és; per exemple, la llei de Y_1 és

$$P\{Y_1 = 1\} = P\{Y_1 = -1\} = \frac{1}{2},$$

mentre que la llei de Y_2 és

$$P\{Y_2 = 2\} = P\{Y_2 = -2\} = \frac{1}{4}, \quad P\{Y_2 = 0\} = \frac{1}{2}.$$

3.2 Alguns models de processos estocàstics estacionaris: el soroll blanc i els processos autoregressius

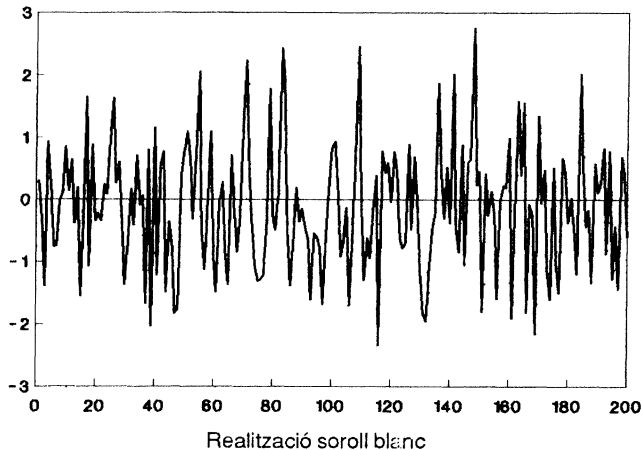
3.2.1 Soroll blanc

El procés estacionari més senzill és aquell que consisteix en una successió de variables aleatòries centrades, amb la mateixa variància i incorrelacionades dos a dos. Es a dir, és un procés $\{Z_n, n \in \mathbf{Z}\}$ tal que:

- (i) $E[Z_n] = 0, \quad \forall n \in \mathbf{Z}.$
- (ii) $E[Z_n^2] = \sigma^2, \quad \forall n \in \mathbf{Z}.$
- (iii) $Cov(Z_n, Z_m) = 0, \quad n \neq m.$

Llavors la funció d'autocorrelació valdrà, com al primer exemple,

$$\rho(0) = 1, \quad \rho(k) = 0, \quad \forall k \geq 1.$$



El terme *soroll blanc* procedeix de l'enginyeria, amb analogia amb la *llum blanca* on intervenen tots els colors amb les mateixes intensitats. Quan vegem l'anàlisi espectral dels processos estocàstics, veurem que en un soroll blanc l'espectrograma dóna exactament la mateixa intensitat a totes les freqüències.

3.2.2 Procés autoregressiu de primer ordre: AR(1)

Un procés estacionari $\{X_n, n \in \mathbf{Z}\}$ direm que és *autoregressiu de primer ordre* (AR(1)) si compleix una relació del tipus següent:

$$X_n = \phi X_{n-1} + Z_n, \quad (1)$$

on $\{Z_n, n \in \mathbf{Z}\}$ és un soroll blanc i ϕ és una constant, $|\phi| < 1$. Noteu que si $\phi = 1$ s'obténdria una passejada aleatòria, que com hem dit no és estacionari.

El nom ve del fet que l'equació que defineix el procés és anàloga a una equació de regressió lineal. Si observem el procés a l'instant $n-1$, aleshores l'observació a l'instant n ve a ser com una regressió sobre l'observació anterior.

Un exemple on podria aparèixer un procés d'aquest estil consistiria en un model (molt simplificat!) dels estalvis d'una persona: suposem que cada any una persona estalvia una quantitat Z_n (aquí suprimirem la condició $E[Z_n] = 0$ per evitar interpretacions difícils), designem per X_n el capital que té. Si no gasta res i Hisenda li reté cada any el 25%, aleshores

$$\begin{aligned} X_1 &= Z_1, \\ X_n &= 0.75X_{n-1} + Z_n \end{aligned}$$

En aquest exemple també es posa de manifest una situació habitual: Tot i que teòricament és molt avantatjós treballar amb un procés estocàstic indexat per \mathbf{Z} , a les aplicacions s'indexa per \mathbf{IN} , i l'equació d'autoregressió (1) s'escriu per a $n \geq 1$, i X_0 una variable inicial.

Si ara iterem l'equació d'autoregressió (1) s'obté:

$$\begin{aligned} X_n &= \phi^2 X_{n-2} + \phi Z_{n-1} + Z_n \\ &= \dots \\ &= \phi^k X_{n-k} + \phi^{k-1} Z_{n-k+1} + \dots + Z_n. \end{aligned}$$

Utilitzant les propietats d'estacionarietat del procés es demostra que

$$X_n = \sum_{k=0}^{\infty} \phi^k Z_{n-k}, \quad \text{convergència en mitjana quadràtica.} \quad (2)$$

Aquesta expressió pot interpretar-se de manera que a cada instant el procés és *causat* per la contribució de tots els *xocs* del soroll blanc anteriors, però, como que $|\phi| < 1$, la contribució és menys important com més antiga és. En general, un procés estacionari que admeti una representació de l'estil

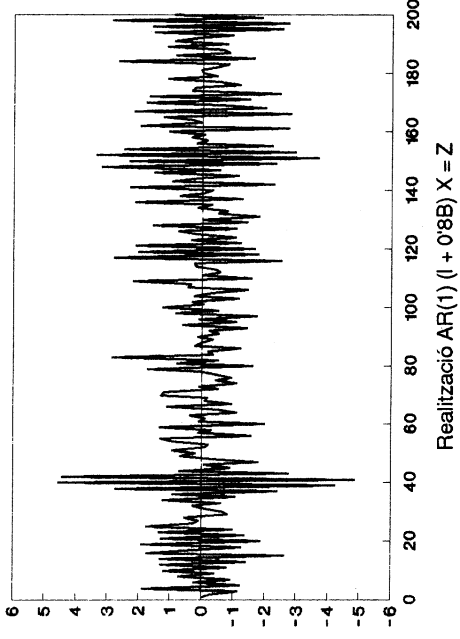
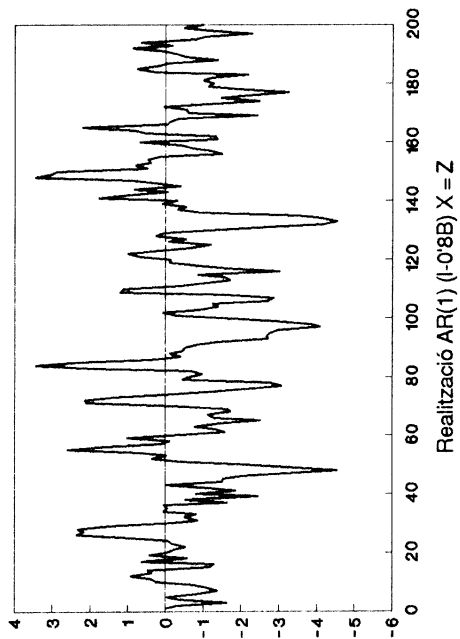
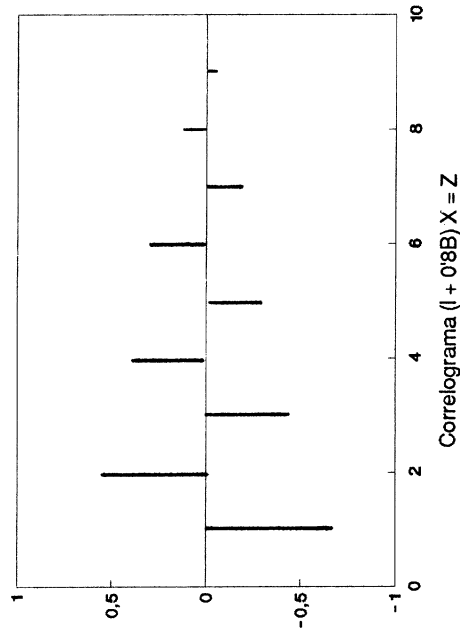
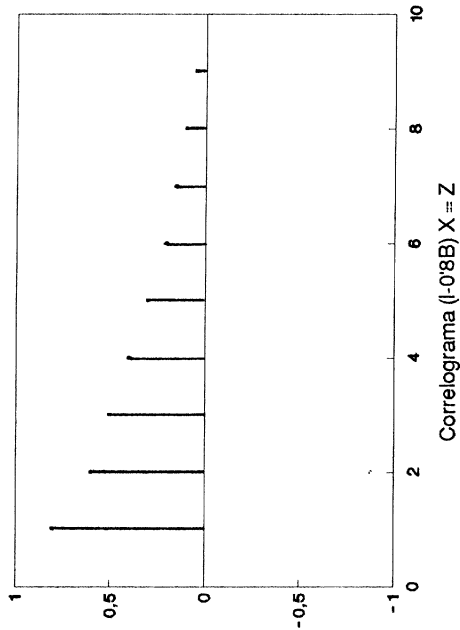
$$X_n = \sum_{k=0}^{\infty} \phi_k Z_k \quad (3)$$

amb $\sum_{k=0}^{\infty} |\phi_k| < \infty$ es diu que és *causal*.

Es senzill de calcular que $E[X_n] = 0, \forall n \in \mathbf{Z}$: també, si treballem una mica amb $E[X_1 X_{k+1}]$ s'obté que l'autocorrelació ρ_k ha de complir l'equació en diferències

$$\rho_k = \phi \rho_{k-1},$$

$$\rho(k) = \phi^k, \quad k \geq 0.$$



Per treballar amb aquests processos és convenient introduir l'operador de retard (*backward*) B definit per

$$BX_n = X_{n-1}.$$

L'equació d'autoregressió (1) s'escriurà:

$$(I - \phi B)X_n = Z_n.$$

Si posem $\Phi(x) = 1 - \phi x$, podem escriure:

$$\Phi(B)X = Z.$$

El polinomi $\Phi(x) = 1 - \phi x$ s'anomena *polinomi característic* del procés. Noteu que la condició $|\phi| < 1$ implica que l'arrel d'aquest polinomi, $\eta = 1/\phi$ valor absolut $|\eta| < 1$ (òbviament suposem $\phi \neq 0$.)

Si introduïm ara els operadors de retard iterats:

$$B^k X_n = B^{k-1}(BX_n) = X_{n-k},$$

l'expressió (2) pot escriure's

$$X_n = \sum_{k=0}^{\infty} \phi^k B^k(Z_n).$$

Notem que, formalment, el que hem fet ha estat invertir l'operador $I - \phi B$, ja que com que $|\phi| < 1, \forall x \in [-1, 1], |\phi x| < 1$, aleshores,

$$(1 - \phi x)^{-1} = \sum_{k=0}^{\infty} \phi^k x^k$$

(suma d'una progressió geomètrica de raó de valor absolut < 1).

3.2.3 Procés autoregressiu de segon ordre (AR(2))

Direm que un procés estacionari $\{X_n, n \in \mathbf{Z}\}$ és *autoregressiu d'ordre 2* si compleix una equació de l'estil

$$X_n = \phi_1 X_{n-1} + \phi_2 X_{n-2} + Z_n,$$

on $\{X_n, n \in \mathbf{Z}\}$ és un soroll blanc. Es a dir, X_n és una regressió lineal sobre X_{n-1} i X_{n-2} . Amb ajuda de l'operador de retard s'escriu

$$(I - \phi_1 B - \phi_2 B^2)X_n = Z_n,$$

o encara

$$\Phi(B)X = Z,$$

on $\Phi(x) = 1 - \phi_1 x - \phi_2 x^2$ és el polinomi característic del procés. Es comprova que les autocorrelacions han de complir l'equació en diferències

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k \geq 1. \quad (4)$$

Particularitzant-la per a $k = 1$ i $k = 2$ s'obtenen les anomenades *equacions de Yule-Walker*

$$\left. \begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 \end{aligned} \right\}.$$

(Cal notar que com que el procés és real, $\rho_{-1} = \rho_1$). Es tracta d'un sistema *bilineal*, és a dir, donats ϕ_1 i ϕ_2 , és lineal en ρ_1 i ρ_2 , i a l'inrevés.

Però no solament ρ_1 i ρ_2 poden obtenir-se en funció de ϕ_1 i ϕ_2 , sinó que l'equació en diferències (4) es pot resoldre i obtenir una expressió general per a ρ_k . Per això es resol l'equació

$$1 - \phi_1 x - \phi_2 x^2 = 0.$$

Si η_1 i η_2 són les equacions (reals distintes o complexes conjugades; si té una arrel doble el raonament és semblant), aleshores les solucions de (4) són

$$\rho_k = a_1 \xi_1^k + a_2 \xi_2^k$$

on $\xi_1 = \frac{1}{\eta_1}$ i $\xi_2 = \frac{1}{\eta_2}$, on a_1 i a_2 es determinen per les condicions inicials

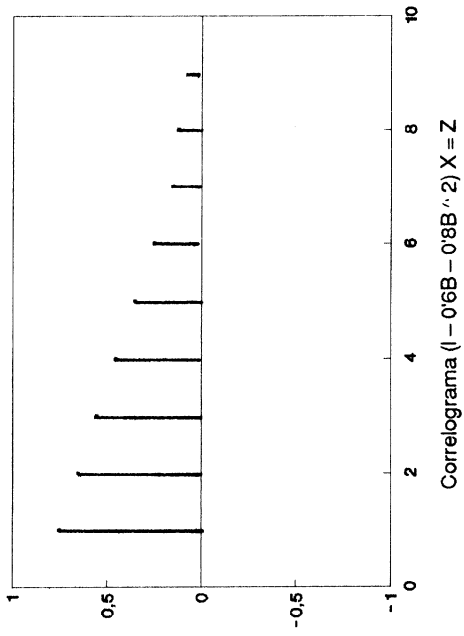
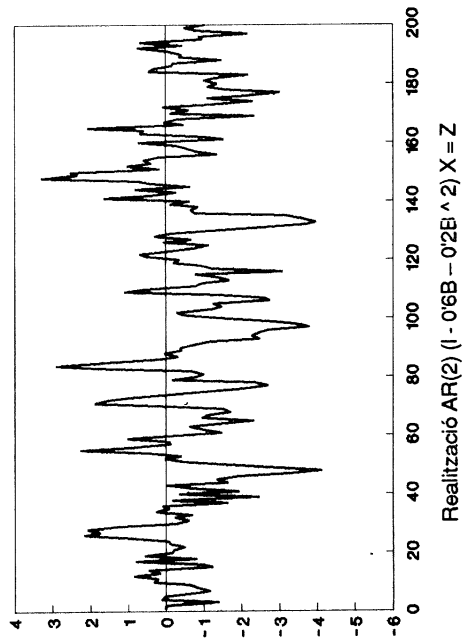
$$\begin{aligned} \rho_0 &= 1, \\ \rho_1 &= \frac{\phi_1}{1 - \phi_2}. \end{aligned}$$

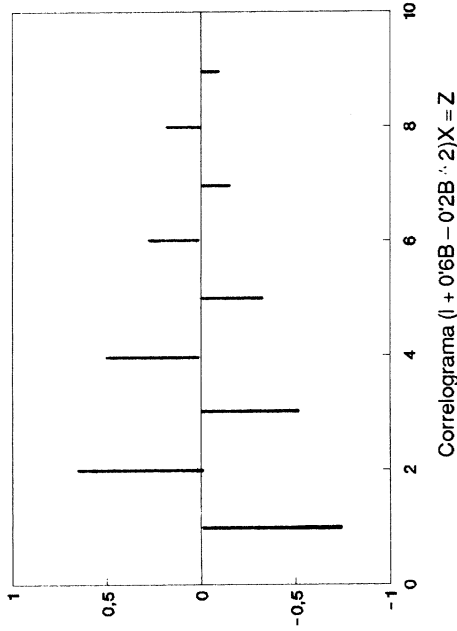
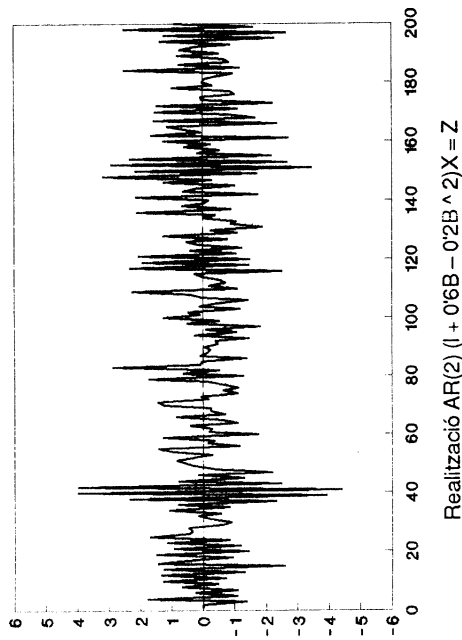
Perquè les solucions ρ_k siguin nombres raonables ($|\rho_k| < 1$) cal que $|\xi_1| < 1$ i $|\xi_2| < 1$, o equivalentment, cal que les arrels del polinomi característic $\Phi(x) = 1 - \phi_1 x - \phi_2 x^2$ estiguin fora del cercle unitat: $|\eta_1| < 1$ i $|\eta_2| > 1$. Sobre els coeficients, tot això exigeix que

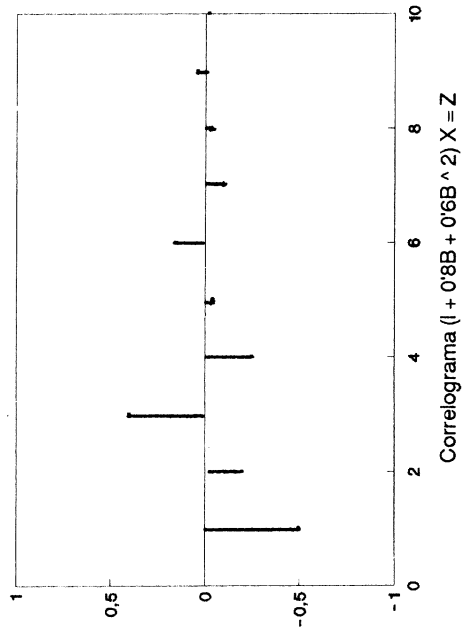
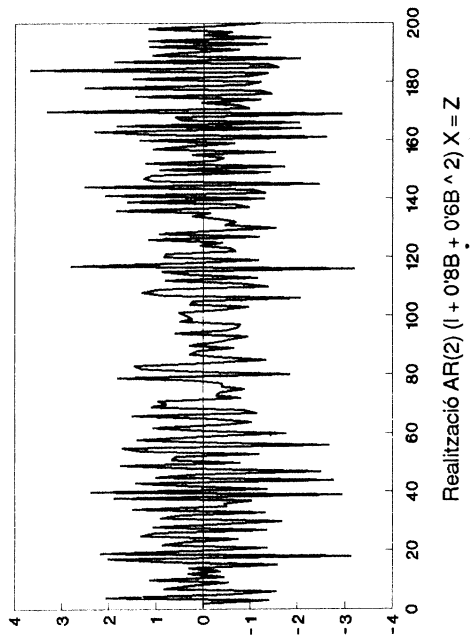
$$\begin{aligned} -1 &< \phi_2 < 1 \\ \phi_1 + \phi_2 &< 1 \\ \phi_2 - \phi_1 &< 1. \end{aligned}$$

Concretant més les solucions ρ_k s'obté:

- Si η_1 i η_2 són reals, aleshores ρ_k és una suma d'exponencials.
- Si η_1 i η_2 són complexos, aleshores ρ_k decreix en forma sinusoidal.







Assenyallem també que, sota aquestes condicions, el procés es podrà escriure

$$X_n = \sum_{k=0}^{\infty} \varphi_k Z_k \quad (3)$$

amb $\sum_{k=0}^{\infty} |\varphi_k| < \infty$ (procés causal). Es tracta, de nou, d'invertir l'operador $\Phi(B)$.

3.2.4 Procés autoregressiu d'ordre p (AR(p))

Es tracta d'un procés estacionari $\{X_n, n \in \mathbf{Z}\}$ que compleix una equació de l'estil

$$X_n = \phi_1 X_{n-1} + \dots + \phi_p X_{n-p} + Z_n,$$

on $\{Z_n, n \in \mathbf{Z}\}$ és un soroll blanc.

S'escriu també

$$\Phi(B)X = Z,$$

on

$$\Phi(x) = 1 - \phi_1 x - \dots - \phi_p x^p.$$

Les autocorrelacions han de complir l'equació en diferències

$$\Phi(B)\rho_k = 0,$$

que per a $k = 1, \dots, p$ dóna les equacions de Yule-Wolker

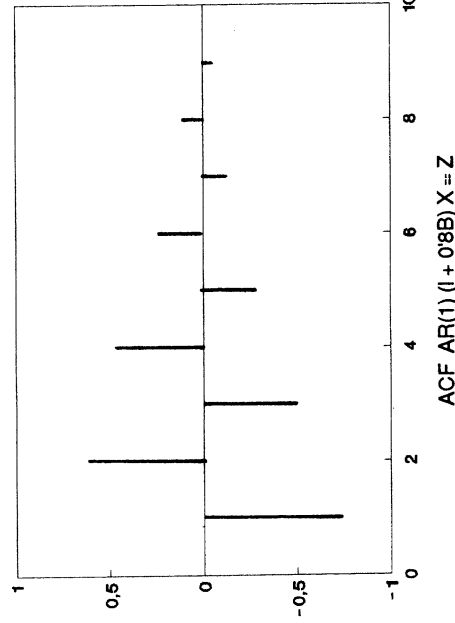
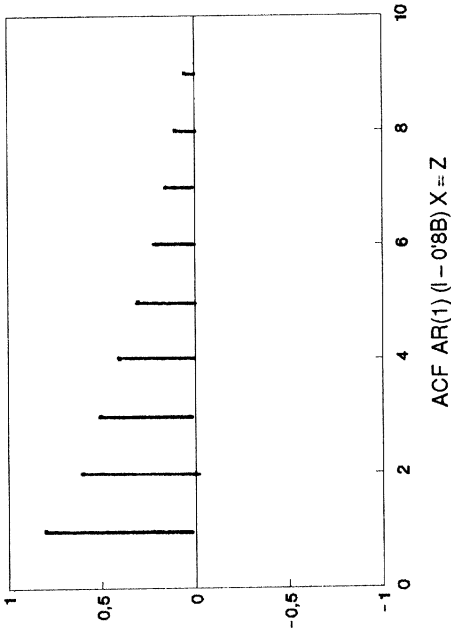
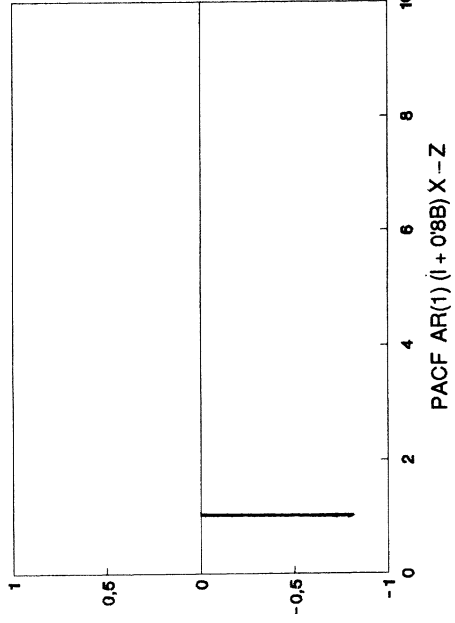
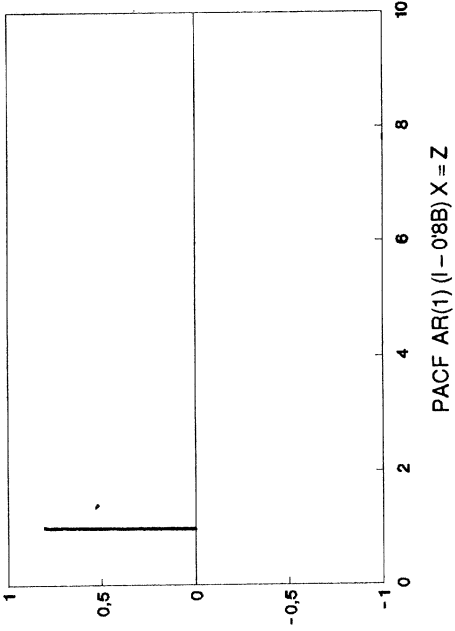
$$\left. \begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \dots + \phi_p \end{aligned} \right\}$$

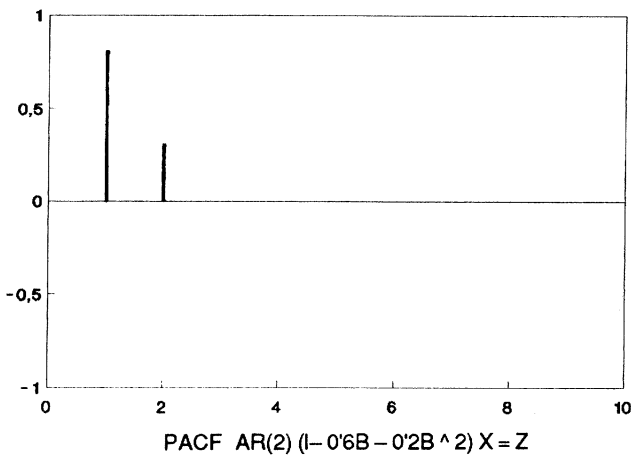
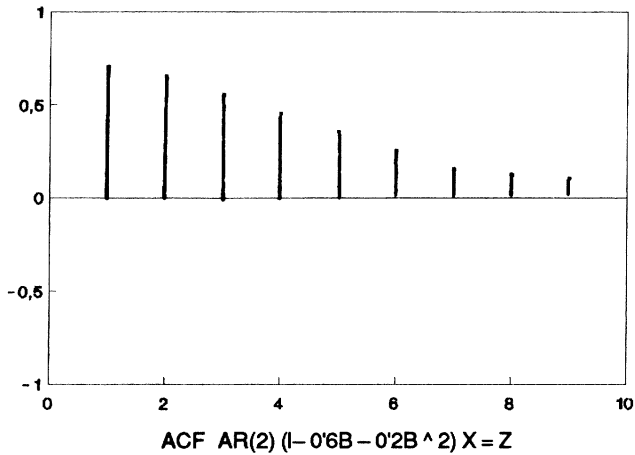
Raonant igual que al procés AR(2), perquè les solucions de l'equació en diferències no explotin, cal que les arrels del polinomi característic $\Phi(x)$ estiguin fora del cercle unitat. També en aquest cas serà un procés causal.

3.3 Funció d'autocorrelació parcial

Per tractar d'afinar més en l'estudi de l'estructura de dependència d'un procés estacionari s'introdueixen uns altres coeficients (i una altra funció). La idea és que quan s'estudia la influència del passat en un procés AR(1) i en un AR(2) la situació és molt diferent. En un procés AR(1)

$$X_n = \phi_1 X_{n-1} + Z_n$$



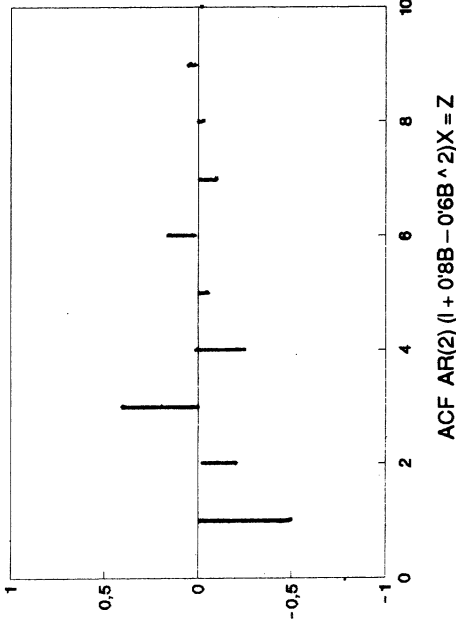
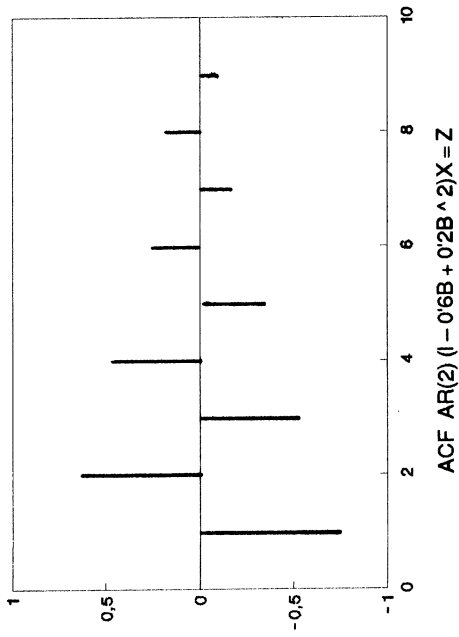
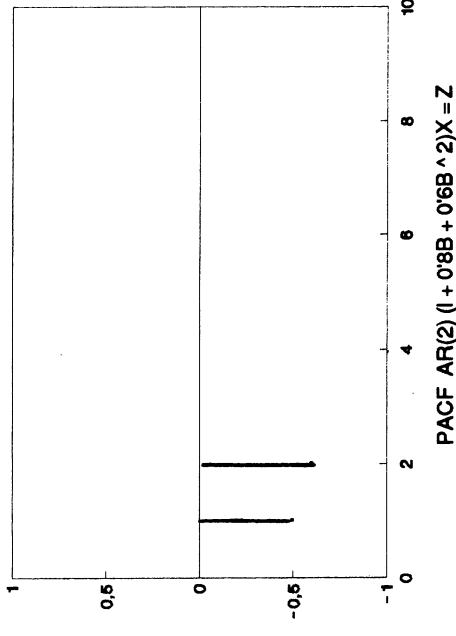
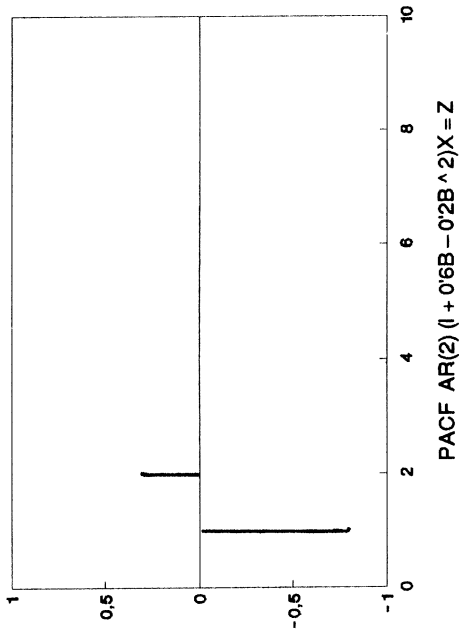


el terme X_{n-2} només influeix en X_n a través de X_{n-1} . En altres paraules, conegut X_{n-1} , el coneixement de X_{n-2} no aporta informació nova (caràcter markovià dels processos AR(1)). Al contrari, en un procés AR(2)

$$X_n = \phi_1 X_{n-1} + \phi_2 X_{n-2} + Z_n$$

X_{n-2} té una influència directa en X_n (i una d'indirecta a través de X_{n-1}). Aquest fet no es detecta amb la funció de correlació. Aleshores, per avaluar la influència directa d'una observació X_{n-k} sobre X_n eliminant totes les influències indirectes a través de $X_{n-k+1}, \dots, X_{n-1}$ s'introdueix el *coeficient d'autocorrelació parcial amb retard* (lag) k $\alpha(k)$: per a $k = 1$,

$$\alpha(1) = \text{Corr}(X_1, X_2) = \rho(1).$$



Per a $k \geq 2$,

$$\alpha(k) = \text{Corr}(X_{k+1} - R[X_{k+1}/X_2, \dots, X_k], X_1 - R[X_1/X_2, \dots, X_k]),$$

on $R[X_{k+1}/X_2, \dots, X_k]$ és la regressió lineal de X_{k+1} sobre X_2, \dots, X_k .

La funció $\alpha(k)$, $k \geq 1$ s'anomena *funció d'autocorrelació parcial*. En la bibliografia anglosaxona, la funció d'autocorrelació s'abreuja ACF (Autocorrelation Function), mentre que la d'autocorrelació parcial és PACF (Partial Autocorrelation Function).

Es demostra que un AR(p) té

$$\alpha(1) \neq 0, \dots, \alpha(p) \neq 0, \quad \alpha(k) = 0, \quad k \geq p + 1.$$

3.4 Processos de mitjana mòbil MA (q)

Direm que un procés estacionari $\{X_n, n \in \mathbf{Z}\}$ és un *procés de mitjana mòbil d'ordre q* si compleix una equació del tipus

$$X_n = Z_n + \theta_1 Z_{n-1} + \dots + \theta_q Z_{n-q},$$

on $\{Z_n, n \in \mathbf{Z}\}$ és un soroll blanc. Amb l'operador de retard s'escriu

$$X = \Theta(B)Z$$

on $\Theta(x) = 1 + \theta_1 x + \dots + \theta_q x^q$. Aquí, d'entrada, no cal imposar cap condició perquè el procés sigui estacionari.

L'expressió *mitjana mòbil* no és gaire afortunada. Fa referència al fet que per calcular X_n hem d'amitjanar Z_n, \dots, Z_{n-q} , ponderats amb uns pesos θ_1, θ_q , i per calcular X_{n+1} hem de fer la mateixa mitjana, però movent tots els Z una unitat. Aquesta és, com veurem més endavant, una tècnica fonamental que s'utilitza per *suavitzar* una sèrie temporal.

La funció d'autocorrelació dona

$$\rho(k) = \begin{cases} \sum_{j=0}^{q-k} \theta_j \theta_{j+k}, & \text{si } k = 0, \dots, q, \\ 0, & \text{si } k \geq q + 1. \end{cases}$$

Estudiem amb més detall el procés MA(1). Tindrem que .

$$X_n = Z_n + \theta_1 Z_{n-1}.$$

Suposem $|\theta_1| < 1$. Si iterem indefinidament el procés, o invertim formalment l'operador $(I + \theta_1 B)$, s'obté:

$$\sum_{j=0}^{\infty} \theta_1^j X_{n-j} = Z_n$$

És a dir, hem escrit el procés com un AR(∞), amb coeficients tals que

$$\sum_{k=0}^{\infty} |\theta_1^k| < \infty$$

Es diu aleshores que el procés és *invertible*. Intuïtivament això vol dir que l'efecte del passat decreix amb el temps. També això posa de manifest que la funció d'autocorrelació parcial serà sempre diferent de zero (qualsevol Z_{n-k} té influència directa sobre X_n), però també que $\alpha(k) \rightarrow 0$ quan $k \rightarrow \infty$.

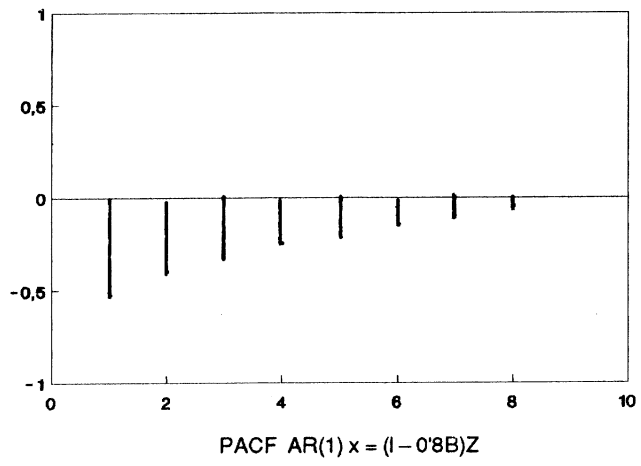
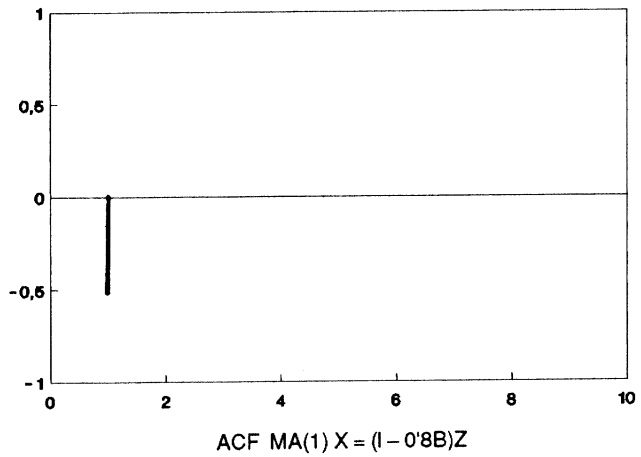
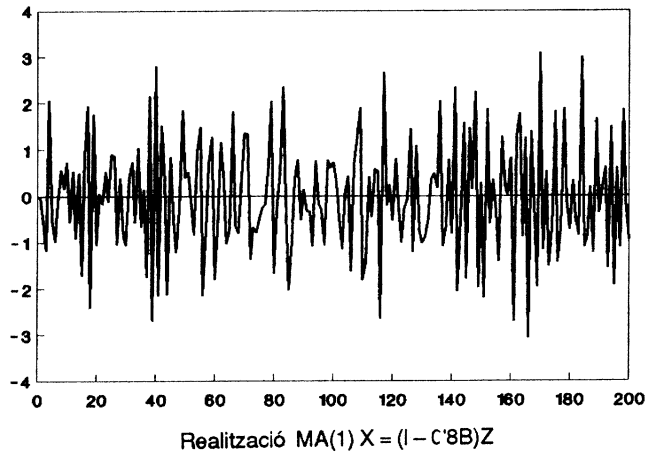
Aquest raonament pot repetir-se per a qualsevol mitjana mòbil MA(q). La condició d'invertibilitat serà que el polinomi característic $\Theta(x)$ tingui totes les arrels fora del cercle unitat. A partir d'ara suposarem que tots els processos de mitjana mòbil amb les quals treballarem compleixen aquesta condició.

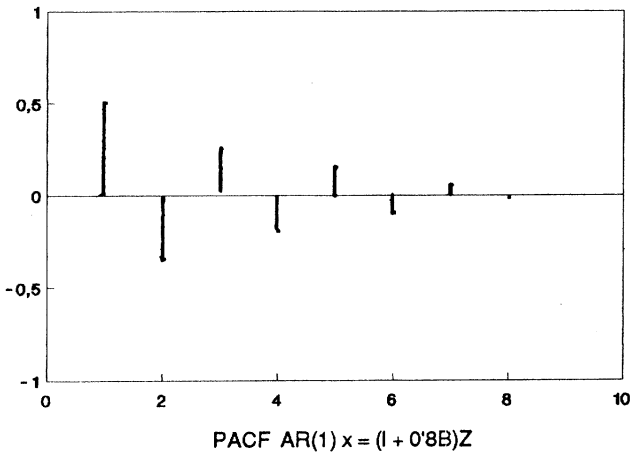
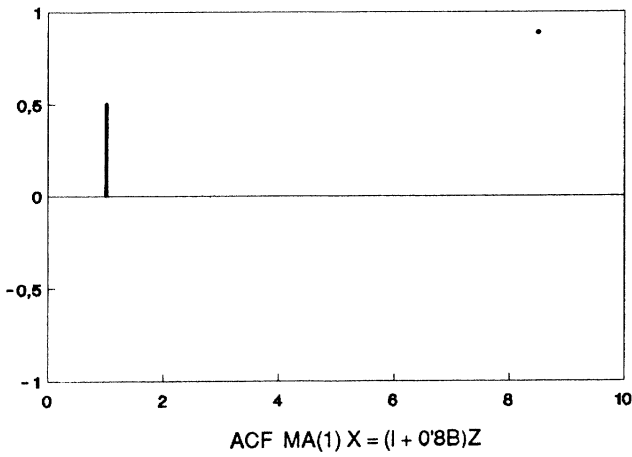
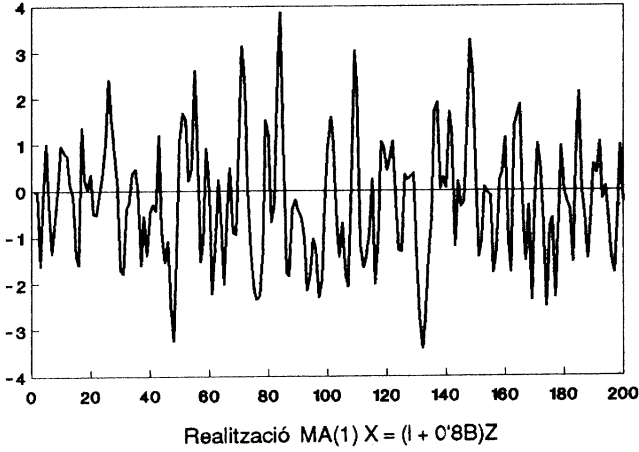
Haviem comentat que donat un procés AR(p) (amb les arrels del polinomi característic fora del cercle unitat) el podem escriure

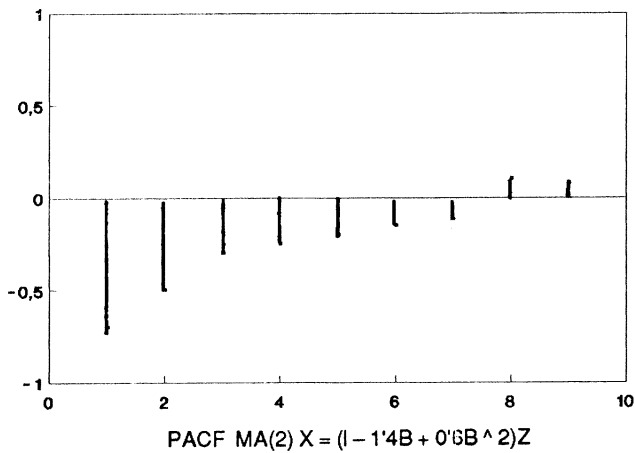
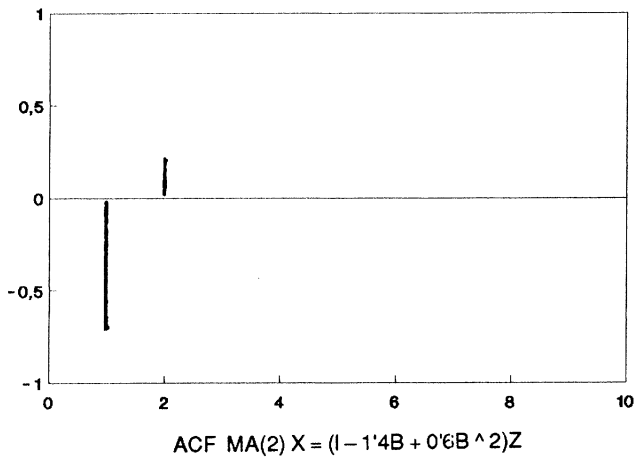
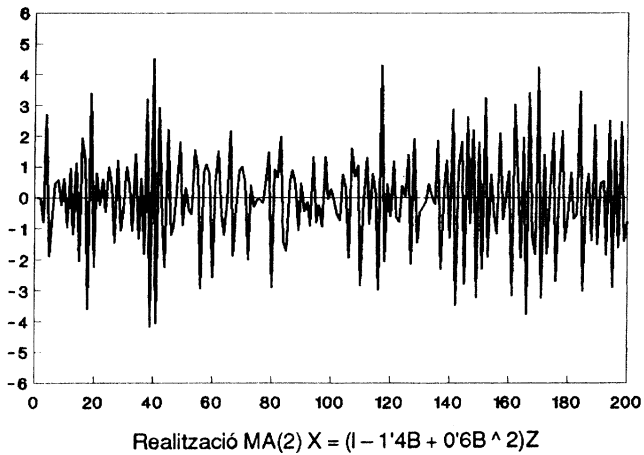
$$X_n = \sum_{k=0}^{\infty} \varphi_k Z_k \tag{3}$$

amb $\sum_{k=0}^{\infty} |\varphi_k| < \infty$, és a dir, com un procés MA(∞). Així podem dir de manera informal que tenim una dualitat:

$$\begin{aligned} AR(\infty) &\leftrightarrow MA(q) \leftrightarrow AR(p) \leftrightarrow MA(\infty) \\ AC F &\leftrightarrow P AC F \\ P AC F &\leftrightarrow AC F \end{aligned}$$







3.5 Processos ARMA

Direm que un procés estacionari $\{X_n, n \in \mathbf{Z}\}$ és un *procés ARMA* (p,q) si compleix una equació del tipus

$$X_n - \phi_1 X_{n-1} - \dots - \phi_p X_{n-p} = Z_n + \theta_1 Z_{n-1} + \dots + \theta_q Z_{n-q},$$

on $\{Z_n, n \in \mathbf{Z}\}$ és un soroll blanc. Amb l'operador de retard s'escriu

$$\Phi(B)X = \Theta(B)Z$$

on

$$\begin{aligned} \Phi(x) &= 1 - \phi_1 x - \dots - \phi_p x^p, \\ \Theta(x) &= 1 + \theta_1 x + \dots + \theta_q x^q \end{aligned}$$

i suposarem que $\Phi(x)$ i $\Theta(x)$ no tenen arrels comunes i tenen totes les arrels fora del cercle unitat.

El procés pot escriure's com un $MA(\infty)$ o com un $AR(\infty)$, i les funcions d'autocorrelació i autocorrelació parcial tenen una estructura complicada. En $\rho(k)$ el decreixement inicial ve dictat per l'ordre MA, i després per AR; en $\alpha(k)$, al principi per l'ordre AR i després per MA. En resum:

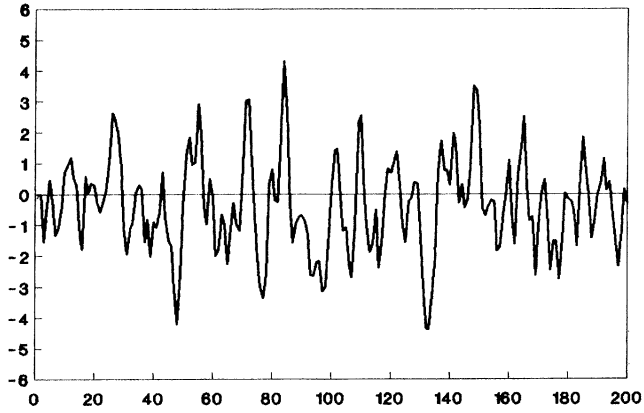
Procés	ACF (ρ)	PACF (α)
AR(p)	Decreixement ràpid a zero	$\alpha(k) \neq 0, k = 1, \dots, p$ $\alpha(k) = 0, k = p + 1, \dots$
MA(q)	$\rho(k) \neq 0, k = 1, \dots, q$ $\rho(k) = 0, k = q + 1, \dots$	Decreixement ràpid a zero
ARMA(p,q)	Decreixement ràpid a zero	Decreixement ràpid a zero

3.6 Processos harmònics

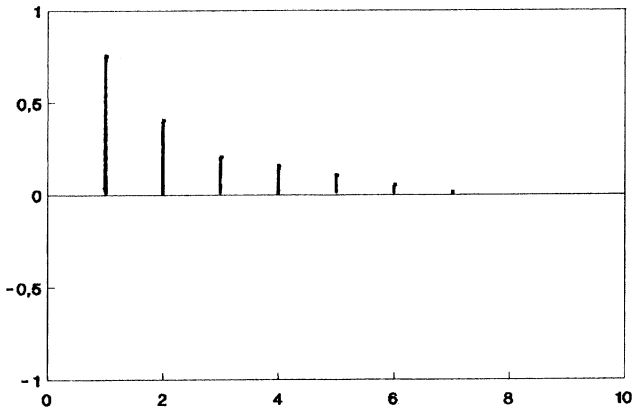
Donat $\lambda \in [0, \pi]$, $A \in \mathbb{R}$, i φ una variable aleatòria amb distribució uniforme en $(-\pi, \pi)$, definim el procés

$$X_n = A \sin(\lambda n + \varphi):$$

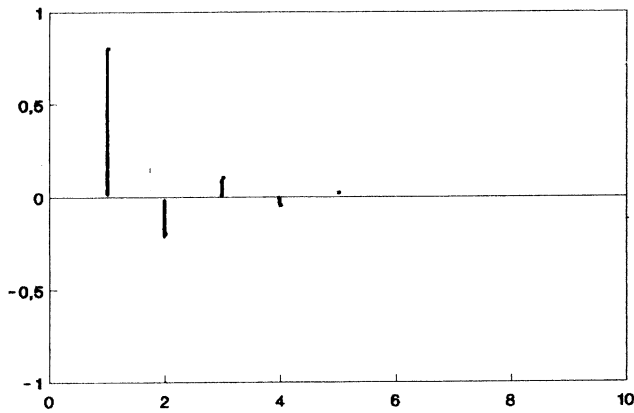
És a dir, es tracta d'una funció periòdica amb la fase aleatòria.



Realització ARMA (1,1) $(1 - 0.5B)X = (1 + 0.5B)Z$



ACF ARMA (1,1) $(1 - 0.5B)X = (1 + 0.5B)Z$



PACF ARMA (1,1) $(1 - 0.5B)X = (1 + 0.5B)Z$

El procés es pot escriure

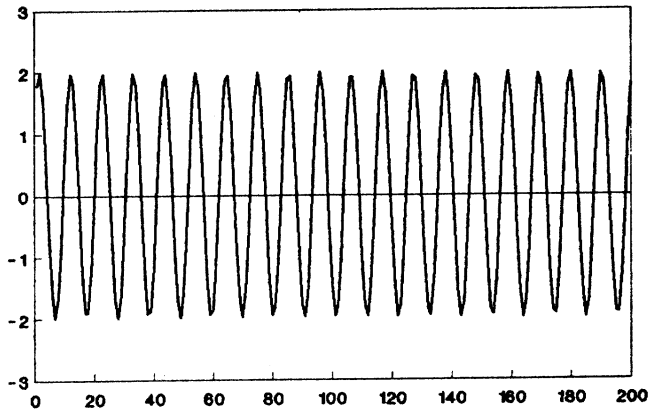
$$X_n = C e^{i\lambda n} + \bar{C} e^{-i\lambda n},$$

on C és una variable aleatòria (complexa). Es tracta d'un procés centrat ($E[X_n] = 0$) i estacionari: la funció de covariància dóna

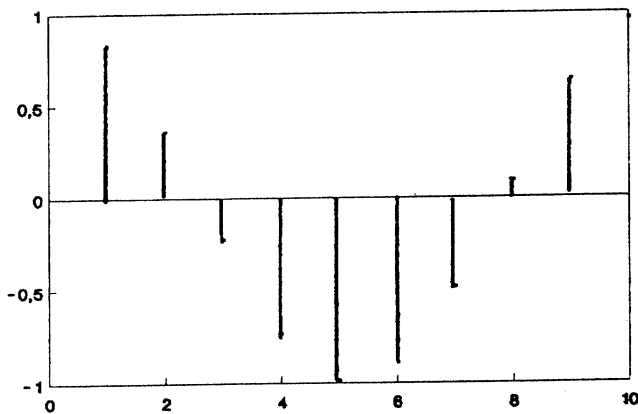
$$\gamma(n) = \frac{A^2}{2} \cos(\lambda n).$$

En general, un procés del tipus

$$X_n = \sum_{j=-k}^k C_j e^{i\lambda_j n},$$



Realització d'un procés harmònic



ACF procés harmònic

amb $\lambda_1, \dots, \lambda_j \in [0, \pi]$, $\lambda_{-j} = \lambda_j$, $C_{-j} = C_j$, amb C_1, \dots, C_k incorrelacionades dos a dos, i amb variància finita σ_j^2 , $j = 1, \dots, k$, s'anomena *procés harmònic*. És un procés centrat i amb funció de covariància

$$\gamma(n) = \sum_{j=-k}^k \sigma_j^2 e^{in\lambda_j},$$

4. Representació espectral de processos estacionaris

Comencem amb el procés harmònic de l'apartat anterior:

$$X_n = \sum_{j=-k}^{j=k} C_j e^{i\lambda_j n},$$

que podem escriure

$$X_n = \int_{-\pi}^{\pi} e^{i\lambda n} dC(\lambda),$$

on C és una mesura discreta (aleatòria) sobre $[-\pi, \pi]$ (a valors en \mathbb{C}) que dóna massa C_j a λ_j (recordeu que C_j és una variable aleatòria). La funció d'autocovariància es pot escriure

$$\gamma(n) = \int_{-\pi}^{\pi} e^{i\lambda n} dG(\lambda),$$

on G és la mesura discreta sobre $[-\pi, \pi]$ (real), que dóna massa σ_j^2 a λ_j .

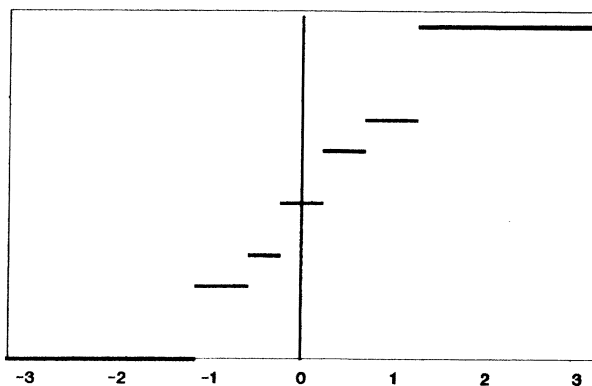
El resultat fonamental de la teoria espectral afirma que tot procés estacionari $\{X_n, n \in \mathbb{Z}\}$ pot representar-se:

$$X_n = \int_{-\pi}^{\pi} e^{i\lambda n} dC(\lambda),$$

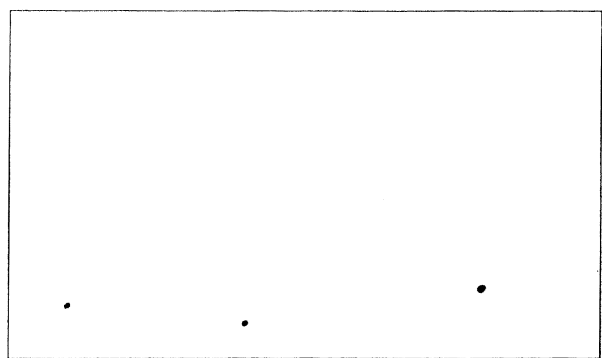
on aquesta integral és una integral estocàstica (que en general no pot interpretar-se de manera senzilla).

De fet, un resultat previ molt important és que del teorema d'Herglotz (1911) es dedueix que la funció d'autocovariància d'un procés estacionari pot escriure's:

$$\gamma(n) = \int_{-\pi}^{\pi} e^{i\lambda n} dG(\lambda),$$



Funció de distribució
de la mesura discreta G



Funció d'assignació de massa
de la mesura discreta G

on G és una mesura (real) en $[-\pi, \pi]$. G s'anomena *mesura espectral*, i si té densitat s'anomena *densitat espectral*, (i és simètrica respecte l'eix de les y). Com en el cas de les variables aleatòries, la manera d'unificar l'estudi de totes aquestes mesures és mitjançant les funcions de distribució: donada una mesura espectral G , anomenem *funció de distribució espectral* la funció $F : [-\pi, \pi] \rightarrow \mathbb{R}$ definida per

$$F(\lambda) = G([-\pi, \lambda]), \quad \lambda \in [-\pi, \pi].$$

Exemples

4.1 Procés harmònic

Com hem dit G és una mesura discreta. Es representa la funció d'assignació de massa

$$g(\lambda) = \begin{cases} \sigma_j^2, & \text{si } \lambda = \lambda_j, \\ 0, & \text{en cas contrari.} \end{cases}$$

(Normalment es representa la funció $g(2\pi c)$, $c \in [0, 0.5]$.)

Els processos harmònics també s'anomenen **processos amb espectre discret**.

4.2 Soroll blanc

En aquest cas,

$$\gamma(n) = \begin{cases} \sigma^2, & n = 0 \\ 0, & n \neq 0. \end{cases}$$

que es pot escriure

$$\gamma(n) = \frac{\sigma^2}{2\pi} \int_{-\pi}^{\pi} e^{i\lambda n} d\lambda$$

és a dir, amb densitat espectral

$$g(\lambda) = \frac{\sigma^2}{2\pi}, \quad \lambda \in [-\pi, \pi],$$

on σ^2 és la variància del soroll blanc.

4.3 Processos ARMA

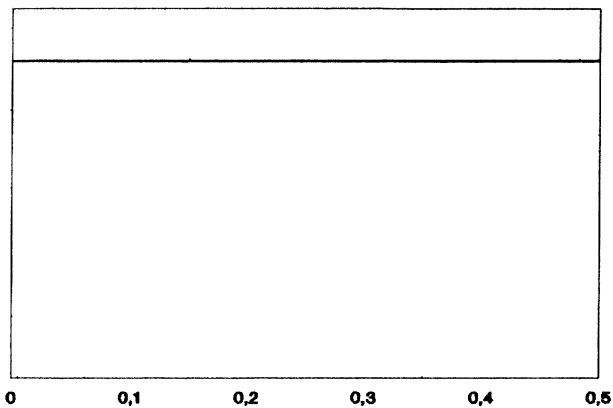
Un procés ARMA (p,q)

$$\Phi(B)X = \Theta(B)Z$$

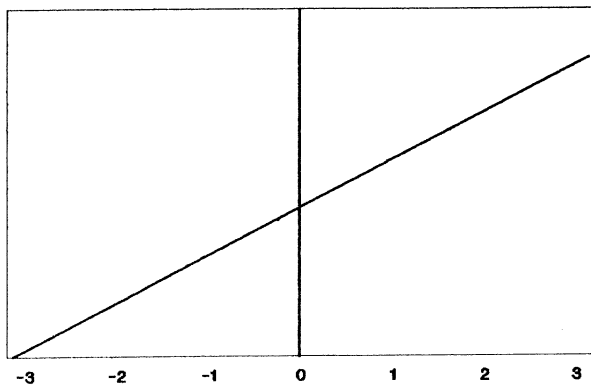
amb les bones condicions habituals té densitat espectral

$$g(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\Theta(e^{-i\lambda})|^2}{|\Phi(e^{-i\lambda})|^2}, \quad \lambda \in [-\pi, \pi],$$

on σ^2 és la variància del soroll blanc. Es diu que són processos amb **densitat espectral racional**.



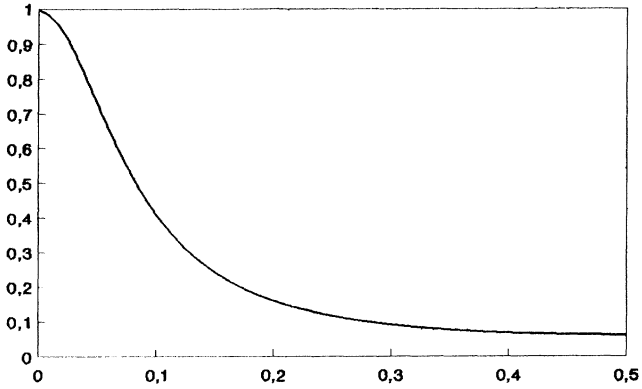
Densitat espectral del soroll blanc



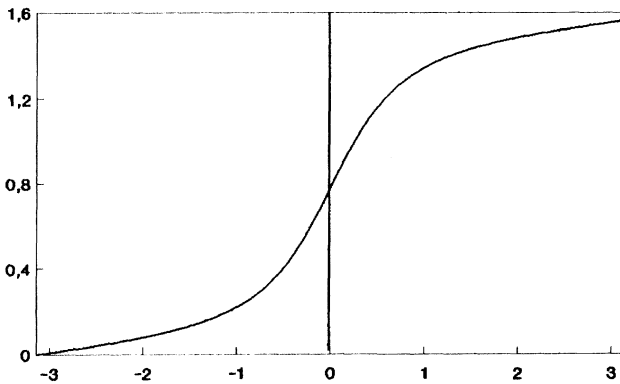
Funció de distribució del soroll blanc

En concret, un procés AR(1) té densitat espectral

$$g(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{1 - 2\phi \cos \lambda + \phi^2} .$$



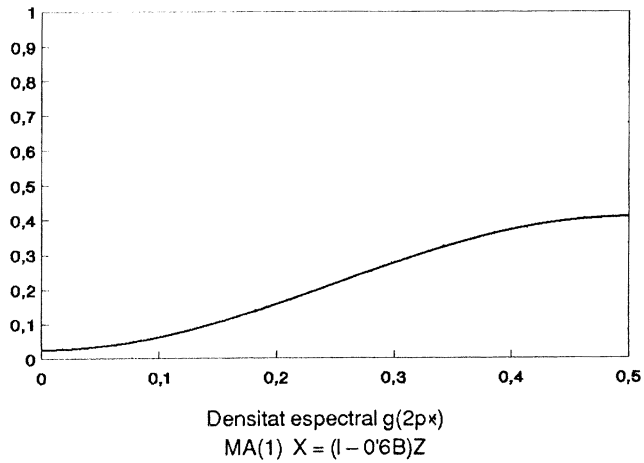
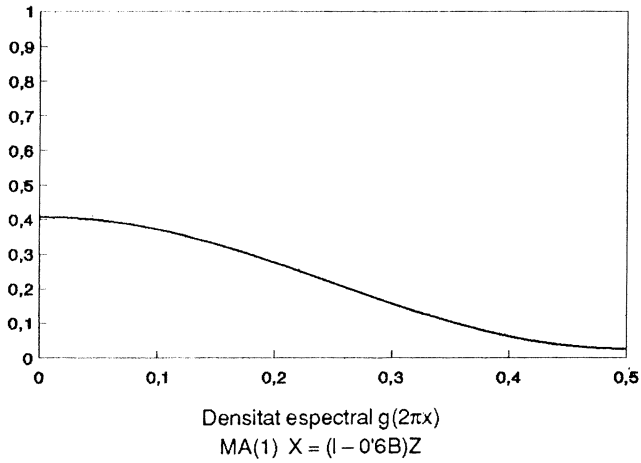
Densitat espectral $g(2\pi\kappa)$
AR(1) ($1 - 0.6B$) $X = Z$



Funció de distribució espectral
Sèrie AR(1)

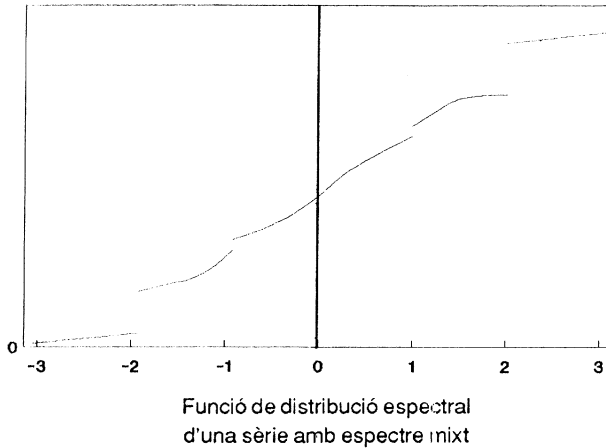
Un procés MA(1) té densitat espectral

$$g(\lambda) = \frac{\sigma^2}{2\pi} (1 + 2\theta \cos \lambda + \theta^2).$$



4.4 Processos amb espectre mixt

Com en el cas de les variables aleatòries, pot donar-se una situació mixta: per exemple una sèrie temporal que sigui suma d'una sèrie harmònica i d'una sèrie ARMA tindrà una mesura espectral que no serà ni discreta ni contínua. Es parla llavors de processos amb *espectre mixt*. La mesura espectral vindrà descrita per la funció de distribució.



Estadística de sèries temporals

5. Estimació de la funció d'autocorrelació d'un procés estacionari

Considerem un procés estacionari $\{X_n, n \in \mathbb{Z}\}$. A la pràctica només tindrem un nombre finit d'observacions X_1, \dots, X_n , i generalment no coneixerem la mitjana teòrica

$$E[X_n] = \mu,$$

ni la funció d'autocovariància. Ens interessarà trobar estimadors d'aquests elements i si és possible construir intervals de confiança i fer tests d'hipòtesis.

5.1 Estimació de la mitjana

Un estimador natural de μ serà

$$\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j.$$

(Si no hi ha confusió escriurem \bar{X} en lloc de \bar{X}_n). Es tracta d'un estimador sense biaix ($E[X_n] = \mu$), consistent en mitjana quadràtica ($\bar{X}_n \rightarrow \mu$ en L^2) i amb condicions addicionals (febles) asimptòticament normal ($\sqrt{n}(\bar{X}_n - \mu) \rightarrow N(0, \nu)$, en distribució).

5.2 Estimació de la funció d'autocovariància

Un estimador raonable de la funció d'autocovariància $\gamma(k)$ serà la funció d'autocovariància mostral

$$\hat{\gamma}(k) = \frac{1}{n-k} \sum_{j=1}^{n-k} (X_j - \bar{X})(X_{j+k} - \bar{X}),$$

$k = 0, \dots, n-1$. Per a la funció d'autocorrelació $\rho(k)$ s'utilitza

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)}.$$

La funció $\hat{\rho}(k)$ s'anomena *funció d'autocorrelació mostral o empírica*.

Aquests estimadors tenen biaix. D'altra banda, quan el retard creix, el nombre de sumands decreix (només un quan $K = n-1$) i aleshores les estimacions són cada cop menys fiables. Box i Jenkins diuen que aquestes estimacions són útils quan $n \geq 50$ i $k \leq n/4$.

Sota condicions raonables (però llargues d'escriure) es té la normalitat asimptòtica (conjunta) del vector $\hat{\rho}_n = (\hat{\rho}_n(1), \dots, \hat{\rho}_n(k))$: Si posem $\rho = (\rho(1), \dots, \rho(k))$, tenim que

$$\sqrt{n}(\hat{\rho}_n - \rho) \rightarrow \mathcal{N}(0, \Sigma), \text{ en distribució.}$$

El fet de conèixer aquesta distribució ens permet de calcular intervals de confiança (aproximat) per la funció d'autocorrelació. Habitualment els programes estadístics calculen intervals de confiança del 0,95.

5.3 Estimació de la funció d'autocorrelació parcial

L'estimació de la funció d'autocorrelació parcial es fa utilitzant les idees de la definició que hem donat: calcular el coeficient de correlació entre dues variables un

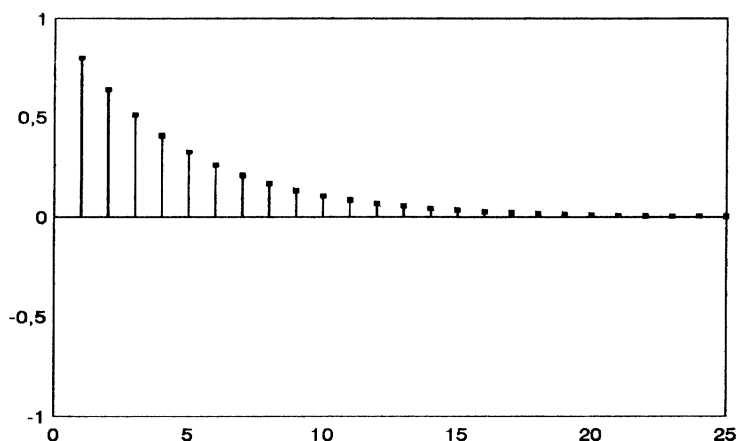
cop eliminada la influència de les variables intermèdies mitjançant regressions lineals. També es construeixen intervals de confiança.

Aquesta estimació s'anomena *funció d'autocorrelació parcial mostral o empírica*.

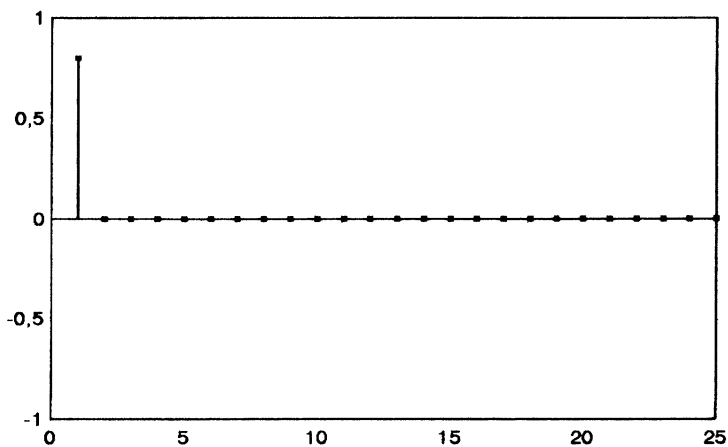
A continuació es representen diferents estimacions de les funcions d'autocorrelació i autocorrelació parcial, així com els intervals de confiança del 0'95, de la sèrie AR(1)

$$(1 - 0'8B)X = Z,$$

prenent diferents realitzacions, i canviant la mida de la mostra. En primer lloc, les funcions d'autocorrelació i autocorrelació parcial teòriques són les següents:

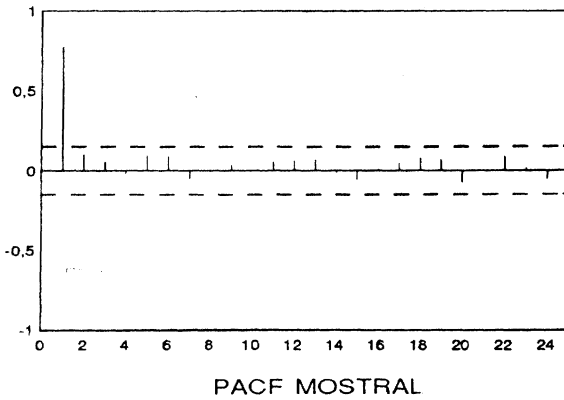
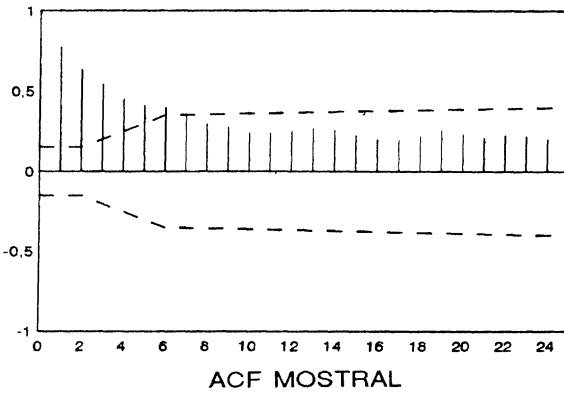
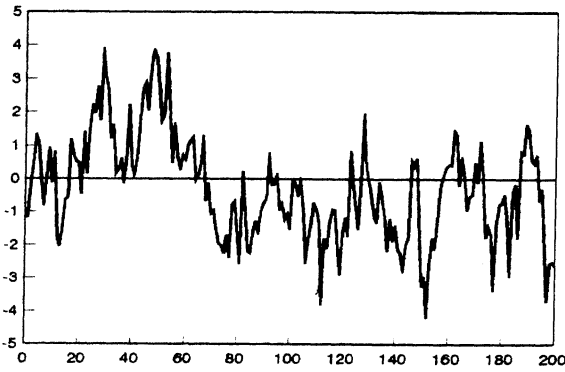


ACF AR(1) (1 - 0'8B)X = Z

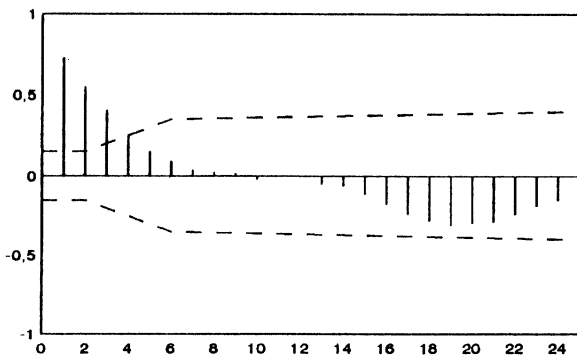
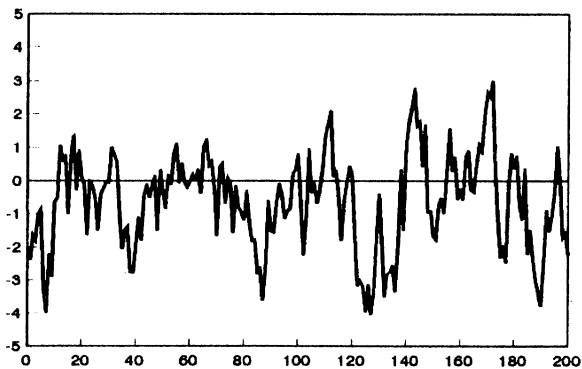


PACF AR(1) (1 - 0'8B)X = Z

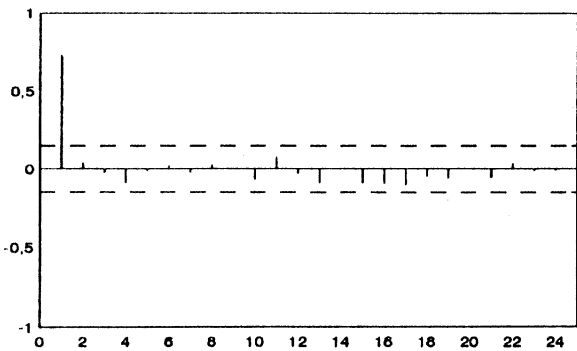
Realització de la sèrie AR(1): $(I - 0.8B)X = Z$
 $n = 200$



Realització de la sèrie AR(1): $(I - 0,8B)X = Z$
 $n = 200$

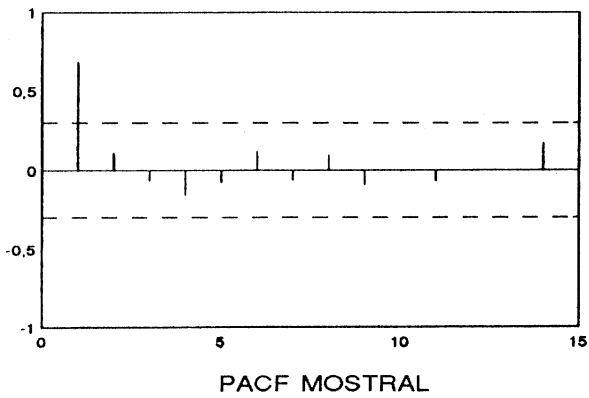
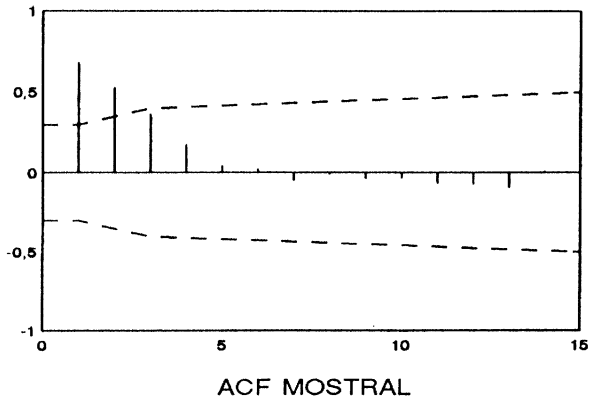
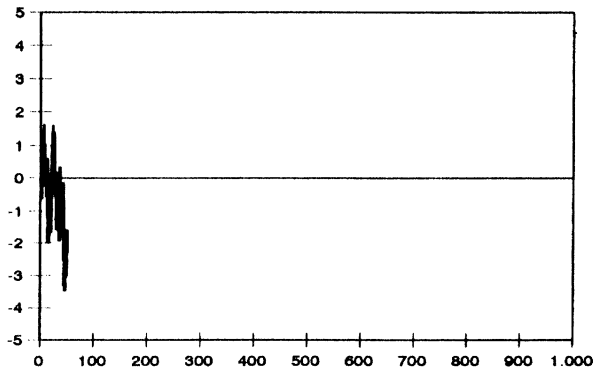


ACF MOSTRAL

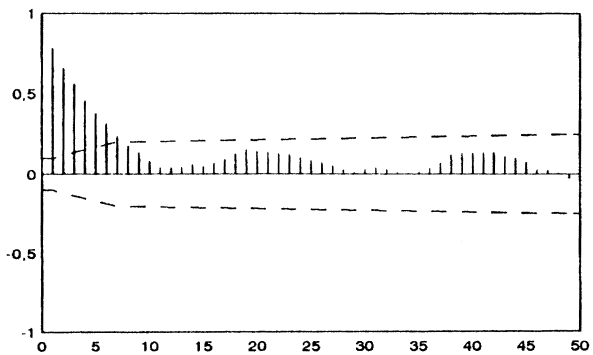
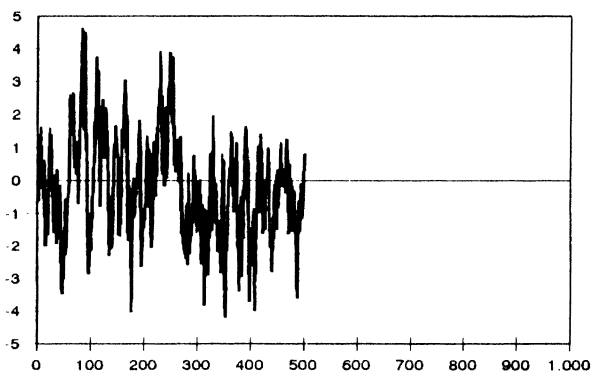


PACF MOSTRAL

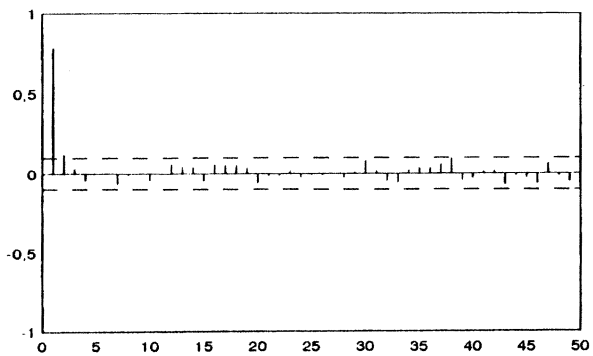
Realització de la sèrie AR(1): $(I - 0.8B)X = Z$
 $n = 50$



Realització de la sèrie AR(1): $(I - 0.8B)X = Z$
 $n = 500$

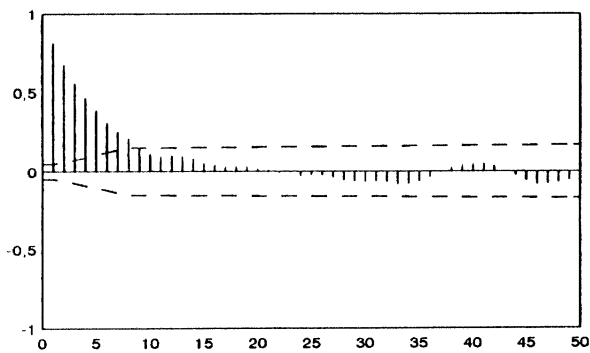
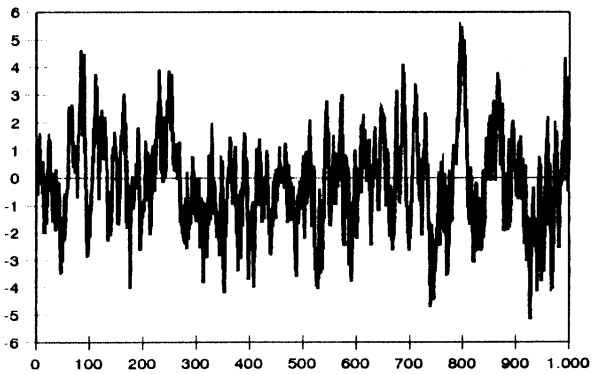


ACF MOSTRAL

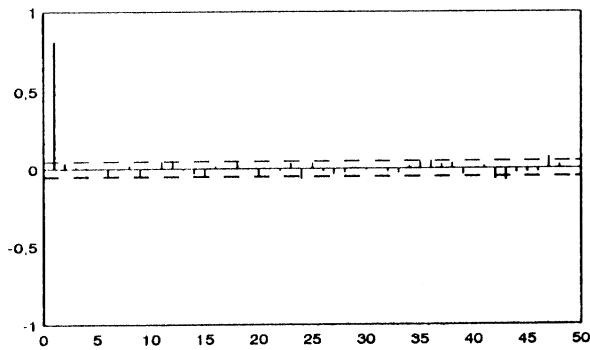


PACF MOSTRAL

Realització de la sèrie AR(1): $(I - 0.8B)X = Z$
 $n = 1000$



ACF MOSTRAL



PACF MOSTRAL

6. Estimació de paràmetres d'un procés ARMA

Les funcions d'autocorrelació i d'autocorrelació parcial mostrals no són estimadors gaire eficients, però ens proporcionen una primera idea sobre com són les funcions d'autocorrelació i autocorrelació parcial. Un cop calculades aquestes funcions mostrals, amb ajut del catàleg d'ACF i PACF teòriques es decideix si algun model ARMA s'adapta a la situació. Aquesta primera etapa s'anomena *identificació*.

Cal tenir present el perill de sobreparametritzar un model. Aparentment, com més grans s'escullen p i q millor serà l'ajust del model. Però això no és així: en primer lloc perquè caldrà estimar més paràmetres (amb el mateix nombre de dades!), i aleshores hi haurà més error. En segon lloc, a efectes de modelitzar la situació, la sobreparametrització pot portar a conseqüències no desitjables: Brockwell i Davis expliquen l'exemple següent: suposem que tenim 100 observacions de la sèrie temporal no estacionària

$$X_n = a + bn + Z_n,$$

on $\{Z_n\}$ és un soroll blanc. Podríem ajustar un polinomi de grau 99, i l'ajust seria perfecte. En utilitzar el model per fer prediccions, el resultat seria molt dolent. Box i Jenkins fan un fi comentari sobre el fet que una cosa és *ajustar* (*fitting*), i l'altra *fer previsions* (*forecasting*). El principi general utilitzat a estadística s'anomena de *parsimònia*: cal considerar el model més simple possible que expliqui les dades. Aleshores es recomana una actitud prudent i conservadora en la identificació, no decidint-se per un model complicat fins que els més simples poden ser rebutjats. Així, es comença sempre considerant models AR(1), MA(1), etc. També cal recordar les paraules de McCullagh i Nelder *Tots els models són dolents, però alguns són més útils que altres, i aquests són els que cal buscar*.

De fet, per resoldre aquest problema d'una manera sistemàtica, i evitar decidir a ull, s'han inventat diferents tècniques, com per exemple el criteri AIC d'Akaike, que intenta avaluar el cost d'introduir un nou paràmetre.

Un cop decidida l'estructura ARMA(p,q) es procedeix a estimar els paràmetres $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ i la variància del soroll blanc σ^2 . No entrarem en el detall del procediment per fer això. Només comentem que s'utilitza el mètode del màxim de versemblança o el dels mínims quadrats. Ambdós mètodes tenen dificultats tècniques (de fet utilitzen algorismes d'optimització no lineal), i arribar a una bona solució depèn de manera important de la qualitat de les estimacions inicials. Aquesta segona etapa s'anomena d'*estimació*. A cavall entre aquesta etapa i la següent (estudiar si el model s'adapta bé a les dades), pot procedir-se a la comprovació que el model proposat satisfà les condicions teòriques exigides: bàsicament, causalitat i invertibilitat, i que consisteixen a veure que els polinomis de les parts AR i MA tenen les arrels fora del cercle unitat. De fet, si alguna arrel té mòdul proper a 1 seria una bona indicació per canviar de model, i proposar un model ARIMA que veurem a la propera secció.

La darrera etapa s'anomena de *diagnòstic*. Es tracta de detectar si el model s'ajusta bé a les dades o no. En general el procediment per jutjar la bondat d'un ajustament consisteix a comparar els valors observats amb els valors que prediu el model, i estudiar si els *residus* es comporten de la manera esperada. Si la sèrie temporal proposada és

$$(I - \phi_1 B - \dots - \phi_p B^p) X_n = (I + \theta_1 B + \dots + \theta_q B^q) Z_n,$$

i estimem $\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q$, aleshores podem calcular la predicció \hat{X}_{j+1} en funció de X_1, \dots, X_j (la manera com es calcula aquesta predicció encara no l'hem vista), i aleshores els residus es defineixen per

$$e_j = \frac{X_j - \hat{X}_j}{r_j},$$

on r_j^2 és una estimació de la variància d'aquesta variable aleatòria. Llavors, si el soroll blanc és gaussià, les variables aleatòries e_1, \dots, e_n han de ser (aproximadament) independents amb distribució $\mathcal{N}(0,1)$. Per contrastar això, a part del gràfic d'aquests residus, pot calcular-se la funció d'autocorrelació mostral, i veure si correspon a la d'un soroll blanc. També hi ha diferents tests construïts per a aquesta situació, com per exemple el test de Portmanteau: En aquest test es contrasta la hipòtesi

$$H_0: \rho'_1 = \rho'_2 = \dots = \rho'_k = 0,$$

on ρ'_j és la funció d'autocorrelació del residu. L'estadístic de contrast, Q , té, sota H_0 , una distribució χ^2 amb d graus de llibertat, que depenen de p i q . Els programes d'ordinador acostumen a donar l'anomenat *valor p*, és a dir $P\{\chi^2(d) > q\}$, on q és el valor observat de Q .

Un comentari sobre els processos ARMA de mitjana no nul·la

Fins ara hem suposat que un procés ARMA tenia mitjana nul·la. Aquesta restricció és innecessària: direm que un procés estacionari $\{X_n, n \in \mathbb{Z}\}$ de mitjana $E[X_n] = \mu$ és un procés ARMA (p,q) si el procés centrat $Y_n = X_n - \mu$ és ARMA(p,q).

Els programes d'ordinador acostumen a donar una estimació de la mitjana, així com el valor d'un estadístic per contrastar que $\mu = 0$, segons veurem a continuació.

Una manera alternativa de tractar un procés ARMA (p,q) de mitjana no nul·la μ és considerar que compleix una equació

$$\Phi(B)X = \delta + \Theta(B)Z,$$

on ara δ és una funció de μ , Φ . Aquest terme δ s'anomena *la constant del model*. Exemple estem en un model AR(p) amb $\Phi(x) = 1 - \phi_1 x - \dots - \phi_p x^p$, aleshores es clar que

$$\delta = \mu(1 - \phi_1 - \dots - \phi_p).$$

Exemple 1: Considerem la simulació ($n = 1000$) del procés AR(1)

$$(I - 0.8B)X = Z$$

Suposant identificada l'estructura AR(1), el programa *STATGRAPHICS* dona les estimacions següents: entre parèntesis es posa una estimació de la desviació típica:

$\hat{\phi}_1 = 0.81328 \quad (0.01844),$ $\hat{\mu} = -1.5801 \quad (0.16).$

A continuació dona uns valors corresponents a un test d'hipòtesi

$$H_0 : \phi_1 = 0.$$

L'estadístic de contrast és $T_1 = \frac{\hat{\phi}_1}{\hat{\sigma}_1}$, que, sota H_0 , té una distribució t de Student (amb uns graus de llibertat a que depenen de n). En aquest cas, el valor observat de T és 44'10426. Aleshores

$$P \{-44'10426 < t(a) < 44'10426\} = 0'000\dots$$

i per tant rebutgem la hipòtesi nul·la (amb qualsevol nivell de significació).
Anàlogament tenim un test per a

$$H_0 : \mu = 0$$

En aquest cas s'observa $T_2 = -0'93$, i com que

$$P\{-0'93 < t(b) < 0'93\} = 0'34,$$

no tenim prou evidència en contra de H_0 (amb els nivells habituals), i, doncs, acceptem H_0 .

Una estimació de la variància del soroll blanc és $\hat{\sigma}^2 = 1'02356$.

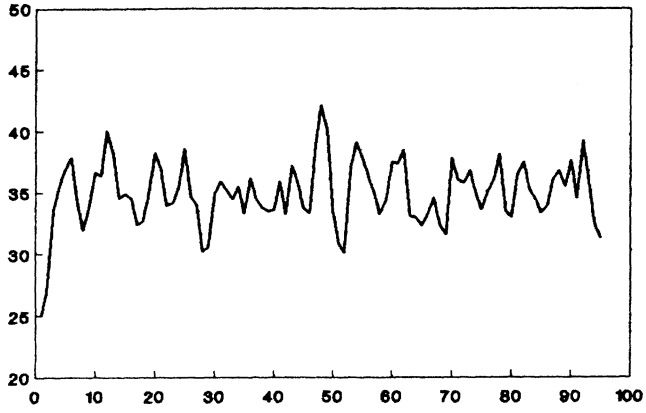
Finalment, per al test de Portmanteau, on es contrasta que les primeres 20 autocorrelacions dels residus són nul·les, dona de valor observat $q = 15'849$, i

$$P\{Q > 15'849\} = 0'603087,$$

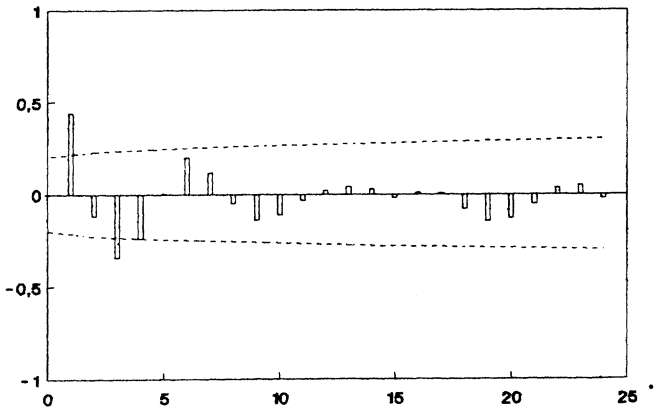
amb la qual cosa no tenim prou evidència en contra de H_0 .

Aquest estudi caldria completar-lo amb una anàlisi visual dels residus, i de la funció d'autocorrelació mostral d'aquests.

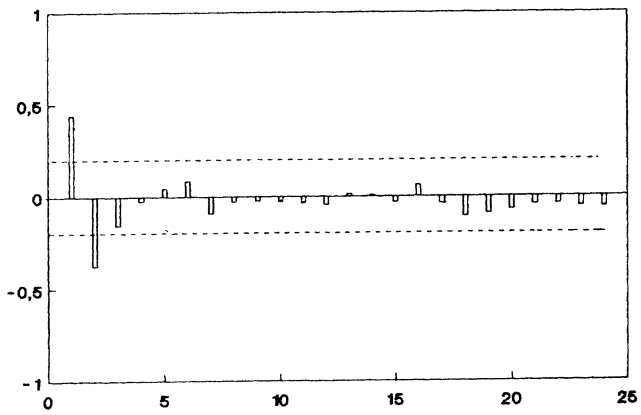
Exemple 2: El gràfic següent mostra les dades de viscositat d'un producte (dades de Bowerman i O'Connell) durant 95 dies. Es presenten també les funcions d'autocorrelació i autocorrelació parcial mostrals.



Sèrie observada



ACF mostral



ACF mostral

Aleshores es proposa un model AR(2), que s'estima

$\hat{\phi}_1 = 0'61$	(0'097)
$\hat{\phi}_2 = -0'37$	(0'098)
$\hat{\mu} = 34'91$	
$\hat{\sigma}^2 = 0'989$	

Tots els tests donen per bo l'ajustament d'aquest model.

7. Sèries temporals no estacionàries: models ARIMA

7.1 Dos exemples de diferenciació d'una sèrie temporal

a) Considerem una sèrie temporal del tipus següent:

$$X_n = a + bn + Z_n,$$

on $\{Z_n, n \in \mathbf{Z}\}$ és un soroll blanc. Òbviament no és estacionària ($E[X_n] = a + bn$). Podríem intentar estimar la tendència (per exemple, mitjançant una regressió) i eliminar-la. Però una altra opció és la següent: considerem l'operador de fer diferències:

$$\nabla X_n = X_n - X_{n-1} = (I - B)X_n,$$

i els seus iterats ∇^k . Aleshores

$$\nabla X_n = b + \nabla Z_n.$$

El procés ∇Z_n és estacionari (de fet és un soroll blanc) i, per tant, el procés $Y_n = \nabla X_n$ és estacionari (no centrat).

En general, qualsevol tendència polinòmica pot ser eliminada mitjançant diferències de l'ordre convenient.

Tot i que els dos procediments proposats semblin igual d'interessants, l'estimació d'una tendència determinista és problemàtica a la pràctica: amb un nombre finit d'observacions, és difícil assegurar com és la tendència.

b) Considerem una passejada aleatòria $\{X_n, n \in \mathbf{Z}\}$, definida a partir d'un soroll blanc $\{Z_n, n \in \mathbf{Z}\}$:

$$X_n = X_{n-1} + Z_n.$$

Aleshores

$$X_n - X_{n-1} = Z_n.$$

o equivalentment,

$$\nabla X_n = Z_n.$$

De nou el procés ∇X_n és estacionari.

7.2 Models ARIMA

Naturalment hi ha nombroses maneres com un procés pot ser no estacionari, però en moltes aplicacions intervenen processos que diferenciant-los un cert nombre de vegades proporcionen un procés estacionari. Aquesta és una de les idees importants de Box i Jenkins.

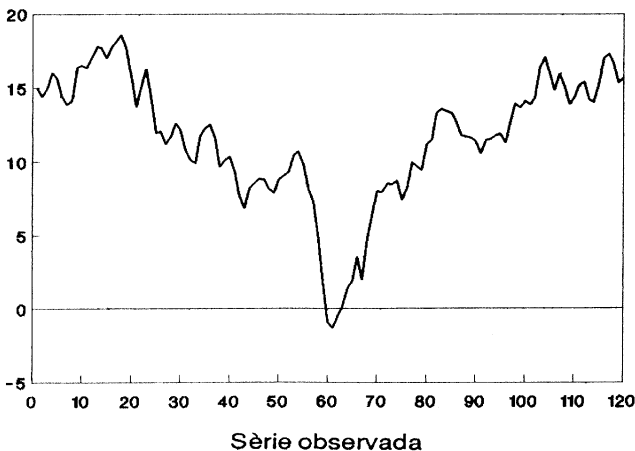
Direm que un procés $\{X_n, n \in \mathbf{Z}\}$ és ARIMA(p,d,q) si el procés $\nabla^d X$ és un procés ARMA(p,q). És a dir, X_n compleix l'equació

$$\Phi(B) (I - B)^d X_n = \Theta(B) Z_n. \quad (5)$$

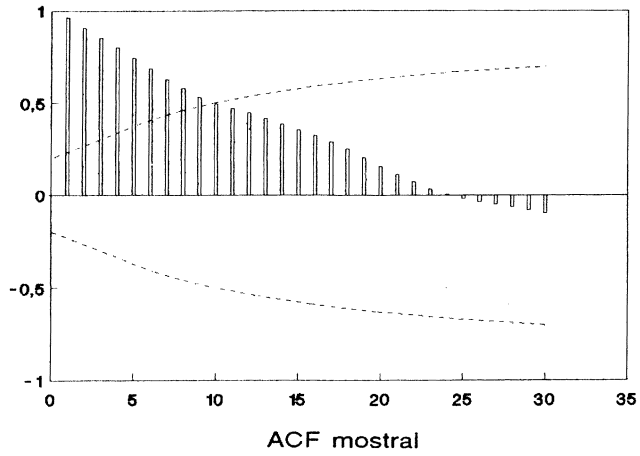
Cal observar que aquesta equació és com la d'un procés ARMA però ara es permet una arrel unitat en el polinomi de la part AR. També cal notar que, d'acord amb les consideracions anteriors, l'equació (5) deixa indeterminada una tendència polinomial d'ordre $d - 1$, i en particular la mitjana.

En aquests processos, com que no són estacionaris, la funció de covariància depèn de les variables. Però si, per exemple, considerem un procés ARIMA(1,1,0) com a límit d'un AR(2) però que té una arrel molt pròxima a 1, aleshores, d'acord amb els càlculs de la secció 3, tindrem que la funció d'autocorrelació tendeix molt lentament a 1. Aquesta és una característica dels processos ARIMA(p,d,q), amb $d \neq 0$: La «funció d'autocorrelació» tendeix lentament a zero, al contrari del que passava amb els processos Arma, en què la funció d'autocorrelació tendia a zero amb velocitat exponencial.

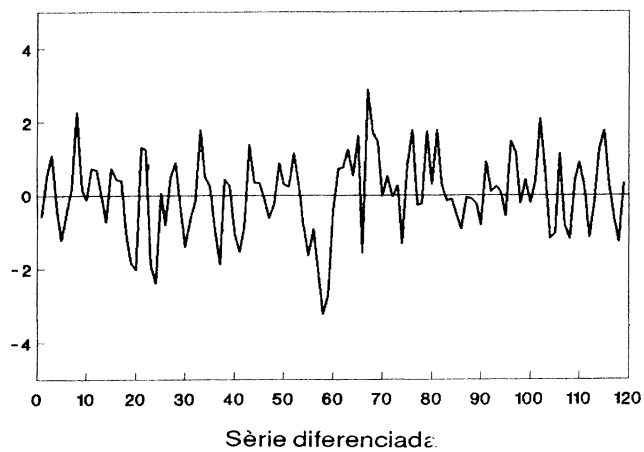
Exemple: El gràfic següent (de Bowerman i O'Connell) mostra les vendes (en unitats de 10.000 rotlles) del nombre de rotlles que passen de 100.000 venuts setmanalment durant 120 setmanes

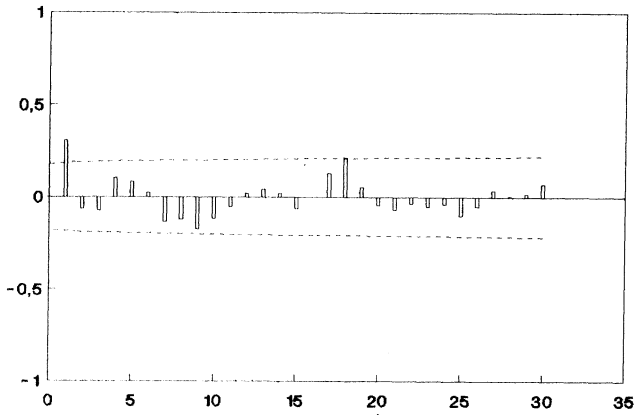


La funció d'autocorrelació mostral és

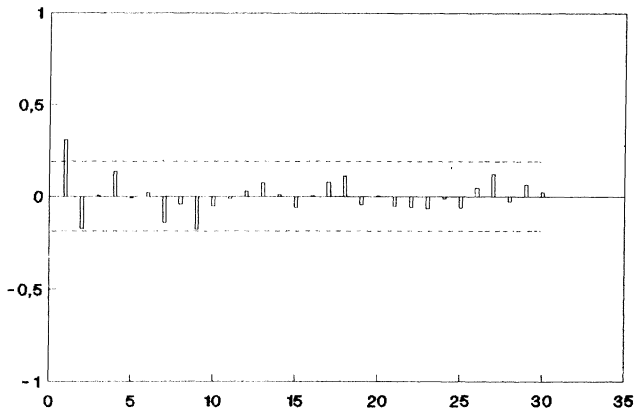


que posa de manifest el caràcter no estacionari d'aquesta sèrie temporal. Diferenciem una vegada, i obtenim





ACF mostral sèrie diferenciada



PACF mostral sèrie diferenciada

Aquesta situació és més difícil d'interpretar, i cal temptejar amb models ARIMA(1,1,0), ARIMA(0,1,1) i ARIMA(1,1,1). De tots aquests, el que dóna millors estimacions és l'ARIMA(0,1,1), que finalment dóna estimacions

$\hat{\theta}_1 = 0'35 \quad (0'088)$
$\hat{\mu} = 0'0048 \quad (0'12)$
$\hat{\sigma}_2 = 1'13$

i la mitjana no és significativament diferent de 0 (valor $p = 0'97$).

7.3 Models ARIMA estacionals

McCleary i Hay diuen que la teoria de les sèries temporals seria molt senzilla sense l'estacionalitat que presenten molts fenòmens. Considerem l'exemple del nombre de passatgers de línies aèries internacionals. La component estacionària és ben evident. Podríem intentar ajustar-hi (per regressió, per exemple) una sinusoidal, i després eliminar-la, tal com havíem dit al principi. Però també dèiem que no és del tot clar que sigui una bona tècnica (encara que sigui molt utilitzada). Una altra possibilitat és deixar més aleatorietat a aquesta component estacional. Aquesta també és una de les idees importants de Box i Jenkins. Suposem que tenim una sèrie temporal amb observacions mensuals, amb una forta component estacional (que s'observa a simple vista). Considerem les subsèries de cada més:

	Mes			
Any	1	2	...	12
1	X_1	X_2	...	X_{12}
2	X_{13}	X_{14}	...	X_{24}
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
n	$X_{1+12(n-1)}$	$X_{2+12(n-1)}$...	$X_{12+12(n-1)}$

Suposem que la sèrie de cada mes té una estructura ARMA(P,Q). Per exemple, la sèrie corresponent al primer mes complirà una equació

$$X_{1+12n} - a_1 X_{1+12(n-1)} + \dots + a_P X_{1+12(n-P)} = U_{1+12n} + c_1 U_{1+12n-1} + \dots + c_Q U_{1+12(n-Q)},$$

on $\{U_{1+12n}, n \in \mathbf{Z}\}$ és un soroll blanc. Però com que el mateix model pot ser aplicat a cada més, s'obté que la sèrie $\{X_n\}$ ha de complir l'equació

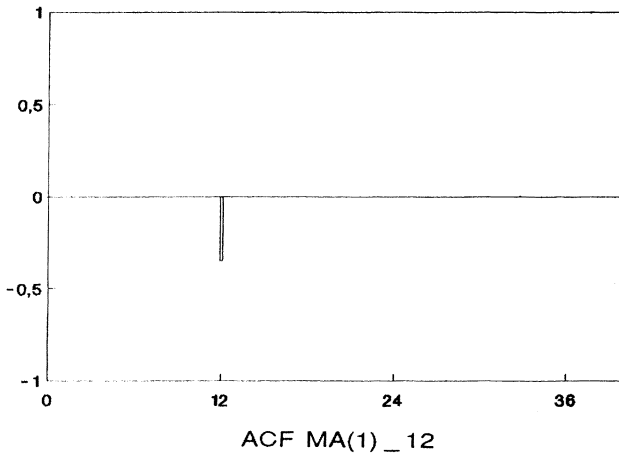
$$X_n - a_1 X_{n-12} + \dots + a_P X_{n-12P} = U_n + c_1 U_{n-12} + \dots + c_Q U_{n-12Q}.$$

que en forma compacta es pot escriure

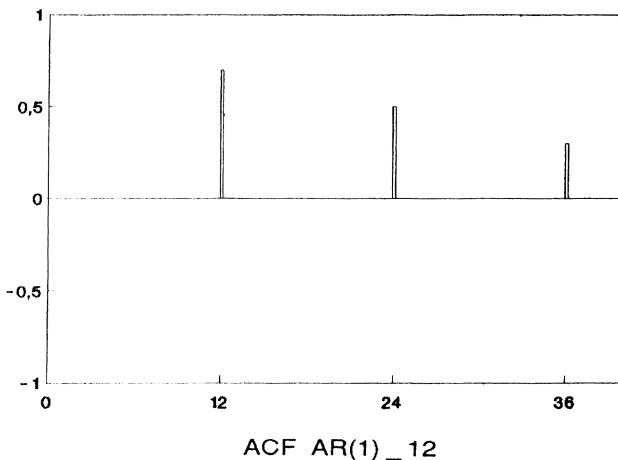
$$A(B^{12})X_n = C(B^{12})Z_n,$$

on $A(x) = 1 - a_1x - \dots - a_Px^P$ i $C(x) = 1 + c_1x + \dots + c_Qx^Q$. Però cal observar que no cal que $E[U_n U_{n+k}]$ sigui igual a zero a menys que k sigui un múltiple de 12.

Exemple 1: Considerem $P = 0$, $Q = 1$, $c_1 = -0.4$. Aleshores la sèrie de cada mes és MA(1). Si els sorolls blancs de cada mes estan incorrelacionats, les diferents columnes estan incorrelacionades.



Exemple 2: Suposem $P = 1$, $Q = 0$, $a_1 = 0.7$. Aleshores les sèries corresponents a cada mes són AR(1). Si els sorolls blancs estan incorrelacionats, la funció d'autocorrelació serà:



Però no és gaire raonable que els sorolls blancs dels diferents mesos estiguin in-correlacionats entre ells. Al contrari, el que es tracta és de suposar relació entre mesos consecutius. Per modelitzar aquesta dependència se suposa que $\{U_n, n \in \mathbf{Z}\}$ segueix un procés ARMA(p,q):

$$\Phi(B)U_n = \Theta(B)Z_n,$$

on $\{Z_n, n \in \mathbf{Z}\}$ és un soroll blanc.

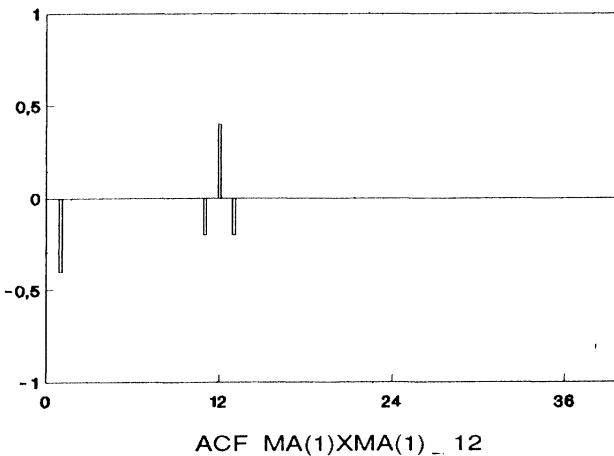
Combinant aquests dos models, i permetent diferenciacions ordinàries i estacionals, direm que un procés $\{Z_n, n \in \mathbf{Z}\}$ és un SARIMA(p,d,q) \times (P,D,Q)_s si compleix una equació del tipus

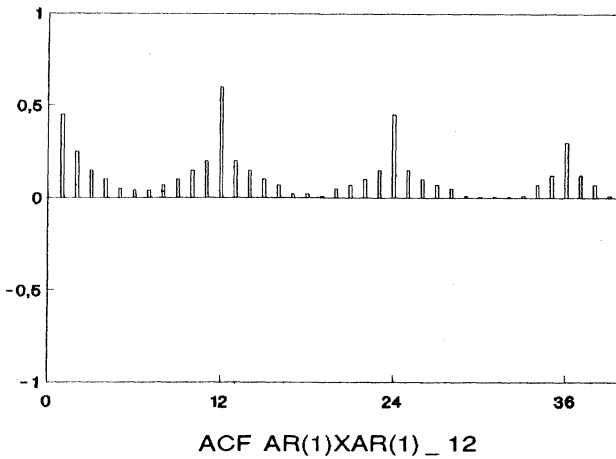
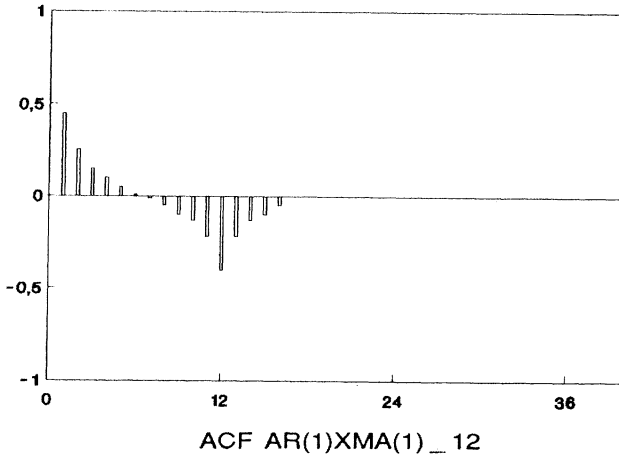
$$\Phi(B)A(B^s)(I-B)^d(I-B^s)^DX_n = \Theta(B)C(B^s)Z_n,$$

on $\Phi(x)$, $A(x)$, $\Theta(x)$ i $C(x)$ són polinomis de graus p , P , q i Q respectivament. Per simplificar l'escriptura s'introdueix l'operador de diferenciació estacional $\nabla_s = I - B^s$. Observeu que $\nabla_s \neq \nabla^s (= (I - B)^s)$.

Donada la interrelació entre la dependència anual i l'estacional, l'estructura de la ACF pot ser molt complicada. Com a línies generals pot dir-se que es comença buscant d i D per tal que la sèrie diferenciada $(I - B)^d (I - B^s)^DX_n$ tingui aparença estacionària. Després es miren les ACF i les PACF als múltiples de s per ajustar la part ARMA estacional. Finalment utilitzant l'ACF en $1, \dots, s - 1$ es tracta d'identificar la part ARMA no estacional.

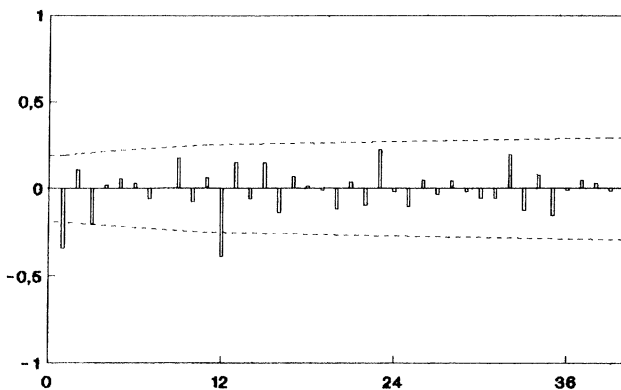
Exemples



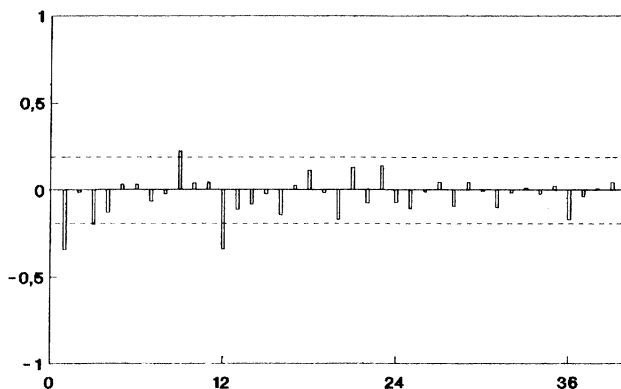


Una crítica que es fa sovint a aquest models és que la condició que tots els mesos tinguin la mateixa estructura ARMA, amb els mateixos coeficients és molt forta.

Exemple 3: Amb les dades del nombre de passatgers de línies aèries internacionals, si considerem la sèrie dels logaritmes (neperians) que designarem per $\{X_n\}$, primer la diferenciem amb un retard de 12 per treure-hi l'estacionalitat, i després fem una diferenciació ordinària per convertir-la en estacionària. La sèrie $\nabla_{12} X$ té les funcions següents d'autocorrelació i autocorrelació parcial mostrals:



Sèrie diferenciada
ACF mostral



Sèrie diferenciada
PACF mostral

Box i Jenkins suggereixen un model multiplicatiu SARIMA(0,1,1) x (0,1,1)₁₂, que estimen

$$(I - B)(I - B^{12})X = (I - 0'396B)(I - 0'614B^{12})Z,$$

on $\{Z_n\}$ és un soroll blanc de variància 0'00134. Brockwell i Davis proposen (per a $\nabla \nabla_{12} X$) un model MA(12), MA(23) o AR(12) amb molts coeficients nuls. Després d'estudiar tots aquests models (amb el criteri AIC d' Akaike) s'inclinen pel model

$$(I - B)(I - B^{12})X = (I - 0'432B - 0'306B^3 + 0'237B^5 - 0'656B^{12} + 0'352B^{23})Z,$$

on $\{Z_n\}$ és un soroll blanc de variància 0'00103. Utilitzant diferents criteris, aquest model s'ajusta millor que el de Box i Jenkins.

8. Previsió (forecasting)

Hem vist les tres etapes que s'anomenen de construcció de models:

- *identificació*
- *estimació*
- *diagnòstic*.

La quarta etapa, i molts autors diuen que la més important, consisteix a utilitzar el model per fer **previsions**, especialment a curt termini. Naturalment les previsions es formularan en termes probabilístics donant un interval de confiança. Per entendre bé la tècnica que utilitzarem ens cal recordar abans algunes nocions.

8.1 Esperança condicionada i prediccions lineals

Considerem dues variables aleatòries, X, Y , de les quals una és observable (posem que la Y) i d'acord amb l'observació de Y volem fer una predicció per a X . Per exemple, suposem que tirem tres cops una moneda, X designa el nombre de cares i Y , la diferència entre el nombre de cares i el de creus, en valor absolut. Si observem $Y = 1$, quina predicció farem per a X ? Y si $Y = 0$?, i si $Y = 3$? Evidentment la predicció, que escriurem \hat{X} , serà una funció de Y . Ara cal precisar en quin sentit volem que la predicció s'aproximi a X . El criteri habitual és el dels mínims quadrats: Cal que

$$E[(X - \hat{X})^2]$$

sigui el mínim possible. La solució a aquest problema es demostra que és prendre l'esperança de X condicionada per Y :

$$\hat{X} = E[X|Y].$$

Formulat en altres paraules: entre totes les funcions $f(Y)$, la que fa mínima

$$E[(X - f(Y))^2]$$

és prendre $f(Y) = E[X|Y]$.

Ara bé, calcular una esperança condicionada sovint no és senzill. Aleshores, en lloc de buscar la millor funció $f(Y)$ que minimitza la pèrdua quadràtica, es busca només la millor funció afí $f(Y) = aY + b$ que minimitza

$$E[(X - (aY + b))^2].$$

Com és ben conegut, aquest és el problema de la regressió lineal, i la solució consisteix a prendre a i b :

$$\hat{a} = \frac{Cov(X, Y)}{Var(Y)}, \quad \hat{b} = E[X] - \frac{Cov(X, Y)}{Var(Y)} E[Y].$$

Noteu que aquí només intervenen variàncies i covariàncies (que suposem que existeixen i són finites), i per tant lliga bé amb tota la teoria que anàvem fent de les sèries temporals.

Llavors

$$\hat{X} = \hat{a}Y + \hat{b}$$

s'anomena el *millor predictor lineal* de X com a funció de Y .

Cal notar que si el vector (X, Y) és normal bidimensional, aleshores aquesta expressió és precisament $E[X|Y]$. Com sempre, el cas gaussià té enormes avantatges.

Tot això es generalitza fàcilment al cas que tenim $n + 1$ variables aleatòries, X, Y_1, \dots, Y_n i observem Y_1, \dots, Y_n i volem una funció afí de les Y tal que

$$E[(X - (a_1 Y_1 + \dots + a_n Y_n + b))^2]$$

sigui mínim.

8.2 Predicció en processos ARMA

Les idees anteriors poden ser aplicades molt fàcilment a un procés estacionari: suposem que hem observat

$$X_1, \dots, X_T$$

i volem fer prediccions

$$\hat{X}_{T+1}, \dots, \hat{X}_{T+h}.$$

Utilitzarem el millor predictor lineal, que s'expressarà en termes de la funció d'autocorrelació, i a partir d'aquí hi ha diferents algorismes que calculen els coeficients de X_1, \dots, X_T per als predictors $\hat{X}_{T+1}, \dots, \hat{X}_{T+h}$ de forma recursiva (algorisme de Durbin-Watson, de les innovacions). La predicció \hat{X}_{T+1} s'anomena **previsió a un pas**, \hat{X}_{T+2} a **dos passos**, i així successivament.

Per a un procés ARMA(p,q) els càlculs són més senzills. Per exemple, per a un AR (el cas més senzill), posem de segon ordre,

$$X_n = \phi_1 X_{n-1} + \phi_2 X_{n-2} + Z_n,$$

la previsió a un pas dona,

$$\hat{X}_{n+1} = \phi_1 X_n + \phi_2 X_{n-1},$$

i la previsió a dos passos

$$\hat{X}_{n+2} = \phi_1 \hat{X}_{n+1} + \phi_2 X_n,$$

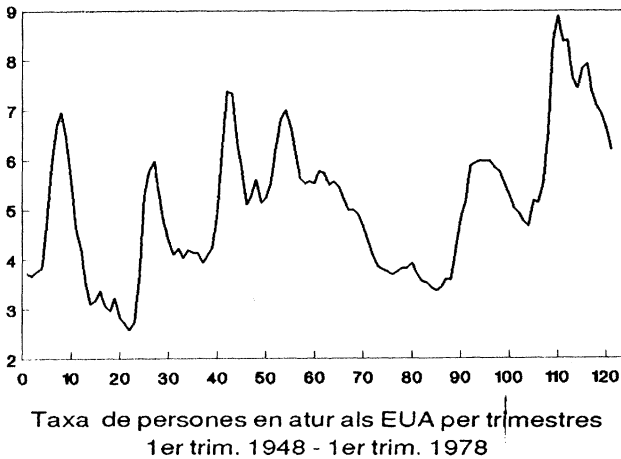
i la de tres passos

$$\hat{X}_{n+3} = \phi_1 \hat{X}_{n+2} + \phi_2 \hat{X}_{n+1}.$$

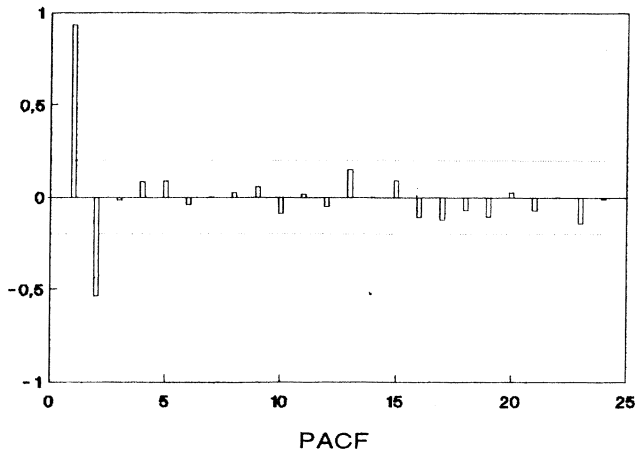
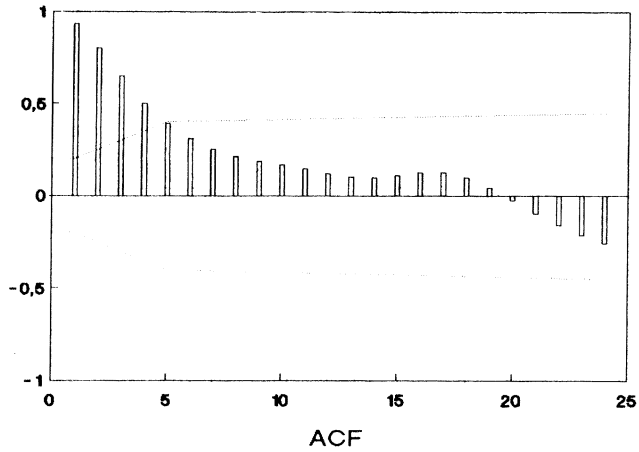
Tornant al cas general ARMA(p,q), si ara suposem que el soroll blanc és gaussià podrem construir intervals de confiança per a les previsions.

A la pràctica, els paràmetres $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ no són coneguts, però es procedeix igual utilitzant les estimacions.

Exemple 1: El següent gràfic mostra les taxes d'atur trimestrals als EUA des del primer trimestre de 1948 al primer trimestre de 1978, després de ser ajustat estacionalment (dades donades per J.D. Cryer)



Les funcions d'autocorrelació i autocorrelació parcial mostrada són les següents:



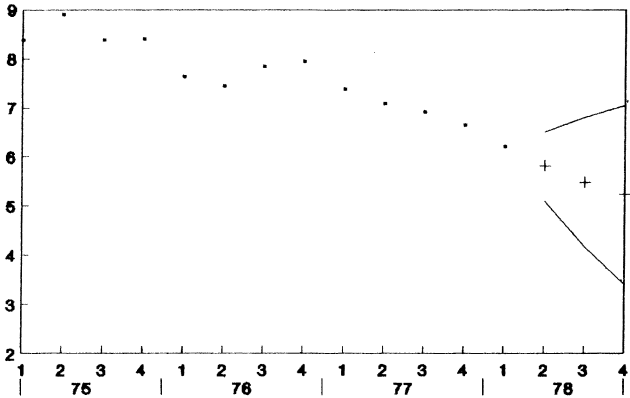
que permeten suposar una estructura AR(1). Les estimacions dels paràmetres donen

$$\begin{aligned}
 \hat{\phi}_1 &= 1'56 \quad (0'07) \\
 \hat{\phi}_2 &= -0'66 \quad (0'07) \\
 \hat{\mu} &= 5'08 \quad (0'34)
 \end{aligned}$$

i tots els tests donen per bo l'ajustament. Les previsions per a la resta de l'any 1978, utilitzant les 121 observacions anteriors són (entre parèntesi hi ha l'interval de confiança de nivell 0'95):

$$\begin{aligned} \hat{X}_{122} &= 5'81 \quad (5'1, 6'52) \\ \hat{X}_{123} &= 5'48 \quad (4'17, 6'8) \\ \hat{X}_{124} &= 5'23 \quad (3'41, 7'05). \end{aligned}$$

Una representació d'aquestes previsions és



Taxa de persones en atur als EUA (1948-1978)
Tres últims anys i previsió any

Observeu que les previsions s'aproximen a la mitjana.

8.3 Previsió en un procés ARIMA

Considerem un procés ARIMA(p,d,q) $\{X_n, n \in \mathbf{Z}\}$, és a dir, que el procés $Y = \nabla^d X$ és un ARMA(p,q). Si suposem que hem observat X_1, \dots, X_n , podem construir Y_{d+1}, \dots, Y_n , i podem calcular les previsions

$$\hat{Y}_{n+1}, \hat{Y}_{n+2}, \dots$$

(aquí la notació es fa una mica confusa ja que no utilitzem n observacions de Y). Ara l'equació

$$Y_m = \nabla^d X_m$$

es pot invertir i dóna

$$X_m = Y_m - \sum_{j=1}^d \binom{d}{j} (-1)^j X_{m-j},$$

en particular,

$$X_{n+1} = Y_{n+1} - \sum_{j=1}^d \binom{d}{j} (-1)^j X_{n+1-j}.$$

Com que X_{n+1-d}, \dots, X_n són observades (naturalment sempre se suposa n molt més gran que d), i podem estimar \hat{Y}_{n+1} amb les tècniques de l'apartat anterior, podem predir que

$$\hat{X}_{n+1} = \hat{Y}_{n+1} - \sum_{j=1}^d \binom{d}{j} (-1)^j X_{n+1-j}.$$

i per a la previsió a dues passes \hat{X}_{n+2} , substituir a la corresponent expressió X_{n+1} per \hat{X}_{n+1} .

Amb alguna hipòtesi addicional, aquest raonament pot fer-se completament rigorós i demostrar que \hat{X}_{n+1} és el millor predictor lineal.

Exemple 2: Considerem un soroll blanc $\{Z_n\}$. Donada la incorrelació, la predicció lineal serà

$$\hat{Z}_{n+1} = 0.$$

(El passat no aporta cap informació.)

Exemple 3: Considerem una passejada aleatòria

$$X_n = X_{n-1} + Z_n.$$

Sigui $Y_n = \nabla X_n = Z_n$ la sèrie diferenciada. Podem també escriure

$$X_{n+1} = Y_{n+1} + X_n,$$

i aleshores, observats X_1, \dots, X_n ,

$$\hat{X}_{n+1} = \hat{Y}_{n+1} + X_n = X_n,$$

d'acord amb l'exemple 2. Es a dir, en una passejada aleatòria, la millor previsió és l'última observació.

Anàlogament,

$$\hat{X}_{n+h} = X_n.$$

8.4 Previsions per suavitzament exponencial

En el món industrial i comercial es fan contínuament previsions partint de sèries temporals. Per exemple, a totes les empreses cada any (o cada mes) han de decidir i planificar la producció dels diferents productes d'acord amb les vendes dels darrers mesos o anys.

El procediment per fer previsions que hem vist era l'últim esglaó de la metodologia Box i Jenkins, la qual no és fàcil d'aprendre ni aplicar. Aleshores ha interessat desenvolupar tècniques senzilles però potents que permetin fer previsions fiables però de manera fàcil. Una d'aquestes tècniques (potser la més utilitzada) s'anomena **de suavitzament exponencial** (en anglès, *exponential weighted moving average*, *EWMA*). La idea és la següent: suposem que tenim observacions X_1, \dots, X_n . De cara a fer una previsió per a X_{n+1} , utilitzarem tota la informació disponible, però la importància d'aquesta disminuirà amb l'antiguitat:

$$\hat{X}_{n+1} = aX_n + a(1-a)X_{n-1} + a(1-a)^2X_{n-2} + \dots, \quad (6)$$

on $a \in (0, 1)$ és una constant que s'anomena factor de suavització, i s'escull d'acord amb la importància que es vol donar al passat immediat. (La notació no és estàndard; alguns llibres i programes utilitzen $1-a$ com a constant de suavització.) És a dir, es tracta d'estimar \hat{X}_{n+1} per una mitjana ponderada del passat, amb pesos que decreixen exponencialment.

De l'equació (6) es dedueix que

$$\hat{X}_{n+1} = aX_n + (1-a)\hat{X}_n, \quad (7)$$

o encara

$$\hat{X}_{n+1} = \hat{X}_n + a(X_n - \hat{X}_n) \quad (8)$$

és a dir, la previsió a l'etapa $n+1$ és la previsió a l'etapa n més un terme proporcional a l'error de previsió.

La fórmula (7) proporciona un mètode molt senzill per calcular i actualitzar les previsions: no cal guardar tota la sèrie temporal sinó només l'última previsió i l'observació actual. (Per inicialitzar l'algorisme es pren $\hat{X}_1 = X_1$.)

A la pràctica, s'utilitzen diferents versions millorades (i molt més complicades) d'aquesta idea, com les de Brown, Winter, Hold.

9. Anàlisi espectral

L'anàlisi espectral tracta d'estudiar la mesura espectral associada a la sèrie temporal (vegeu la secció 4). Aquesta mesura indica de quina manera la variabilitat de la sèrie es reparteix entre les diferents freqüències, i és un enfocament complementari

(més que alternatiu) a l'anàlisi en el domini del temps que hem fet fins ara. Podria ocórrer que la sèrie tingués una component periòdica, però que quedés difuminada entre tota la variació de la sèrie i que no es detectés estudiant la funció d'autocorrelació. Tot i que la funció d'autocorrelació i la mesura espectral són dues maneres d'expressar el mateix, la informació que trameten és distinta. Així, l'objectiu d'aquesta secció és construir estimadors de la densitat espectral (en cas que existeixi), estudiar-ne les propietats, fer tests d'hipòtesis i construir bandes de confiança per a la densitat espectral (l'anàleg d'un interval de confiança, però per a una funció). *A priori*, aquests problemes són, des del punt de vista estadístic, molt difícils de resoldre: com estimar una funció real de variable real utilitzant la informació d'un nombre finit d'observacions?

L'anàlisi espectral s'ha mostrat com una eina extremadament potent, i ha permès fer importants descobriments, com per exemple mostrar una forta evidència que les grans variacions climàtiques a la terra (tals com les èpoques glacials) estaven lligades a les variacions en l'òrbita de la terra (Hays et al., 1976).

9.1 El periodograma

Hem vist que donada una sèrie temporal estacionària $\{X_n, n \in \mathbf{Z}\}$, amb funció d'autocovariància $\gamma(k)$, podem representar

$$\gamma(k) = \int_{-\pi}^{\pi} e^{i\lambda k} G(d\lambda),$$

on G és la mesura espectral. Si G té densitat g , aleshores

$$\gamma(k) = \int_{-\pi}^{\pi} e^{i\lambda k} g(\lambda) d\lambda$$

i es pot invertir l'expressió i obtenir

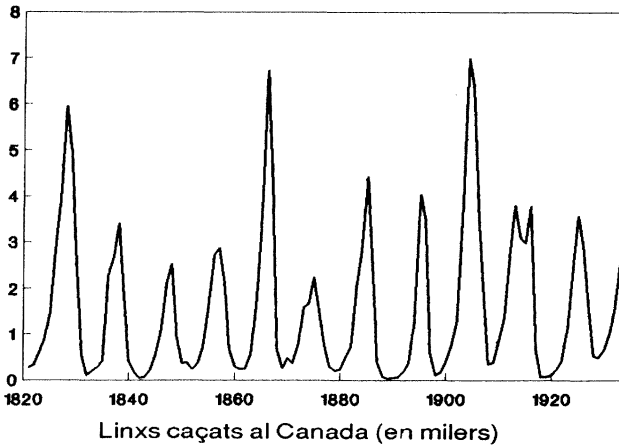
$$g(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \gamma(k) e^{-i\lambda k}, \quad \lambda \in [-\pi, \pi],$$

(que és simètrica respecte a l'eix d'ordenades perquè la sèrie és real). Això ens dona una idea per estimar la densitat espectral: donades les observacions X_1, \dots, X_n , es defineix el periodograma $I_n(\lambda)$ per

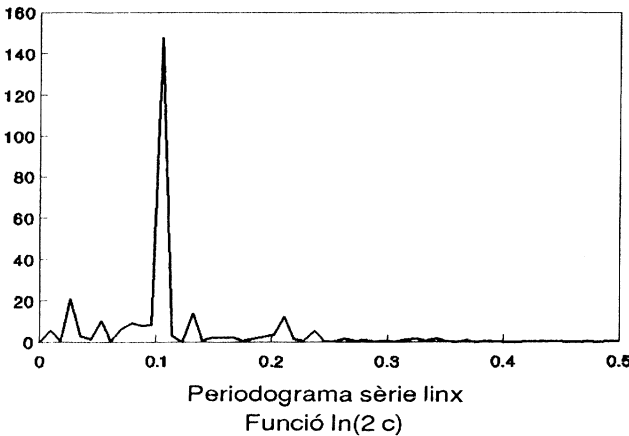
$$I_n(\lambda) = \frac{1}{2\pi} \sum_{k=-(n-1)}^{n-1} \hat{\gamma}(k) e^{-i\lambda k}, \quad \lambda \in [-\pi, \pi],$$

on $\hat{\gamma}(k)$ és la funció d'autocorrelació mostral (alguns autors ometen el factor 2π). El periodograma, de fet, es calcula només en els punts $\lambda = \frac{2\pi j}{n}$ que pertanyen a l'interval $[0, \pi]$, és a dir, $j = 0, \dots, [\frac{n}{2}]$. (Recordeu el fenomen *alias* que impedeix estimar freqüències superiors a π). Aquestes freqüències $\omega_j = \frac{2\pi j}{n} \in [0, \pi]$ s'anomenen *freqüències de Fourier*. Com més observacions tinguem, en més punts calcularem el periodograma. El càlcul directe del periodograma és força laboriós. Afortunadament es disposa d'un algorisme anomenat *transformada ràpida de Fourier (FFT)* que permet realitzar els càlculs de forma ràpida.

Com a exemple de periodograma estudiem el corresponent a la sèrie temporal del nombre de linxs caçats a un districte del Canadà de 1821 a 1934 (vegeu les dades a l'apèndix).



Aquesta és una sèrie temporal clàssica, força difícil d'analitzar, i el seu estudi admet distints enfocaments (vegeu Campbell i Walker, 1977, o Priestley, 1981). El gràfic següent és la representació del periodograma (mòdul d'un factor constant multiplicatiu). (Habitualment es representa $I_n(2\pi c)$, $c \in [0,0.5]$.)



Hi ha un pic a la freqüència de Fourier $\omega_{12} = 2\pi 12/114 = 0'66$, que correspon a un període $p = 2\pi/0'66$, que aproximadament és de 9'5 anys. També hi ha pics —molt més petits— en alguns harmònics d'aquesta freqüència, cosa que posa de manifest que l'estructura periòdica no és una sinusoide pura.

El principal problema amb el periodograma és l'alta variabilitat que presenta: sota bones hipòtesis, cada $I_n(\omega_j)$ és asimptòticament proporcional a una llei $\chi^2(2)$, i els valors en dues freqüències distintes són asimptòticament independents. Com veurem més endavant, les estimacions es milloren amitjanant els valors del periodograma en un entorn de cada freqüència (*suavització*).

9.2 Test per a la presència de periodicitats ocultes

La primera aplicació del periodograma consistirà a construir un test per contrastar la presència d'una periodicitat oculta a la sèrie temporal. Suposem que el nostre model és

$$X_n = \mu + A \cos(\lambda n) + B \sin(\lambda n) + Z_n;$$

on $\{Z_n\}$ és un soroll blanc gaussià. Volem contrastar

$$H_0 : A = B = 0$$

contra

$$H_1 : A \text{ o } B \text{ són no nuls.}$$

Si H_0 fos certa, el nostre procés seria un soroll blanc, i tindria una densitat espectral constant. Si tingués una component periòdica, parlant molt informalment, aquesta faria aparèixer un pic a la freqüència corresponent, o més exactament, faria aparèixer un salt a la funció de distribució espectral (es tractaria d'una sèrie amb espectre mixt). Aleshores la idea és rebutjar H_0 si alguna $I_n(\omega_j)$ és significativament més gran que les altres. Aquestes idees es poden estendre per contrastar l'existència d'una periodicitat de període p conegut o no. En aquesta darrera situació s'utilitza l'anomenat test de Fisher (Fisher, 1929), que consisteix a considerar l'estadístic

$$g = \frac{\max_{j=1, \dots, q} I(\omega_j)}{\sum_{j=1}^q I_n(\omega_j)}, \quad \text{on } q = \left\lfloor \frac{n-1}{2} \right\rfloor,$$

i mirar si g és significativament gran.

També, si la sèrie és un soroll blanc, és a dir, sota H_0 , es demostra que les variables aleatòries Y_1, \dots, Y_{q-1} definides per

$$Y_j = \frac{\sum_{k=1}^j I(\omega_k)}{\sum_{k=1}^q I_n(\omega_j)}$$

estan distribuïdes com si fossin una mostra ordenada d'una llei uniforme en (0,1). En altres paraules, si representem la funció de distribució empírica de Y_1, \dots, Y_{q-1} , (que correspon a l'histograma de freqüències relatives acumulades):

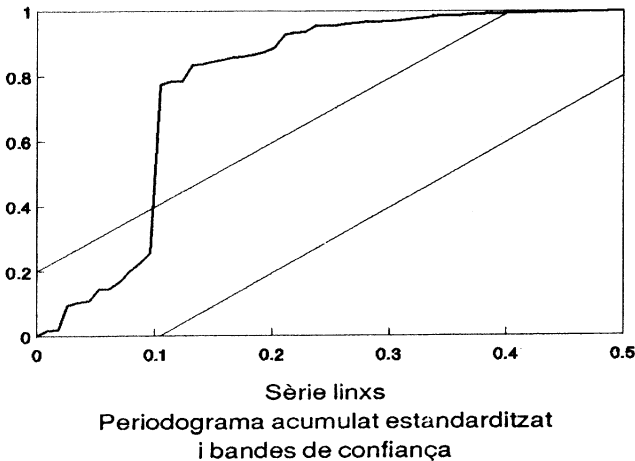
$$F_n(x) = \begin{cases} 0, & x < Y_1, \\ \frac{1}{q-1}, & Y_1 \leq x < Y_2, \\ \frac{2}{q-1}, & Y_2 \leq x < Y_3, \\ \vdots & \vdots \\ 1, & Y_{q-1} \leq x, \end{cases}$$

ha de ser aproximadament la funció de distribució d'una llei uniforme en (0,1):

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x \leq 1, \\ 1, & 1 < x. \end{cases}$$

A part de la impressió visual de la linealitat o no de la gràfica, s'utilitza un test de Kolmogorov-Smirnov, que dóna unes bandes de confiança per a cada nivell α . Si la funció F_n no està sempre dintre de les bandes, es rebutja H_0 . En general els programes d'ordinador dibuixen un gràfic equivalent a F_n que s'anomena *peridograma acumulant estandarditzat*.

El gràfic següent és el peridograma acumulat estandarditzat corresponent a la sèrie temporal dels linxs, i la banda de confiança del 95%.



Naturalment rebutgem la hipòtesi nul·la.

9.3 Suavitació del periodograma

Hem definit el periodograma per

$$I_n(\lambda) = \frac{1}{2\pi} \sum_{k=-(n-1)}^{n-1} \hat{g}(k) e^{-i\lambda k}.$$

Però quan k és pròxim a n , la funció d'autocovariància mostrada $\gamma(k)$ es calcula amb poques observacions, i és aleshores una estimació força grollera. Tot i que $I_n(\lambda)$ és un estimador asimptòticament sense biaix, és a dir, $E[I_n(\lambda)] \rightarrow g(\lambda)$ (g és la densitat espectral), el fet que $I_n(\lambda)$ tingui asimptòticament una llei proporcional a una χ^2 posa de manifest que $I_n(\lambda)$ no és un estimador consistent de la densitat espectral, és a dir, no convergeix (en probabilitat) a $g(\lambda)$, tal com seria de desitjar (la consistència és gairebé la mínima propietat exigible a un estimador). Però com indicàvem abans, això pot arreglar-se amitjanant el periodograma en punts pròxims a λ . L'entorn que s'agafa per amitjanar s'anomena *banda*. La manera més senzilla de fer-ho és elegint una amplada de banda (*bandwidth*) (i, doncs, el nombre de punts que amitjanarem). Si la llargada de la banda és ampla, la suavització és important, es redueix la variabilitat, però es perd informació. Si l'amplada és petita passa a l'inrevés. Es tracta, com quasi sempre, d'arribar a un punt d'equilibri.

De manera més general, per suavitzar el periodograma es considera una funció de pes $\{W(k), k = -m, \dots, m\}$, tal que

$$W(-k) = W(k) \quad \text{i} \quad \sum_{k=-m}^m W(k) = 1,$$

i definim el *periodograma suavitzat*

$$I_n^W(\omega_j) = \sum_{k=-m}^m W(k) I(\omega_{j+k}).$$

Si $\omega_{j+k} \notin [0, \pi]$, es considera la corresponent freqüència de Fourier de manera que I_n sigui periòdica de període 2π .

Per exemple, la situació d'amitjanar els punts pròxims a la qual ens referíem abans, suposem r punts a la dreta i r punts a l'esquerra, consistiria a prendre $m = r$ i una funció de pes

$$W(k) = \frac{1}{2m+1}, \quad k = -m, \dots, m.$$

Cal notar que el valor de m dependrà del nombre d'observacions n , i que en general

s'exigeixen, per obtenir bones propietats, més condicions sobre la funció de pes W per a $n \rightarrow \infty$

Una pràctica força habitual és aplicar successivament diverses suavitzacions. Per exemple, si primer apliquem

$$W_1(k) = \frac{1}{3}, \quad k = -1, 0, 1,$$

i després

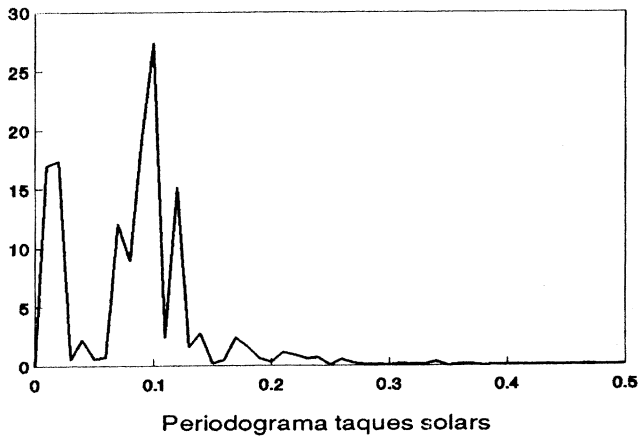
$$W_2(k) = \frac{1}{7}, \quad k = -3, -2, \dots, 3,$$

s'obindrà el mateix que aplicant

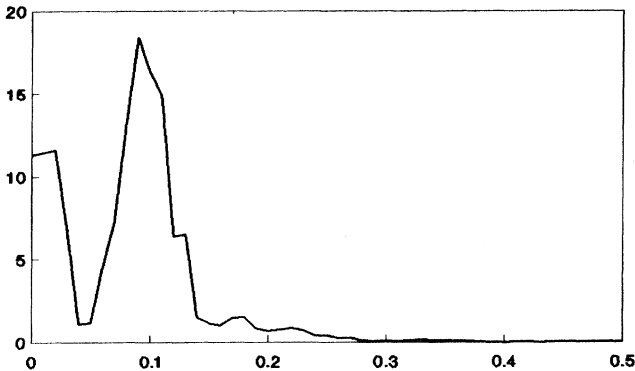
$$W_3(0) = W_3(1) = W_3(2) = \frac{3}{21}, \quad W_3(3) = \frac{2}{21}, \quad W_3(4) = \frac{1}{21}, \quad \text{i } W_3(-k) = W_3(k),$$

que, com es veu, dóna menys importància (menys pes) als punts més allunyats.

Per exemple, considerem el periodograma (mòdul d'un factor multiplicatiu) de la sèrie de les taques solars comentat a la introducció:

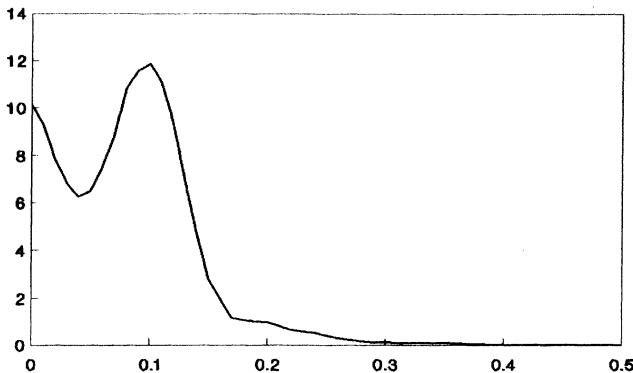


Una primera suavització realitzada amb la funció de pes W_1 anterior proporciona:



Peridograma de taques solars
1a suavització

Amb la funció de pes W_3 s'obté:



Peridograma de taques solars
2a suavització

Una altra manera de tractar el problema consisteix a posar els pesos en la definició del periodograma. Concretament, s'anomena *estimator amb finestra de retard* (*lag window estimator*)

$$I_n^L(\lambda) = \frac{1}{2\pi} \sum_{k=-h}^h \omega\left(\frac{k}{h}\right) \hat{g}(k) e^{-ik\lambda},$$

on ω és una funció (anomenada *finestra de retard*) contínua a trossos,

$$\omega(-x) = \omega(x), \quad \omega(0) = 1, \quad |\omega(x)| \leq 1, \quad \omega(x) = 0 \quad \text{si } |x| > 1.$$

Aleshores, prenent una bona finestra de retard es té $I_n^L(\lambda) \rightarrow g(\lambda)$ en L^2 (consistència en mitjana quadràtica).

Les dues idees, la dels pesos i la de les finestres, estan molt relacionades. Concretament, a partir d'una finestra de retard ω podem definir una funció de pes de tipus continu, anomenada *finestra espectral*

$$W(\lambda) = \frac{1}{2\pi} \sum_{k=-h}^h \omega\left(\frac{k}{h}\right) e^{-ik\lambda}, \quad \lambda \in [-\pi, \pi],$$

i llavors

$$I_n^L(\lambda) = \int_{-\pi}^{\pi} W(t) I_n(\lambda + t) dt,$$

és a dir, I_n^L és una suavització de tipus continu.

El nom de *finestra* és degut al fet que la funció $W(\lambda)$ que acabem de definir és pràcticament nul·la fora d'un petit interval $(\lambda - \varepsilon, \lambda + \varepsilon)$ (vegeu-ne els exemples), i llavors la integral que dóna la suavització és com veure $I_n(\lambda)$ a través d'una finestra que va de $\lambda - \varepsilon$ a $\lambda + \varepsilon$.

Exemples

1. La finestra de truncació

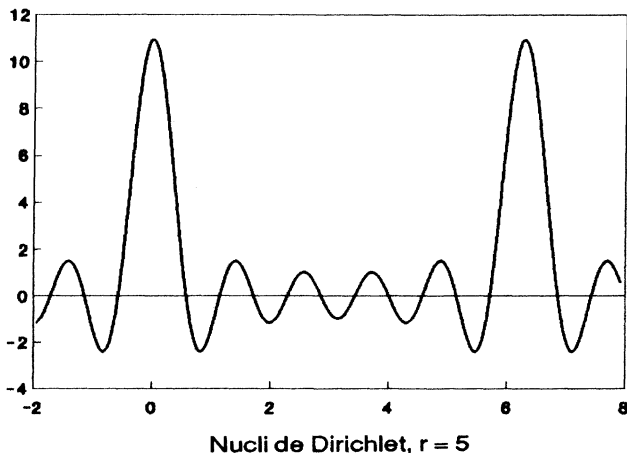
Es considera la funció rectangular (per aquest motiu també a vegades s'anomena *finestra rectangular*)

$$\omega(x) = \begin{cases} 1, & \text{si } |x| \leq 1, \\ 0, & \text{si } |x| > 1. \end{cases}$$

La corresponent finestra espectral és el nucli de Dirichlet D_h ,

$$W(\lambda) = \begin{cases} \frac{1}{2\pi} \frac{\sin((h + \frac{1}{2})\lambda)}{\sin(\frac{\lambda}{2})}, & \text{si } \lambda \neq 0, \\ 2n + 1, & \text{si } \lambda = 0. \end{cases}$$

Si escrivim $I^L(\lambda)$ veurem que consisteix a utilitzar només les autocorrelacions fins al retard *lag* h en la definició de periodograma (d'aquí el nom de *truncació*). El



valor h s'anomena *punt de truncació*. Posats, aleshores, a manipular la definició de periodograma, sembla ja més interessant fixar el punt de truncació, i donar uns pesos a les diferents autocorrelacions que hi fem intervenir. Això és el que fan les altres finestres.

2. La finestra triangular

La finestra de retard és la funció triangular

$$\omega(x) = \begin{cases} 1 - |x|, & \text{si } |x| \leq 1, \\ 0, & \text{si } |x| > 1, \end{cases}$$

i la finestra espectral és el nucli de Fejer F_h ,

$$W(\lambda) = \begin{cases} \frac{1}{2\pi r} \frac{\sin^2(\frac{h\lambda}{2})}{\sin^2(\frac{\lambda}{2})}, & \text{si } \lambda \neq 0, \\ \frac{n}{2\pi}, & \text{si } \lambda = 0. \end{cases}$$

3. La finestra de Daniell

En aquest cas es parteix d'una finestra espectral rectangular

$$W(\lambda) = \begin{cases} \frac{r}{2\pi}, & \text{si } |\lambda| \leq \frac{\pi}{r}, \\ 0, & \text{si } |\lambda| > \frac{\pi}{r}, \end{cases}$$

que correspon a la idea d'avitjar els periodogrames en un interval de longitud $2\pi/r$ al voltant de λ . La finestra de retard és

$$\omega(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x}, & \text{si } |x| \leq 1, \\ 0, & \text{si } |x| > 1. \end{cases}$$

4. La finestra de Tuckey

En aquest cas es considera la finestra de retard

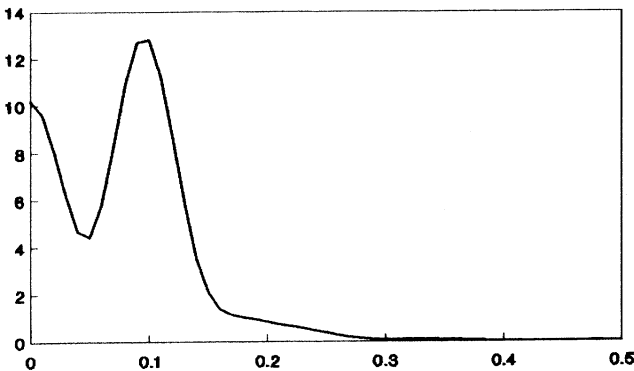
$$\omega(x) = \begin{cases} 1 - 2a + 2a \cos x, & \text{si } |x| \leq 1, \\ 0, & \text{si } |x| > 1. \end{cases}$$

I la corresponent finestra espectral és

$$W(\lambda) = aD_h(\lambda - \frac{\pi}{h}) + (1 - 2a)D_h(\lambda) + aD_h(\lambda + \frac{\pi}{h}),$$

on D_h és el nucli de Dirichlet. Habitualment es pren $a = 0'23$ o $a = 0'25$.

El gràfic següent dóna la suavització del periodograma de la sèrie de les taques solars (mòdul d'un factor multiplicatiu), realitzat amb la finestra de Tuckey ($a=0'23$), amb punt de truncació $h = 20$.



Periodograma de taques solars
Finestra de Tukey-Hanning, $r = 20$

Naturalment, hi ha nombrosos estudis sobre els avantatges d'una o d'altra finestra. Malgrat tot, el factor realment crític és l'elecció del punt de truncació h .

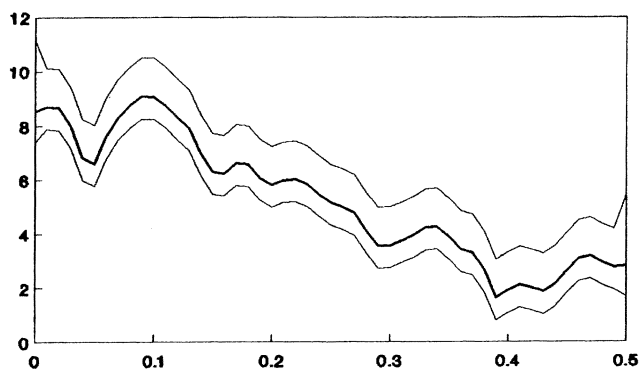
9.4 Bandes de confiança per a la densitat espectral

El coneixement de la distribució asimptòtica de les components del periodograma permet construir bandes de confiança per a la densitat espectral, això és, donat un nivell de confiança δ (normalment 0'95 o 0'99), es defineixen, mitjançant una funció de pes o una finestra, dues funcions $a(\lambda)$, $b(\lambda)$ tals que

$$a(\omega_j) \leq g(\omega_j) \leq b(\omega_j), \quad \forall_j = 0, \dots, \left[\frac{n}{2} \right],$$

amb probabilitat $\geq \delta$. Habitualment els programes d'ordinador donen les bandes de confiança per a $\ln g(\lambda)$.

El gràfic següent consisteix en el periodograma suavitzat per a la sèrie de les taques solars i les bandes del 95% (en logaritmes neperians).



Periodograma suavitzat i bandes de confiança
(Log. neperians)

Apèndix

Conjunts de dades

Passatgers de Línies Aèries Internacionals (1949-1960) (en milers)

	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960
Gener	112	115	145	172	197	205	243	285	316	341	361	418
Febrer	118	126	151	181	197	189	234	278	302	319	343	392
Març	132	141	179	194	237	236	268	318	357	363	407	420
Abril	129	135	164	182	236	228	270	314	349	349	397	462
Maig	121	125	173	184	230	235	271	319	356	364	421	473
Juny	135	150	179	219	244	265	316	375	423	436	473	536
Juliol	148	171	200	231	265	303	365	414	466	492	549	623
Agost	148	171	200	243	273	294	348	406	468	506	560	607
Setembre	136	159	185	210	238	260	313	356	405	405	464	509
Octubre	119	133	163	192	212	230	275	307	348	360	408	462
Novembre	104	114	146	173	181	204	238	272	306	311	363	391
Desembre	118	140	167	195	202	230	279	307	337	338	406	433

**Taxa de persones en atur als EUA per trimestres
1r trimestre 1948-1r trimestre 1978
després de ser ajustada estacionalment
(per columnes)**

3,73	2,73	5,83	5,00	4,17	8,37
3,67	3,70	5,10	5,00	4,80	8,40
3,77	5,27	5,27	4,90	5,17	7,63
3,83	5,80	5,60	4,67	5,87	7,43
4,67	5,97	5,13	4,37	5,93	7,83
5,87	5,33	5,23	4,10	5,97	7,93
6,70	4,73	5,53	3,87	5,97	7,37
6,97	4,40	6,27	3,80	5,97	7,07
6,40	4,10	6,80	3,77	5,83	6,90
5,57	4,23	7,00	3,70	5,77	6,63
4,63	4,03	6,67	3,77	5,53	6,20
4,23	4,20	6,20	3,83	5,27	
3,50	4,13	5,63	3,83	5,03	
3,10	4,13	5,53	3,93	4,93	
3,17	3,93	5,57	3,73	4,77	
3,37	4,10	5,53	3,57	4,67	
3,07	4,23	5,77	3,53	5,17	
2,97	4,93	5,73	3,43	5,13	
3,23	6,30	5,50	3,37	5,50	
2,83	7,37	5,57	3,43	6,57	
2,70	7,33	5,47	3,60	8,37	
2,57	6,37	5,20	3,60	8,90'	

Dades de viscositat d'un producte (durant 95 dies)

1	25,000	21	36,893	41	35,922	61	37,334	81	36,512
2	27,000	22	33,894	42	33,213	62	38,468	82	37,424
3	33,514	23	34,171	43	37,167	63	33,098	83	35,155
4	35,496	24	35,427	44	35,814	64	32,929	84	34,480
5	36,903	25	38,583	45	33,685	65	32,275	85	33,290
6	37,836	26	34,618	46	33,276	66	33,231	86	33,925
7	34,265	27	33,974	47	38,816	67	34,579	87	36,104
8	31,898	28	30,207	48	42,084	68	32,345	88	36,735
9	33,757	29	30,543	49	40,007	69	31,532	89	35,458
10	36,630	30	34,869	50	33,451	70	37,804	90	37,592
11	36,352	31	35,889	51	30,841	71	36,054	91	34,490
12	40,076	32	35,204	52	30,066	72	35,730	92	39,169
13	38,093	33	34,434	53	37,054	73	36,799	93	35,824
14	34,541	34	35,484	54	39,098	74	34,950	94	32,388
15	34,857	35	33,238	55	37,908	75	33,525	95	31,285
16	34,532	36	36,168	56	36,239	76	35,101		
17	32,385	37	34,412	57	34,954	77	35,977		
18	32,606	38	33,767	58	33,206	78	38,098		
19	34,891	39	33,425	59	34,426	79	33,460		
20	38,242	40	33,572	60	37,451	80	32,928		

**Vendes setmanals que passen de 100.000 rotlles
d'una fàbrica de paper durant 120 setmanes
(unitats de 10.000 rotlles)**

1	1500	26	1203	51	908	76	819	101	1389
2	1441	27	1121	52	932	77	997	102	1429
3	1494	28	1170	53	1047	78	969	103	1639
4	1604	29	1260	54	1069	79	945	104	1709
5	1563	30	1220	55	984	80	1121	105	1589
6	1440	31	1078	56	818	81	1150	106	1482
7	1390	32	1011	57	725	82	1328	107	1595
8	1408	33	993	58	508	83	1359	108	1510
9	1637	34	1174	59	183	84	1343	109	1388
10	1653	35	1226	60	-91	85	1331	110	1427
11	1638	36	1250	61	-132	86	1274	111	1517
12	1710	37	1154	62	-60	87	1180	112	1538
13	1779	38	966	63	14	88	1173	113	1419
14	1774	39	1010	64	140	89	1162	114	1401
15	1700	40	1035	65	193	90	1137	115	1525
16	1775	41	928	66	356	91	1055	116	1702
17	1819	42	772	67	196	92	1147	117	1729
18	1860	43	683	68	485	93	1156	118	1664
19	1759	44	820	69	655	94	1180	119	1534
20	1574	45	853	70	801	95	1189	120	1565
21	1370	46	887	71	797	96	1130		
22	1501	47	879	72	850	97	1278		
23	1626	48	816	73	845	98	1394		
24	1435	49	791	74	871	99	1369		
25	1196	50	880	75	738	100	1411		

Linxs caçats a Mackenzie River Distric, Canadà (1821-1934)

Any	Linxs	Any	Linxs	Any	Linxs	Any	Linxs	Any	Linxs
1821	269	1845	546	1869	255	1893	377	1917	674
1822	321	1846	1033	1870	473	1894	1292	1918	81
1823	585	1847	2129	1871	358	1895	4031	1919	80
1824	871	1848	2536	1872	784	4896	3495	1920	108
1825	1475	1849	957	1873	1594	1897	587	1921	229
1826	2821	1850	361	1874	1676	1898	105	1922	399
1827	3928	1851	377	1875	2251	1899	153	1923	1132
1828	5943	1852	225	1876	1426	1900	387	1924	2432
1829	4950	1853	360	1877	756	1901	758	1925	3574
1830	2577	1854	731	1878	299	1902	1307	1926	2935
1831	523	1855	1638	1879	201	1903	3465	1927	1537
1832	98	1856	2725	1880	229	1904	6991	1928	529
1833	184	1857	2871	1881	469	1905	6313	1929	485
1834	279	1858	2119	1882	736	1906	3794	1930	662
1835	409	1859	684	1883	2042	1907	1836	1931	1000
1836	2285	1860	299	1884	2811	1908	345	1932	1590
1837	2685	1861	236	1885	4431	1909	382	1933	2657
1838	3409	1862	245	1886	2511	1910	808	1934	3396
1839	1824	1863	552	1887	389	1911	1388		
1840	409	1864	1623	1888	73	1912	2713		
1841	151	1865	3311	1889	39	1913	3800		
1842	45	1866	6721	1890	49	1914	3091		
1843	68	1867	4254	1891	59	1915	2985		

Bibliografia

- Azencott, R.; Dacunha-Castelle, D., *Séries d'observations irrégulières*, Masson, París, 1984.
- Bloomfield, P., *Fourier Analysis of Time Series: An Introduction*, Wiley, Nova York, 1976.
- Bowerman, B.L.; O'Connell, *Time Series Forecasting*, Duxbury Press, Boston, 1979.
- Box, G.E.P.; Jenkins, G.M., *Time Series Analysis: Forecasting and Control*, Holden-Day, Oakland, Califòrnia, 1976.
- Brockwell, P.J.; Davis, R.A., *Time Series: Theory and Methods*, Springer, Nova York, 1987.

- Campbell, M.J.; Walker, A.M., A Survey of Statistical Work on the Mackenzie River Series of Annual Canadian Lynx Trappings for the Years 1821-1934 and a New Analysis, *J.R. Statist. Soc. A* (1977), 140, Part A, pàg. 411-431.
- Cryer, J.D., *Time Series Analysis*, Duxbury Press, Boston, 1986.
- Dixon, W.J., (Edit.) *BMDP Statistical Software Manual*.
- Gourieroux, C.; Montfort, A., *Séries temporelles et modèles dynamiques*, Economica, París, 1990.
- Granger, C.W.J.; Newbold, P., *Forecasting Economic Time Series* 2a ed., Academic Press, Orlando, 1986.
- Harvey, A.C., *Time Series Models*, Ph. Allan Publ. Lted, Oxford, 1981.
- Hays, J.D. et al., Variations in the Earth's Orbit: Pacemaker of the Ice Ages, *Science* (1976), Vol. 194, Núm. 4270, pàg. 1121-1132.
- Koopmans, L.H., *The Spectral Analysis of Time Series*, Academic Press, Nova York, 1974.
- McCleary, R.; Hay, R.A., *Applied Time Series Analysis*, Sage, Nova York, 1980.
- McCullagh, P.; Nelder, J.A., *Generalized Linear Models*, 2a ed., Chapman & Hall, Londres, 1989.
- Moore, D.S., *Statistics, Concepts and Controversies*, 2a ed. W.H. Freeman and Co., Nova York, 1985.
- Morris, M.J., Forecasting the Sunspot Cycle, *J.R. Statist. Soc. A* (1977), 140, Part 4, pàg. 437-468.
- Pankratz, A., *Forecasting with Univariate Box-Jenkins Models*, Wiley, Nova York, 1983.
- Peña, D., *Estadística: Modelos y métodos. 2. Modelos lineales y series temporales*, Alianza Editorial, Madrid, 1987.
- Priestley, M.B., *Spectral Analysis and Time Series*, Academic Press, Londres, 1981.
- Uriel, E., *Análisis de series temporales. Modelos ARIMA*, Paraninfo, Madrid, 1985.