



Stockholms
universitet

Informationsfönster som kvalitetsmått på prediktioner

Sandra Malkey

Kandidatuppsats 2013:10
Matematisk statistik
Oktober 2013

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Informationsfönster som kvalitetsmått på prediktioner

Sandra Malkey*

Oktober 2013

Sammanfattning

I detta arbete presenteras Öller och Teterukovskys metod för att kvantifiera kvaliteten på en stationär makroekonomisk statistisk tidsserie, där åtgärderna grundar sig på en kombination av hur förutsägbar serien är och hur mycket statistikan behöver revideras. Öller och Teterukovsky presenterade (2007) diagramstypen informationsfönster som är baserad på prediktionsindex och som ger en illustration av prediktionernas kvalitet. Detta möjliggör att man kan testa hurvida informationsrika prediktioner är. En generell analys av informationsfönster görs, i syfte om att undersöka informationsfönstrets tillförlitlighet genom att simulera data från autoregressiv- och moving average processer. Inledningsvis simuleras 10000 oberoende tidsserier där det sanna värdet är tidpunkten 2000 på varje tidsserie och 5 prediktioner görs av det sanna värdet. Vi ser då att informationsfönstret ger en god illustration av prediktionernas kvalitet. Avslutningsvis görs ett mer verklighetsbaserat exempel där vi simulerar en tidsserie och gör 5 prediktioner av 10 olika tidpunkter i tidsserien. Vi ser då att Öller och Teterukovskys teori inte fungerar och att informationsfönstret inte är tillförlitligt.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post:sandramalkey@hotmail.com . Handledare: Anders Björkström.

Abstract

In this presentation Öller och Teterukovskys method is presented to quantify the quality of a stationary macroeconomical statistic timeseries. Their actions are based on a combination of how predictable the series is and how much the statistics need to be revised. Öller and Teterukovsky presented (2007) the graphtype information window that is based on predictionindex and that gives an illustration of the quality of the predictions. This makes it possible to test how rich of information the predictions are. A general analysis of information windows are made, in order to examine the reliability of the information window, by simulating data from autoregressive- and moving average processes. We begin with simulating 10000 independent time series where the true value is the date of 2000 on each time serie and five predictions are made of the true value. Then we see that the information window provides a good illustration of the quality of the predictions. Finally we made a more reality-based example where we simulate one time serie and make 5 predictions of 10 different points in the time serie. We see that Öller and Teterukovskys theory does not work and that the information window is not reliable.

Förord

Detta är en kandidatuppsats på 15hp som leder till en kandidatexamen i matematisk statistik på Stockholms universitet.

Jag vill rikta ett stort tack till min handledare Anders Björkström på matematiska institutionen på Stockholms universitet för alla värdefulla råd och vägledning under arbetets gång.

Innehåll

1 Inledning	6
1.1 Beskrivning av Öller och Teterukovskys diagram	7
1.2 Centrala begrepp	10
1.2.1 Stationäritet	10
1.2.2 Prediktion	10
2 Beskrivning av data och prediktionsmetoder	10
2.1 Data	10
2.2 Undersökta prediktionsmetoder	11
3 Oberoende replikat	12
3.1 Autoregressiv process	12
3.1.1 Autoregressiv process av grad 1	13
3.1.2 Härledning av prediktionsformler för Autoregressiva processen av grad 1	13
3.1.3 Autoregressiv process av grad 2	18
3.1.4 Härledning av prediktionsformler för Autoregressiva processen av grad 2	20
3.2 Moving average process	26
3.2.1 Moving average process av grad 1	26
3.2.2 Härledning av prediktionsformler för Moving average av grad 1	26
3.2.3 Moving average process av grad 2	32
3.2.4 Härledning av prediktionsformler för Moving average processen av grad 2	32
4 Utan replikat	36
5 Diskussion/Slutsats	40
6 Referenser	43

1 Inledning

För att ett samhälle ska fungera är det både av intresse och stor vikt att försöka förutse framtiden. Ett av de mest klassiska tillvägagångssätten för att förutse framtiden är genom att skapa prognoser. Med hjälp av prognoser kan en fingervisning ges avseende svar på de frågor som önskas besvaras. I många sammanhang bearbetas en prognos fram utifrån nödvändig fakta och ofta i kombination med en matematisk modell som grund.

Ett ekonomiskt område där viktiga prognoser görs är inom makroekonomi där exempelvis inflation, konjunktursvängningar, arbetslöshet och tillväxt ingår. Makroekonomi skulle bli mer användbar om det sker en förbättring kring modeller som säger hur ekonomin fungerar, vilket troligtvis skulle leda till bättre prognoser. Detta har varit ekonomernas dröm ända sedan 1800-talet. Men även i modern tid har de flesta experter varit skeptiska till att det skulle finnas ett optimalt sätt att undersöka hur bra statistiska tidsserier, ger oss det vi förväntar oss av dem. (Öller och Teterukovsky 2007)

Förutom att prognoser alltid är osäkra så är en ytterligare komplikation att det sanna värdet på en variabel X_t inte blir känt så fort tidpunkten t passerar. Det krävs tid för inrapportering, efterkontroller med mera. Vissa prognoser av X_t är alltså inte förutsägelser om framtiden utan snarare revisioner i efterhand. Från matematisk-statistisk synpunkt sett spelar skillnaden mellan prognos och revision ingen roll. Båda typerna kan betraktas som prediktioner av den okända variabeln X_t baserat på mer och mer information.

Öller och Teterukovsky presenterade (2007) diagramstypen informationsfönster i samband med prognos av privat och offentlig konsumtion, där metoder för att kvantifiera kvaliteten på makroekonomiska statistiska tidsserier presenteras. Åtgärderna grundar sig på en kombination av hur förutsägbar serien är och hur mycket statistiken behöver revideras. I artikeln introduceras ett slags diagram, kallat informationsfönster som är baserat på signal-brus kvot, vars avsikt är att ge en illustration av prognosernas och revisionernas kvalitet.

Kvaliten på makroekonomiska data uttrycks vanligen genom att vi i ord beskriver datas innehåll, noggrannhet, aktualitet, jämförbarhet, tillgänglighet och tydlighet. Öller och Teterukovsky gör ett försök till att uttrycka kvaliteten i termer av hur tidigt och hur exakt värdet av en makroekonomisk variabel kan bedömas med en viss säkerhet.

Öller och Teterukovsky anser att antalet och omfattningen revideringar återspeglar tillförlitligheten hos en variabel. Om dess värde måste förändras från grunden upp till 2 år efteråt, hur trovärdig är den första (och vik-

tigaste) publicerade siffran och hur relevant är den slutgiltiga siffran? En studie om revideringar av svenska Nationalräkneskaper, utgifter och variabler, Öller och Hansson (2004), avslöjar att revideringar har en betydelse. De utgår ifrån förhållandet mellan prognosens fel/revideringar och variationen av tidsserierna. Prognosfelet och variationen av tidsserier, lider av det faktum att det korrekta resultatet är okänt. Den slutgiltiga siffran används för att representera det värdet. (Öller och Teterukovsky 2007)

1.1 Beskrivning av Öller och Teterukovskys diagram

Betrakta en stationär tidsserie X_t , där t genomlöper alla heltal från $-\infty$ till ∞ . För vissa t -värden, säg $t = 1, \dots, T$, har vi tillgång till prediktioner gjorda med olika lång framförhållning. Vi har alltså en matris av data, y_{lt} , $l = 1, \dots, L$, $t = 1, \dots, T$:

$$\begin{pmatrix} y_{11} & y_{21} & \cdots & y_{l1} & \cdots & y_{L1} \\ y_{12} & y_{22} & \cdots & y_{l2} & \cdots & y_{L2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{1t} & y_{2t} & \cdots & y_{lt} & \cdots & y_{Lt} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ y_{1T} & y_{2T} & \cdots & y_{lT} & \cdots & y_{LT} \end{pmatrix}$$

där y_{lt} är en prediktion av X_t gjord vid tiden $t - \Delta_l$, där talen Δ_l , $l = 1, \dots, L$ är en avtagande följd av "framförhållningar". Kvaliteten på prognoser med framförhållningen Δ_l kan illustreras genom medelkvadratfelet

$$MSE_l = \frac{1}{T} \sum_{t=1}^T (y_{lt} - y_{Lt})^2, l = 1, \dots, L - 1. \quad (1)$$

Värdena på MSE_l standardiseras via variansen av den slutgiltiga siffran, det vill säga genom division med s_L^2 ,

$$\overline{MSE}_l = \frac{MSE_l}{s_L^2} \quad (2)$$

där

$$s_L^2 = \frac{1}{T-1} \sum_{t=1}^T (y_{Lt} - \bar{y}_L)^2 \quad (3)$$

Skattningen s_L^2 antas vara nära den sanna men ändå okända variansen av X_t . En stationär process tillåter oss standardisera med s_L^2 , för att tidsserien har en konstant varians vilket då gör det meningsfullt att beräkna s_L^2 . Då

gäller det även att X_t måste vara oberoende.

Låt oss beteckna prediktionsindex med I_l som beräknas

$$I_l = 1 - \frac{MSE_l}{s_L^2} \text{ där } l = 1, \dots, L. \quad (4)$$

Beteckningarna i (4) är tagna från Öller och Teterukovsky (2007). Vi kommer i det följande ofta att tala om en prediktion som avser processen k steg framåt i tiden. Motsvarande prediktionsindex blir med deras beteckningar I_{L-k} där L har samma betydelse som ovan.

Öller och Teterukovsky definerade ovannämnda, I_l , till signal-brus-kvot (SNR), men vi väljer att använda oss av benämningen prediktionsindex istället, då det finns en del andra "index" i statistiken som är definerade på likartat sätt. Denna statistiska, I_l , estimerar hur mycket av osäkerheten, dvs variansen av variabeln som har eliminerats genom prediktionen y_{lt} . Prediktionsfel och revideringar kan, för vissa prediktioner, vara så stora att de överstiger variansen hos den variabeln, vilket gör det motsvarande värdet av I_l , mindre än noll. Detta skulle vara en ledtråd till att antingen helt ändra sätten på vilka prediktionerna produceras, eller avstå från att producera prediktioner helt. Negativa värden på I_l trunckeras därför till noll, därför gäller

$$0 \leq I_l \leq 1. \quad (5)$$

Ett naturligt krav är att prediktionerna ska konvergera. Dessutom önskar man att konvergensen ska vara monoton och snabb vid början av processen och stabilisera sig vid slutet. En icke-monoton konvergens innebär att skattningens fel blir större mellan två på varandra följande prediktioner. Detta kan hända om de senaste prediktionerna i tidsserien är så pass dåliga till den grad att en prediktion är reviderad i fel riktning.

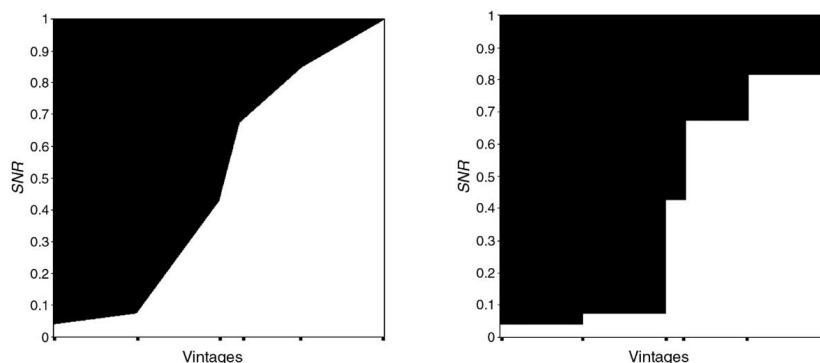
Om snabb konvergens endast uppstår vid ett sent skede av skattningens process, väcker det misstankar om tillförlitligheten i den slutliga siffran. Om inte y_{Lt} är nära det okända värdet av X_t , då gäller inte analysen, men vi kan konstatera att inget är känt om variabeln i fråga.

I Öller och Teterukovsky (2007) presenteras diagramstypen informationsfönster, som visar förändringar av information för X_t med tiden och hur prediktionsfelet minskar när tidshorisonten minskar. Informationsfönstret illustrerar med andra ord prediktionsindex för tidsaxeln l , där $l = 1, \dots, L$.

Det finns inget självklart sätt att binda ihop punkterna till en kurva, men Öller och Teterukovsky har två förslag, nämligen utifrån en producents och

en användares synvinkel. En producent har statistisk kunskap, vilket innebär att producenten får in ny data successivt och därför gjorde de ett godtyckligt antagande att informationen ökar relativt linjärt från l till $l + 1$, medan för användaren som ej erhåller statistisk kunskap, ökar inte informationen.

Nedan visas exempel på en producents och en användares informationsfönster.



Figur 1: Bilden t.v illustrerar en producents informationsfönster och t.h en användares.

Det vita området markerar informationen vi har om X_t och det svarta området betraktas som en gardin som täcker för den okända informationen. Ambitionen är därför att få så ljusa fönster som möjligt.

Låt oss beteckna den vita arean i bilden till vänster, i Figur 1, med \mathcal{I}_p , som är den integrerade prediktionsindexen. Värdet på \mathcal{I}_p beräknas genom formeln

$$\mathcal{I}_p = \frac{1}{2} \sum_{l=1}^{L-1} (I_l + I_{l+1}) \tau(l, l+1) \quad (6)$$

där $\tau(l, l+1)$ är längden av tidsintervallet mellan prediktionerna l och $l+1$, där de summeras till 1. Både de vertikala och horisontella axlarna av diagrammet är av längden 1. Det totala kvalitetsmättet, \mathcal{I}_p kommer vara mellan $\tau(L, L+1)/2$ och 1. Medan för den integrerade prediktionsindexen, \mathcal{I}_a , i bilden till höger (användarens informationsfönster) beräknas enligt följande,

$$\mathcal{I}_a = \sum_{l=0}^{L-1} I_l \tau(l, l+1). \quad (7)$$

Kvaliten på dessa prognoser skulle eventuellt kunna granskas i ett informationsfönster. Huvudsyftet med detta examensarbete är att göra en generell analys av informationsfönstrets tillförlitlighet genom att simulera data från autoregressiv- och moving average processer. Mer specifikt undersöker vi frågeställningarna: är informationsfönster tillförlitlig? Kan ett informationsfönster

eventuellt hjälpa oss att välja modell för en dataserie? Beroende på hur man väljer koefficienterna i en *AR*-eller *MA*-process blir den mer eller mindre predikterbar, återspeglas detta i hur informationsfönsterna ser ut?

1.2 Centrala begrepp

I denna uppsats är stationäritet och prediktion centrala begrepp. Därför följer en kort beskrivning av dessa.

1.2.1 Stationäritet

Man skiljer på två typer av stationäritet, dels strikt stationär och dels svagt stationär. En stokastisk process är strikt stationär om den simultana fördelningen är konstant över tiden. För en stokastisk process som är svagt stationär måste följande villkor uppfyllas:

$$E[X_t] = \mu < \infty \quad (8)$$

$$Var(X_t) = \sigma^2 < \infty \quad (9)$$

$$\psi(k) = E[(X_t - \mu)(Y_{t+k} - \mu)] < \infty \quad (10)$$

Enligt första och andra villkoret gäller det att tidsserien har konstant medelvärde och varians över tiden och kovariansen beror endast på avståden mellan observationerna och inte på läget, dvs tiden påverkar inte observationerna. (Ross 2010).

Att en stokastisk process är icke-stationär kan bero på trend, drift och även slumpvandring. Man kan exempelvis justera tidsserien genom att bilda differenser, med tex en periods tidsförskjutning. Detta medför till att en serie med linjär trend blir stationär.

1.2.2 Prediktion

En prediktion är en förutsägelse av ett okänt värde i en stokastisk process, baserat på tidigare observerade värden.

2 Beskrivning av data och prediktionsmetoder

2.1 Data

För att illustrera metoden ovan, har data simulerats enligt Autoregression och Moving Average i matlab. Vi kommer att undersöka två sätt att simulera tidsserier. Nedan följer en beskrivning av dessa.

Det första sättet går ut på att man simulerar en serie från X_0 till X_{2000} och gör prognos av X_{2000} vid tiderna 1999,1998,...,1995. Sen simulerar man en

ny serie från X_0 till X_{2000} , oberoende av den första, och gör nya prognoser av X_{2000} vid tiderna 1999,1998,...,1995; osv tills man har T oberoende upprepningar. Då är s_L^2 en väntevärdesriktig skattning av $\text{Var}(X_t)$.

Det andra sättet går ut på att man simulerar en enda serie från X_0 till X_{2000} och gör prognos av

X_{2000} vid tiderna 1999,1998,...,1995.

X_{1999} vid tiderna 1998,1997,...,1994.

.

.

.

X_{1991} vid tiderna 1990,1998,...,1986.

Detta ger oss $T = 10$ i formeln (3), men hur ska man tolka s_L^2 ?

I verkligheten är det egentligen bara en serie man har, dvs det andra sättet som är beskriven här ovan. Därför kan man ifrågasätta vad s_L^2 egentligen är en skattning av, för att det är ju inte en väntevärdesriktig skattning av den stationära processens varians. (Mer om detta kommer i avsnitt 4). Vi kommer att undersöka tidsserier som är simulerade enligt båda sätten. För det första sättet (Oberoende replikat) kommer vi att betrakta $L = 6$ prediktioner och $T = 10000$ oberoende upprepningar och för det andra sättet (En serie) $L = 6$ prediktioner och $T = 10$. Punkterna i informationsfönstret kommer att bindas enligt en producents perspektiv.

2.2 Undersökta prediktionsmetoder

Vi gör en serie prediktioner av X_{2000} genom att beräkna det betingade väntevärdet av X_{2000} givet de tidigare observerade värdena.

Den första prediktionen ($l = 1$) görs när vi just har observerat $X_{2000-(L-1)}$, vilket ger oss $y_{1t} = E[X_{2000}|X_{2000-(L-1)}]$ och de övriga ($l = 2, \dots, L$) precis på samma sätt när vi observerar $X_{2000-(L-2)}$, $X_{2000-(L-3)}, \dots, X_{2000}$. Det sanna värdet X_t antas vara y_{Lt} , det vill säga den sista prediktionen som görs, $y_{Lt} = X_{2000}$.

Vi kommer även studera en serie prediktioner av X_{2000} , där prediktionerna är det senaste observerade värdet, dvs $y_{lt} = X_{2000-(L-l)}$. Alltså då ingen förändring sker.

V har därmed fått två prediktionsmetoder att jämföra. För att särskilja dessa två prediktionsmetoder, kommer den förstnämnda prediktionsmeto-

den, dvs ”den betingade väntevärdesmetoden”, i fortsättningen att kallas för MSE-metoden, medan den andra för, persistensmetoden.

3 Oberoende replikat

3.1 Autoregressiv process

En autoregressiv (AR) modell är en typ av slumpmässig process som ofta används för att modellera och förutsäga tidsserier. Utgångspunkten är att förklara ett framtida värde med hjälp av tidigare värden. (Brockwell och Davis 1996).

Den autoregressiva processen betecknas $AR(p)$ och indikerar en autoregressiv modell av graden p . $AR(p)$ modellen definieras enligt följande

$$X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t \quad (11)$$

där ϕ_1, \dots, ϕ_p är parametrar, c är en konstant och ϵ_t är vitt-brus, $\epsilon_t \sim WN(0, \sigma^2)$, $E[\epsilon_t] = 0$, $Var(\epsilon_t) = \sigma^2$.

$AR(p)$ modellen kan även skrivas på formen

$$\phi_p(B)X_t = \epsilon_t \quad (12)$$

där $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ och B är backshift operator som funktionerar på ett värde i en tidsserie för att producera föregående värde.

För att den autoregressiva processen ska vara stationär, krävs vissa begränsningar. Generellt gäller, för att en $AR(p)$ modell ska vara stationär måste rötterna till $\phi_p(B) = 0$ ligga utanför enhetscirkeln, det vill säga, varje rot B_i måste uppfylla $|B_i| > 1$.

Den enklaste Autoregressiva modellen är $AR(0)$ och det är då endast felet som bidrar till X_t . Med andra ord motsvarar $AR(0)$ fel-termen, ϵ_t . För den högre graden, $AR(1)$ -modellen, är det endast den föregående termen och feltermen som bidrar X_t . Ju närmare ϕ_1 är 0, desto mer ser processen ut som vitt-brus, men ju mer ϕ_1 närmar sig 1, får X_t ett större bidrag från den tidigare termen i förhållande till felet. För en $AR(2)$ -modell är det de två tidigare termerna och fel-termen som bidrar till X_t . Vi kommer endast se på den första och den andra graden av den autoregressiva modellen.

3.1.1 Autoregressiv process av grad 1

Vi inleder med att undersöka en AR(1) process som ser ut som följande

$$(1 - \phi_1 B)X_t = \epsilon_t \text{ eller } X_t = \phi_1 X_{t-1} + \epsilon_t \quad (13)$$

Som vi såg ovan, måste roten av $(1 - \phi_1 B)$ ligga utanför enhetscirkeln för att den ska vara stationär, vilket innebär att $|\phi_1| \leq 1$. AR(1) processen är en Markov process för att fördelningen för X_t bestäms helt utifrån kunskapen av den förgående termen, X_{t-1} .

Nedan följer en härledning av prediktionerna och prediktionsindex för $l=1, \dots, 6$ när vi utgår ifrån MSE-metoden.

3.1.2 Härledning av prediktionsformler för Autoregressiva processen av grad 1

Antag att vi har observerat information fram till tidpunkten Z_0 . Låt \hat{Z}_t vara den prediktion vi gör av Z_t , vid tiden 0, alltså $\hat{Z}_t = E[Z_t | \mathbf{Z}_0]$. Låt δ_t vara prediktionsfelet, $\delta_t = Z_t - \hat{Z}_t$.

Låt oss betrakta en AR(1)-process som är på formen, $Z_1 = \phi_1 Z_0 + \epsilon_1$.

Ett-steps prediktionen blir

$$\hat{Z}_1 = \phi_1 Z_0 \quad (14)$$

med prediktionsfelet $\delta_1 = Z_1 - \phi_1 Z_0$ som ger $\delta_1 = \epsilon_1$ och alltså är prediktionsfelets kvadratiska förväntade storlek $E[\epsilon_1^2] = \sigma^2$.

Om processen är stationär är $Var(Z_t) = \frac{\sigma^2}{1 - \phi_1^2}$ för alla t och alltså borde kvoten $\frac{MSE_{L-1}}{s_L^2}$ vara $\frac{\sigma^2}{\sigma^2/(1 - \phi_1^2)} = (1 - \phi_1^2)$.

Prediktionsindex för en ett-steps prognos, alltså det som med beteckningar enligt (4) är I_{L-1} borde därför bli ungefär

$$I_{L-1} \approx 1 - (1 - \phi_1^2) = \phi_1^2 \quad (15)$$

Vidare är två-steps prediktionen

$$\hat{Z}_2 = E[Z_2 | \mathbf{Z}_0] = \phi_1 E[Z_1 | \mathbf{Z}_0] + E[\epsilon_1 | \mathbf{Z}_0] = \phi_1 \hat{Z}_1 \quad (16)$$

som ger prediktionsfelet

$$\delta_2 = Z_2 - \hat{Z}_2 = Z_2 - \phi_1 \hat{Z}_1 = \phi_1 Z_1 + \epsilon_2 - \phi_1 \hat{Z}_1 = \phi_1 \delta_1 + \epsilon_2 = \phi_1 \epsilon_1 + \epsilon_2. \quad (17)$$

Prediktionsfelets kvadratiska förväntade storlek blir då

$$E[\delta_1^2] = \phi_1^2 E[\epsilon_1^2] + E[\epsilon_2^2] = (\phi_1^2 + 1)\sigma^2 \quad (18)$$

vilket ger oss kvoten

$$\frac{MSE_{L-2}}{s_L^2} \approx \frac{(\phi_1^2 + 1)\sigma^2}{\sigma^2/(1 - \phi_1^2)}. \quad (19)$$

och prediktionsindex blir därmed

$$I_{L-2} = \phi_1^4. \quad (20)$$

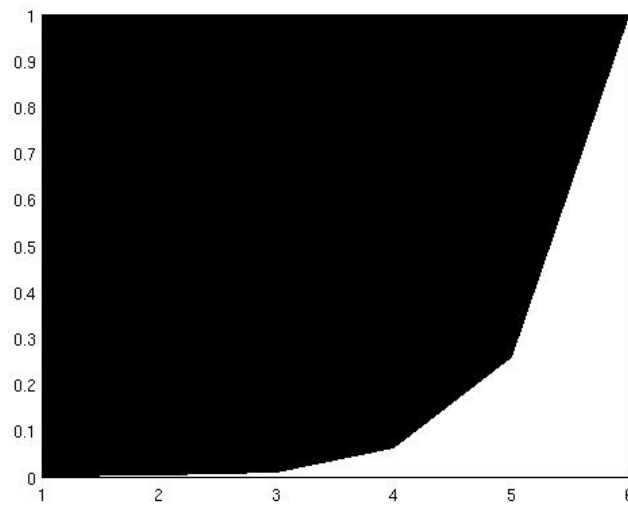
Det är nu lätt att inse att

$$I_{L-k} = \phi_1^{2k} \quad (21)$$

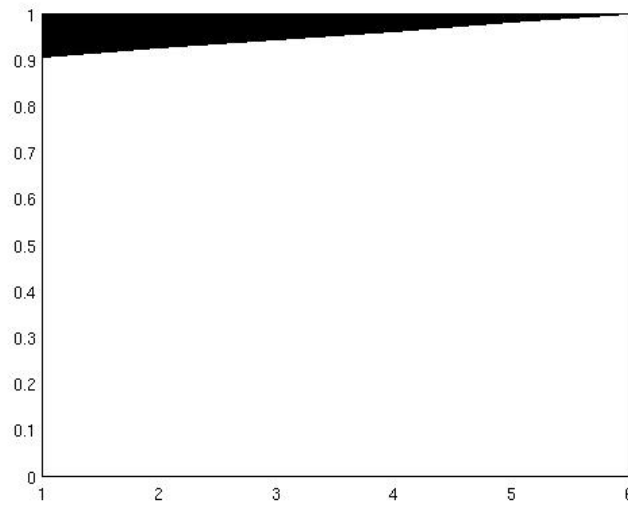


Tack vare stationäritet hos processen X_t kan vi använda oss av det härledda resultatet för att undersöka prediktionerna av X_{2000} .

Låt oss nu se på två informationsfönster med den autoregressiva processen av grad 1, där en serie prediktioner görs av X_{2000} genom att beräkna det betingade väntevärdet av X_{2000} givet de tidigare observerade värdena, dvs MSE-metoden.



Figur 2: Informationsfönster för modellen $X_t = 0.5X_{t-1} + \epsilon_t$ med arean 0.1674.



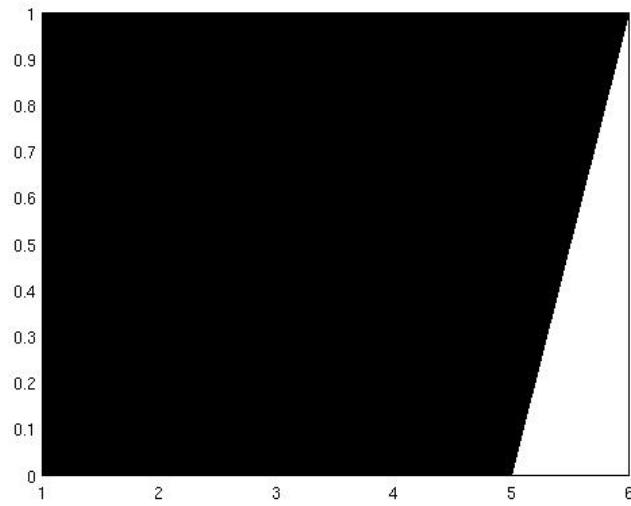
Figur 3: Informationsfönster för modellen $X_t = 0.99X_{t-1} + \epsilon_t$ med arean 0.9514.

Tabellerna nedan illustrerar prediktionsindexen vi erhöill från simuleringen som vi väljer att beteckna, $I_{l,S}$, samt teoretiska, $I_{l,T}$. Den första tabellen jämför modellen $X_t = 0.5X_{t-1} + \epsilon_t$, där $\phi_1 = 0.5$ och den andra tabellen $X_t = 0.99X_{t-1} + \epsilon_t$, där $\phi_1 = 0.99$.

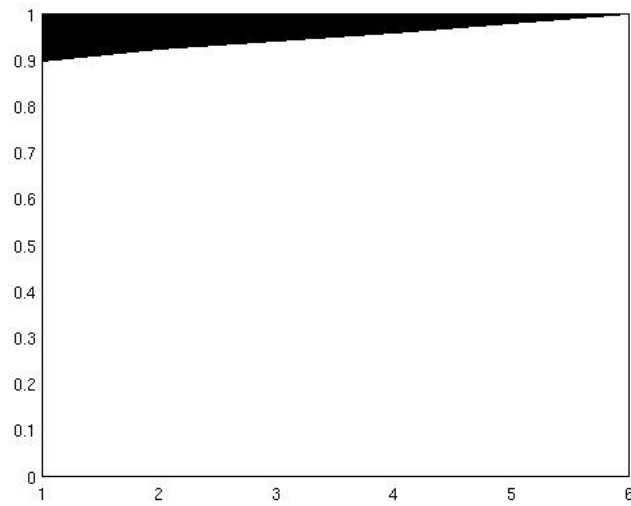
I_l	$I_{l,S}$	$I_{l,T}$	I_l	$I_{l,S}$	$I_{l,T}$
1	0	0.0001	1	0.9060	0.9044
2	0.0043	0.0039	2	0.9257	0.9227
3	0.0112	0.0156	3	0.9441	0.9415
4	0.0615	0.0625	4	0.9618	0.9606
5	0.2600	0.2500	5	0.9809	0.9801
6	1	1	6	1	1

Prediktionsindexen som erhöills av serien prediktioner som har gjorts via den simulerade AR(1)-process överensstämmer väl med teorin. Detta säger oss att informationsfönsterna ger den förväntade informationen om prediktionerna. Vi ser även att prediktionerna är mycket bättre när $\phi_1 = 0.99$ än $\phi_1 = 0.5$.

Vi fortsätter med att illustrera exempel på två informationsfönster med persistensmetoden, det vill säga då vi predikterar det sanna värdet med det senaste observerade värdet, $y_{lt} = X_{2000-(L-l)}$.



Figur 4: Informationsfönster för modellen $X_t = 0.5X_{t-1} + \epsilon_t$ med arean 0.1.



Figur 5: Informationsfönster för modellen $X_t = 0.99X_{t-1} + \epsilon_t$ med arean 0.9493.

I figur 4 ser vi att prediktionsindex är mindre än eller lika med noll för alla prediktioner, vilket innebär att prediktionerna är helt informationslösa. Vi noterar även att $I_6=1$, vilket just beror på att vi identifierar $y_{6,t}$ med X_{2000} och att arean är 0.1 i det här fallet, säger oss egentligen ingenting. Medan i figur 5 ser vi ett betydligt ljusare fönster där prediktionsindex ökar linjärt fram till den slutgiltiga siffran. Vi kan utifrån detta konstatera att koefficienten, ϕ_1 har en stor påverkan på modellerna, ju större den är, desto ljusare och informationsrikare fönster får vi.

En ökning av modellen skulle eventuellt kunna ge en förbättring. Låt oss fortsätta med att undersöka den Autoregressiva modellen av grad två.

3.1.3 Autoregressiv process av grad 2

För den andra graden av den Autoregressiva processen gäller

$$(1 - \phi_1 B - \phi_2 B^2)X_t = \epsilon_t \text{ eller } X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \epsilon_t$$

Rötterna till polynomet $1 - \phi_1 B - \phi_2 B^2 = 0$ måste ligga utanför enhetscirkeln för att tidsserien ska vara stationär. Kriterierna för stationaritet för AR(2) modellen kan även uttryckas i termer av dess parameter värden. Låt B_1 och B_2 vara rötterna till $1 - \phi_1 B - \phi_2 B^2 = 0$ som är ekvivalent med $\phi_2 B^2 + \phi_1 B - 1 = 0$.

Vi har

$$B_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \text{ och } B_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}$$

Detta kan skrivas om till

$$\frac{1}{B_1} = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} \text{ och } \frac{1}{B_2} = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

Det nödvändiga villkoret $|B_i| > 1$ medför att $|\frac{1}{B_i}| < 1$ för $i=1,2$, vilket ger oss

$$\begin{aligned} \left| \frac{1}{B_1} \cdot \frac{1}{B_2} \right| &= |\phi_2| < 1 \\ \left| \frac{1}{B_1} + \frac{1}{B_2} \right| &= |\phi_1| < 2 \end{aligned}$$

Vi undersöker olika fall som uppfyller detta. Vi får följande nödvändiga villkor för stationaritet som måste uppfyllas oavsett om rötterna är reella eller komplexa

$$\begin{cases} -1 < \phi_2 < 1 \\ -2 < \phi_1 < 2 \end{cases}$$

För reella rötter måste $\phi_1 + 4\phi_2 \geq 0$ vilket ger oss,

$$-1 < \frac{1}{B_2} = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2} \leq \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} = \frac{1}{B_1} < 1,$$

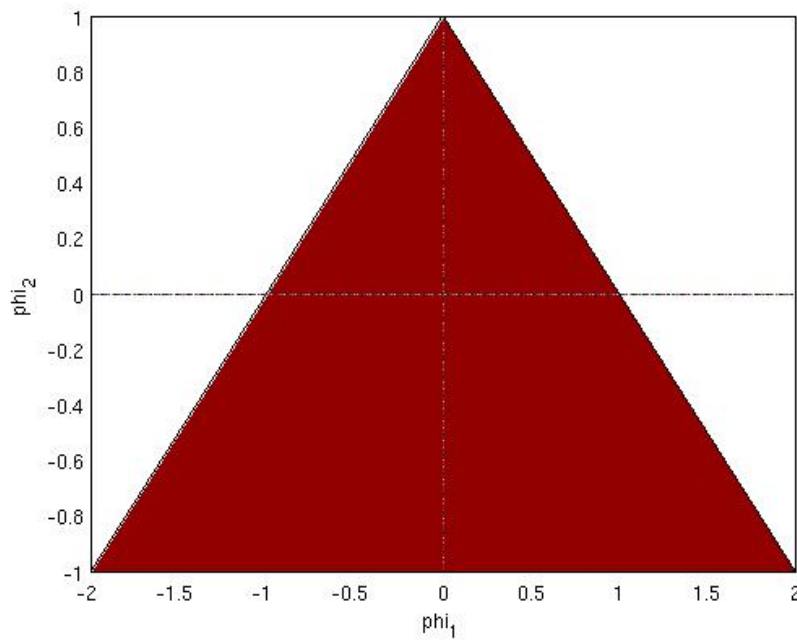
som är ekvivalent med

$$\begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \end{cases}$$

För komplexa rötter har vi $\phi_2 < 0$ och $\phi_1 + 4\phi_2 < 0$ som ger oss

$$\begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \\ -1 < \phi_2 < 1 \end{cases}$$

Området i figuren nedan markerar villkoren för reella rötter.



Figur 6

Nedan följer en härledning av prediktionerna och prediktionsindex för $l=1, \dots, 6$ när vi utgår ifrån MSE-metoden.

3.1.4 Härledning av prediktionsformler för Autoregressiva processen av grad 2

Antag att vi har observerat information fram till tidpunkten Z_0 . Låt \hat{Z}_t vara den prediktion vi gör av Z_t , vid tiden 0, alltså $\hat{Z}_t = E[Z_t | \mathbf{Z}_0]$. Låt δ_n vara prediktionsfelet, $\delta_t = Z_t - \hat{Z}_t$.

Betrakta en AR(2)-process. Den allmänna formeln lyder

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \epsilon_t.$$

Första-steps prediktionen blir alltså

$$\hat{Z}_1 = E[\phi_1 Z_0 + \phi_2 Z_{-1} + \epsilon_1 | \mathbf{Z}_0] = \phi_1 Z_0 + \phi_2 Z_{-1}$$

med prediktionsfelet

$$\delta_1 = Z_1 - \phi_1 Z_0 - \phi_2 Z_{-1} = \epsilon_1. \quad (22)$$

Det kvadratiske prediktionsfelet blir då σ^2 .

Variansen av en AR(2)-modell är mer komplicerad att härleda än en AR(1)-modell. Låt oss titta på härledningen av AR(2)-modellens varians. Vi har

$$\text{Var}(Z_t) = \text{Var}(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \epsilon_t)$$

$$\text{Var}(Z_t) = \text{Var}(\phi_1 Z_{t-1}) + \text{Var}(\phi_2 Z_{t-2}) + 2\text{Cov}(\phi_1 Z_{t-1}, \phi_2 Z_{t-2}) + \text{Var}(\epsilon_t)$$

Tack vare stationäritet är variansen av Z_t densamma för alla t . Kovariansen är

$$\text{Cov}(\phi_1 Z_{t-1}, \phi_2 Z_{t-2}) = \frac{\phi_1^2 \phi_2}{1 - \phi_2} \text{Var}(Z_t)$$

och vi får då

$$\text{Var}(Z_t)(1 - \phi_1^2 - \phi_2^2) = 2 \frac{\phi_1^2 \phi_2}{1 - \phi_2} \text{Var}(Z_t) + \sigma^2.$$

Löser vi ut $\text{Var}(Z_t)$ får vi

$$\text{Var}(Z_t) = \frac{\sigma^2}{(1 - \phi_1^2 - \phi_2^2) - 2 \frac{\phi_1^2 \phi_2}{1 - \phi_2}}.$$

Prediktionsindexen blir alltså

$$1 - \frac{\sigma^2}{\sigma^2/(1 - \phi_1^2 - \phi_2^2 - 2\frac{\phi_1^2\phi_2}{1-\phi_2})} = \phi_1^2 + \phi_2^2 + 2\frac{\phi_1^2\phi_2}{1 - \phi_2} \quad (23)$$

Vidare är

$$Z_2 = \phi_1 Z_1 + \phi_2 Z_0 + \epsilon_2.$$

Två-steps prediktionen blir

$$\hat{Z}_2 = E[\phi_1 Z_1 + \phi_2 Z_0 + \epsilon_2 | \mathbf{Z}_0] = \phi_1 E[Z_1 | \mathbf{Z}_0] + \phi_2 E[Z_0 | \mathbf{Z}_0] = \phi_1 \hat{Z}_1 + \phi_2 Z_0.$$

Prediktionsfelet blir

$$\delta_2 = Z_2 - \hat{Z}_2 = \phi_1 Z_1 + \phi_2 Z_0 + \epsilon_2 - \phi_1 \hat{Z}_1 - \phi_2 Z_0 = \phi_1 \epsilon_1 + \epsilon_2 \quad (24)$$

och därmed får vi det kvadratiske prediktionsfelet till

$$E[\delta_2^2] = (1 + \phi_1^2)\sigma^2.$$

Slutligen får vi prediktionsindexen till

$$I_{L-2} = 1 - (1 + \phi_1^2)(1 - \phi_1^2 - \phi_2^2 - 2\frac{\phi_1^2\phi_2}{1 - \phi_2}).$$

På samma sätt får vi tre-steps prediktionen

$$\hat{Z}_3 = E[\phi_1 Z_2 + \phi_2 Z_1 + \epsilon_3 | \mathbf{Z}_0] = \phi_1 \hat{Z}_2 + \phi_2 \hat{Z}_1.$$

Prediktionsfelet blir

$$\begin{aligned} \delta_3 &= Z_3 - \hat{Z}_3 = \phi_1 Z_2 + \phi_2 Z_1 + \epsilon_3 - \phi_1 \hat{Z}_2 - \phi_2 \hat{Z}_1 \\ &= \phi_1 (Z_2 - \hat{Z}_2) + \phi_2 (Z_1 - \hat{Z}_1) + \epsilon_3 \end{aligned}$$

Från ekvationerna (22) och (24) får vi det prediktionsfelet till

$$\delta_3 = \phi_2 \epsilon_1 + \phi_1^2 \epsilon_1 + \phi_1 \epsilon_2 + \epsilon_3 \quad (25)$$

och väntevärdet av det kvadratiske prediktionsfelet blir därmed

$$E[\delta_3^2] = (1 + \phi_1 + (\phi_1^2 + \phi_2)^2)\sigma^2.$$

Prediktionsindex blir då

$$I_{L-3} = 1 - (1 + \phi_1 + (\phi_1^2 + \phi_2)^2)(1 - \phi_1^2 - \phi_2^2 - 2\frac{\phi_1^2\phi_2}{1 - \phi_2}) \quad (26)$$

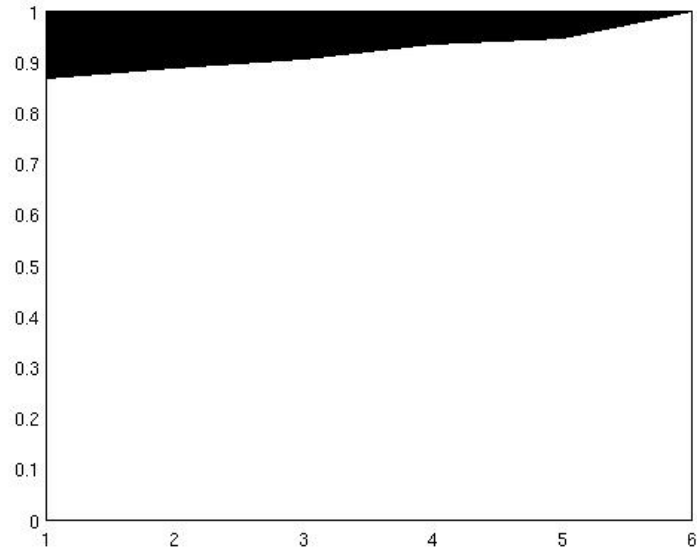
Principen är densamma för I_{L-4} och I_{L-5} .

$$I_{L-4} = 1 - (1 + \phi_1^2 + (\phi_1^3 + 2\phi_1\phi_2)^2 + (\phi_1^2 + \phi_2)^2)(1 - \phi_1^2 - \phi_2^2 - 2\frac{\phi_1^2\phi_2}{1 - \phi_2}) \quad (27)$$

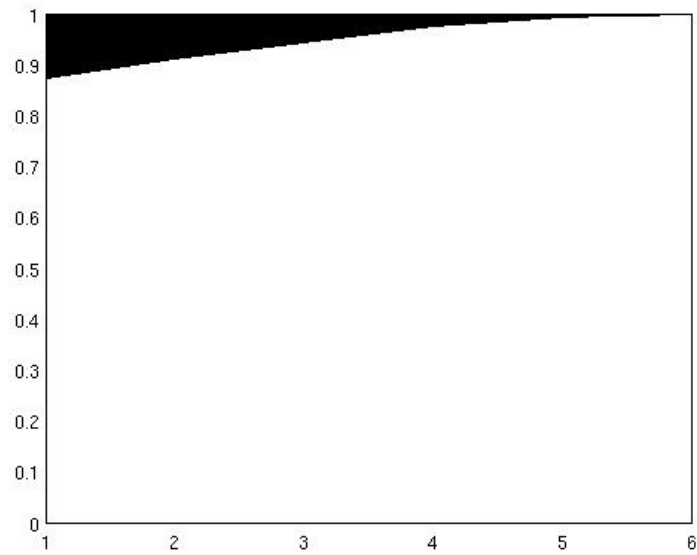
$$I_{L-5} = 1 - (1 + \phi_1^2 + (\phi_1^2 + \phi_2)^2 + (2\phi_1\phi_2 + \phi_1^3)^2 + (\phi_1^4 + 3\phi_1^2\phi_2 + \phi_2^2)^2)(1 - \phi_1^2 - \phi_2^2 - 2\frac{\phi_1^2\phi_2}{1 - \phi_2}) \quad (28)$$

■

Låt oss nu se på två informationsfönster med den autoregressiva processen av grad 2, där en serie prediktioner görs av X_{2000} genom att betinga väntevärdet av X_{2000} på de tidigare observerade värden, dvs MSE-metoden.



Figur 7: Informationsfönster för modellen $X_t = 0.5X_{t-1} + 0.48X_{t-2} + \epsilon_t$ med arean 0.9213.



Figur 8: Informationsfönster för modellen $X_t = 1.62X_{t-1} - 0.63X_{t-2} + \epsilon_t$ med arean 0.9514.

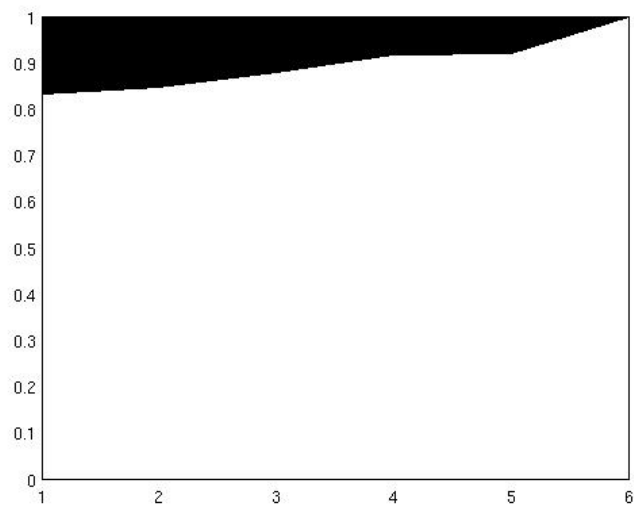
Tabellerna nedan illustrerar prediktionsindex vi erhö ll fr n simuleringen samt de teoretiska. Den f rsta tabellen j mf r modellen $X_t = 0.5X_{t-1} + 0.48X_{t-2} + \epsilon_t$ och den andra tabellen $X_t = 1.62X_{t-1} - 0.63X_{t-2} + \epsilon_t$.

I_l	$I_{l,S}$	$I_{l,T}$
1	0.8661	0.8505
2	0.8882	0.8752
3	0.9063	0.8965
4	0.9328	0.9274
5	0.9462	0.9419
6	1	1

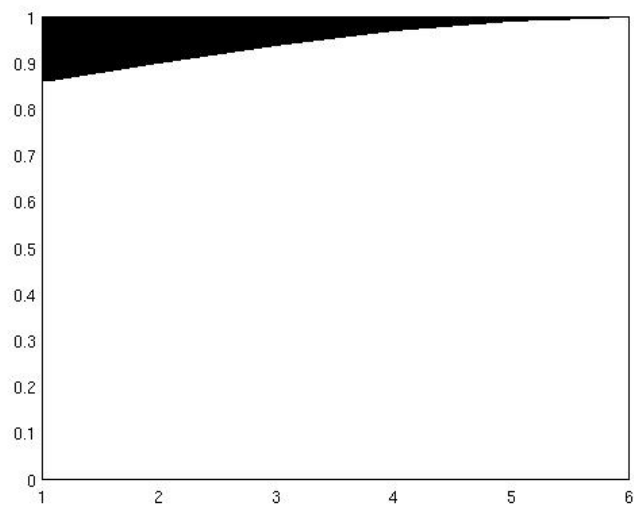
I_l	$I_{l,S}$	$I_{l,T}$
1	0.8723	0.8680
2	0.9100	0.9079
3	0.9444	0.9439
4	0.9738	0.9733
5	0.9926	0.9926
6	1	1

Vi ser att prediktionsindexen  ven f r AR(2)-processen  verensst mmer bra med teorin, vilket s ger oss att informationsf nstret ger oss den f rväntade illustrationen av kvaliteten.

P  n sta sida f ljer tv  exempel p  informationsf nster, n r vi utg r fr n persistensmetoden.



Figur 9: Informationsfönster för modellen $X_t = 0.5X_{t-1} + 0.48X_{t-2} + \epsilon_t$ med arean 0.8951.



Figur 10: Informationsfönster för modellen $X_t = 1.62X_{t-1} - 0.63X_{t-2} + \epsilon_t$ med arean 0.9451.

Informationsfönstrena ovan säger oss att prediktionerna av X_{2000} är av god kvalite. Som vi spekulerade innan gav en utökning av modellen en förbättring av kvaliteten, för att precisionen av estimatoren ökar ju fler parametrar man arbetar med.

3.2 Moving average process

X_t är en Moving average process av ordning q om

$$X_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \dots + \theta_q\epsilon_{t-q}, \quad (29)$$

där $\epsilon_t \sim WN(0, \sigma^2)$ och $\theta_1, \dots, \theta_q$ är konstanter. Moving average processen av ordningen q betecknas med $MA(q)$.

3.2.1 Moving average process av grad 1

En simpel Moving Average är en serie X_t från en white noise serie ϵ_t ,

$$X_t = \theta_1\epsilon_{t-1} + \epsilon_t. \quad (30)$$

som även kan skrivas på formen

$$X_t = (1 - \theta_1 B)\epsilon_t. \quad (31)$$

Den enkla Moving Average processen betecknas $MA(1)$, eftersom den endast innehåller en parameter, θ_1 . (Brockwell och Davis 1996).

Nedan följer en härledning av prediktionerna och prediktionsindex för $l=1, \dots, 6$ när vi utgår ifrån MSE-metoden.

3.2.2 Härledning av prediktionsformler för Moving average av grad 1

Antag att vi har observerat information fram till tidpunkten Z_0 . Låt \hat{Z}_t vara den prediktion vi gör av Z_t , vid tiden 0, alltså $\hat{Z}_t = E[Z_t | \mathbf{Z}_0]$. Låt δ_t vara prediktionsfelet, $\delta_t = Z_t - \hat{Z}_t$.

För alla t gäller

$$Z_t = \theta_1\epsilon_{t-1} + \epsilon_t \quad (32)$$

Alltså

$$Z_1 = \theta_1\epsilon_0 + \epsilon_1 \quad (33)$$

och vi får då ett-steps prediktionen till

$$\hat{Z}_1 = E[\theta_1 \epsilon_0 + \epsilon_1 | \mathbf{Z}_0] = \theta_1 E[\epsilon_0 | \mathbf{Z}_0] \quad (34)$$

$$Z_0 = \theta_1 \epsilon_{-1} + \epsilon_0$$

Då gäller

$$\begin{pmatrix} Z_0 \\ \epsilon_0 \end{pmatrix} = \begin{pmatrix} \theta & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_{-1} \\ \epsilon_0 \end{pmatrix}$$

där

$$\begin{pmatrix} \epsilon_{-1} \\ \epsilon_0 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma^2 I_2 \right)$$

Kovarians matrisen är

$$\begin{aligned} \text{Var} \begin{pmatrix} Z_0 \\ \epsilon_0 \end{pmatrix} &= \begin{pmatrix} \text{Var}(Z_0) & \text{Cov}(Z_0, \epsilon_0) \\ \text{Cov}(Z_0, \epsilon_0) & \text{Var}(\epsilon_0) \end{pmatrix} = \\ &= \begin{pmatrix} (\theta_1^2 + 1)\sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 \end{pmatrix} \end{aligned}$$

och från kovarians matrisen kan man enkelt härleda $E[\epsilon_0 | \mathbf{Z}_0]$.

$$E[\epsilon_0 | Z_0 = z_0] = \mu_{\epsilon_0} + \frac{\text{Cov}(Z_0, \epsilon_0)}{(\theta_1^2 + 1)\sigma^2} (z_0 - \mu_{z_0}) = 0 + \frac{1}{(\theta_1^2 + 1)} (\theta_1 \epsilon_{-1} + \epsilon_0 - 0) \quad (35)$$

Vi lägger in ovanstående uttryck i equation (33) och får då

$$\hat{Z}_1 = \frac{\theta_1}{(\theta_1^2 + 1)} (\theta_1 \epsilon_{-1} + \epsilon_0)$$

Prediktionsfelet blir

$$\delta_1 = Z_1 - \hat{Z}_1 = \theta_1 \epsilon_0 + \epsilon_1 - \frac{\theta_1}{(\theta_1^2 + 1)} (\theta_1 \epsilon_{-1} + \epsilon_0)$$

och det förväntade kvadratiske prediktionsfelet blir

$$\begin{aligned} E[\delta_1^2] &= \left(\theta_1 - \frac{\theta_1}{\theta_1^2 + 1} \right)^2 + 1 + \left(\frac{\theta_1^2}{\theta_1^2 + 1} \right) = \dots = \\ &= \frac{\theta_1^4 + \theta_1^2 + 1}{\theta_1^2 + 1} \end{aligned}$$

vilket ger oss följande

$$I_{L-1} = 1 - \left(\frac{\theta_1^4 + \theta_1^2 + 1}{(\theta_1^2 + 1)^2} \right) = \frac{\theta_1^2}{(\theta_1^2 + 1)^2}$$

Vidare är

$$Z_2 = \theta_1 \epsilon_1 + \epsilon_2$$

och två-steps prediktionen blir

$$\hat{Z}_2 = E[\theta_1 \epsilon_1 + \epsilon_2 | \mathbf{Z}_0] = 0.$$

Prediktionsfelet blir då

$$\delta_2 = Z_2 - \hat{Z}_2 = \theta_1 \epsilon_1 + \epsilon_2 - 0 \quad (36)$$

och därmed blir det kvadratiske prediktionsfelet

$$E[\delta^2] = (\theta_1^2 + 1)\sigma^2$$

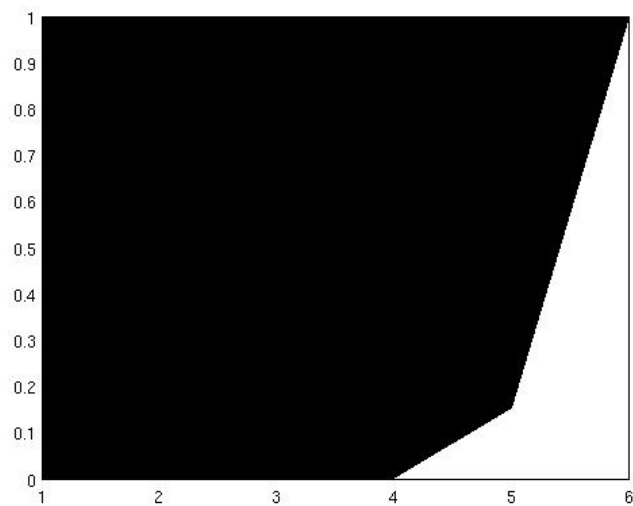
vilket till slut ger oss

$$I_{L-2} = 1 - \frac{(\theta_1^2 + 1)\sigma^2}{(\theta_1^2 + 1)\sigma^2} = 0 \quad (37)$$

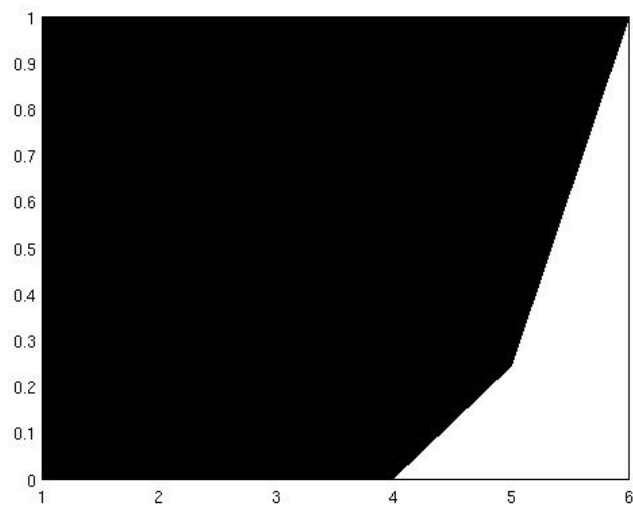
Detta gäller för alla I_{L-k} , där $k > 1$.

■

Låt oss se på två informationsfönster av moving average process av grad 1, där vi gör en serie prediktioner av X_{2000} genom att betinga väntevärdet av X_{2000} på de tidigare observerade värdena, dvs MSE-metoden.



Figur 11: Informationsfönster för modellen $X_t = 0.5\epsilon_t + \epsilon_{t-1}$ med arean 0.1309.



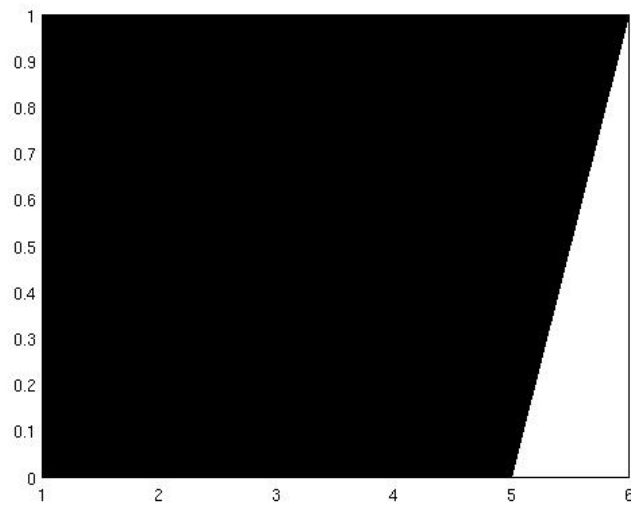
Figur 12: Informationsfönster för modellen $X_t = 0.9\epsilon_t + \epsilon_{t-1}$ med arean 0.1492.

Tabellerna nedan illustrerar de teoretiska värdena på prediktionsindex samt de värden vi får från simuleringen. Den första tabellen jämför $X_t = 0.5\epsilon_{t-1} + \epsilon_t$, där $\theta_1=0.5$ och den andra tabellen $X_t = 0.9\epsilon_{t-1} + \epsilon_t$, där $\theta_1=0.9$.

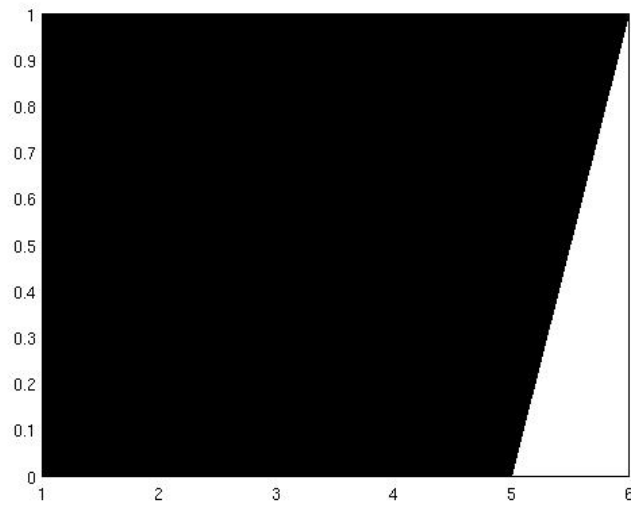
I_l	$I_{l,S}$	$I_{l,T}$	I_l	$I_{l,S}$	$I_{l,T}$
1	0.001	0	1	0.0001	0
2	0.0001	0	2	0.0001	0
3	0.0001	0	3	0.0001	0
4	0.0001	0	4	0.0001	0
5	0.1541	0.16	5	0.2455	0.2472
6	1	1	6	1	1

I tabellerna ovan ser vi att prediktionsindex från den simulerade datan överensstämmer väl med teorin, vilket innebär att informationsfönstret ger en förväntad bild av prediktionernas kvalite. Vi ser även att ett högre värde på koefficienten ger ett högre prediktionsindex.

Vi fortsätter med att illustrera två exempel på informationsfönster med persistensmetoden.



Figur 13: Informationsfönster för modellen $X_t = \epsilon_t + 0.5\epsilon_{t-1}$ med arean 0.1.



Figur 14: Informationsfönster för modellen $X_t = \epsilon_t + 0.9\epsilon_{t-1}$ med arean 0.1.

Informationsfönstrena ovan, visar tydligt att prediktionerna av det sanna värdet är olämpliga. Vi kan utifrån detta konstatera att detta troligtvis gäller alla moving average modeller av grad 1.

3.2.3 Moving average process av grad 2

En Moving Average process av grad 2 defineras enligt följande,

$$X_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} \quad (38)$$

som är ekvivalent med

$$X_t = (1 - \theta_1 B - \theta_2 B^2) \epsilon_t, \quad (39)$$

där $\epsilon_{t-2}, \epsilon_{t-1}, \epsilon_t$ och θ_1, θ_2 är definerade enligt tidigare. (Brockwell och Davis 1996).

Nedan följer en härledning av prediktionerna och prediktionsindex för $l=1, \dots, 6$ när vi utgår ifrån MSE-metoden.

3.2.4 Härledning av prediktionsformler för Moving average processen av grad 2

Härledning av prediktionerna samt prediktionsindex fungerar på samma sätt som vi har sett i dem tidigare avsnitten av *Härledning*. Nedan kommer enbart en påbörjad härledning av ett-steps-prediktion. Resterande överläts till den intresserade läsaren.

Vi gör samma antaganden som innan, det vill säga att vi har observerat information fram till tidpunkten Z_0 och låter \hat{Z}_t vara den prediktion vi gör av Z_t , vid tiden 0, alltså $\hat{Z}_t = E[Z_t | \mathbf{Z}_0]$. Låt δ_t vara prediktionsfelet, $\delta_t = Z_t - \hat{Z}_t$.

Vi vill prediktera

$$Z_1 = \theta_1 \epsilon_0 + \theta_2 \epsilon_{-1} + \epsilon_1 \quad (40)$$

och ett-steps-prediktionen blir då

$$\hat{Z}_1 = E[Z_1 | \mathbf{Z}_0] = \theta_1 E[\epsilon_0 | \mathbf{Z}_0] + \theta_2 E[\epsilon_{-1} | \mathbf{Z}_0] \quad (41)$$

givet

$$Z_0 = \theta_1 \epsilon_{-1} + \theta_2 \epsilon_{-2} + \epsilon_0 \quad (42)$$

Det som behövs är $E[\epsilon_0|\mathbf{Z}_0]$ och $E[\epsilon_{-1}|\mathbf{Z}_0]$.

Kovariansmatrisen för Z_0 och ϵ_0 är

$$\begin{aligned} \text{Var} \begin{pmatrix} Z_0 \\ \epsilon_0 \end{pmatrix} &= \begin{pmatrix} \text{Var}(Z_0) & \text{Cov}(Z_0, \epsilon_0) \\ \text{Cov}(Z_0, \epsilon_0) & \text{Var}(\epsilon_0) \end{pmatrix} = \\ &= \begin{pmatrix} (\theta_1^2 + \theta_2^2 + 1)\sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 \end{pmatrix} \end{aligned}$$

vilket ger oss att

$$E[\epsilon_0|Z_0 = z_0] = \frac{\theta_1\epsilon_{-1} + \theta_2\epsilon_{-2} + \epsilon_0}{\theta_1^2 + \theta_2^2 + 1}. \quad (43)$$

Kovariansmatrisen för Z_0 och ϵ_{-1} är

$$\begin{aligned} \text{Var} \begin{pmatrix} Z_0 \\ \epsilon_{-1} \end{pmatrix} &= \begin{pmatrix} \text{Var}(Z_0) & \text{Cov}(Z_0, \epsilon_{-1}) \\ \text{Cov}(Z_0, \epsilon_{-1}) & \text{Var}(\epsilon_{-1}) \end{pmatrix} = \\ &= \begin{pmatrix} (\theta_1^2 + \theta_2^2 + 1)\sigma^2 & \theta_1\sigma^2 \\ \theta_1\sigma^2 & \sigma^2 \end{pmatrix} \end{aligned}$$

och vi får då

$$E[\epsilon_{-1}|Z_0] = \frac{\theta_1^2\epsilon_{-1} + \theta_1\theta_2\epsilon_{-2} + \theta_1\epsilon_0}{1 + \theta_1^2 + \theta_2^2}. \quad (44)$$

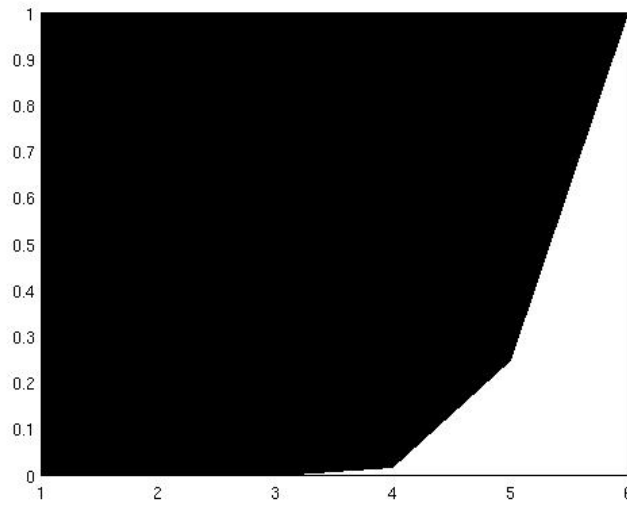
Lägger vi in ekvationerna (42) och (43) i (40) får vi

$$\hat{Z}_1 = \theta_1 \left(\frac{\theta_1\epsilon_{-1} + \epsilon_0}{\theta_1^2 + \theta_2^2 + 1} \right) + \theta_2 \left(\frac{\theta_1^2\epsilon_{-1} + \theta_1\theta_2\epsilon_{-2} + \theta_1\epsilon_0}{1 + \theta_1^2 + \theta_2^2} \right) \quad (45)$$

$$\delta_1 = Z_1 - \hat{Z}_1 = \theta_1\epsilon_0 + \theta_2\epsilon_{-1} + \epsilon_1 - \theta_1 \left(\frac{\theta_1\epsilon_{-1} + \epsilon_0}{\theta_1^2 + \theta_2^2 + 1} \right) - \theta_2 \left(\frac{\theta_1^2\epsilon_{-1} + \theta_1\theta_2\epsilon_{-2} + \theta_1\epsilon_0}{1 + \theta_1^2 + \theta_2^2} \right)$$

■

På nästa sida följer exempel på ett informationsfönster där vi har predikerat X_{2000} genom att beräkna det betingade väntevärdet av X_{2000} givet tidigare observerade värden.



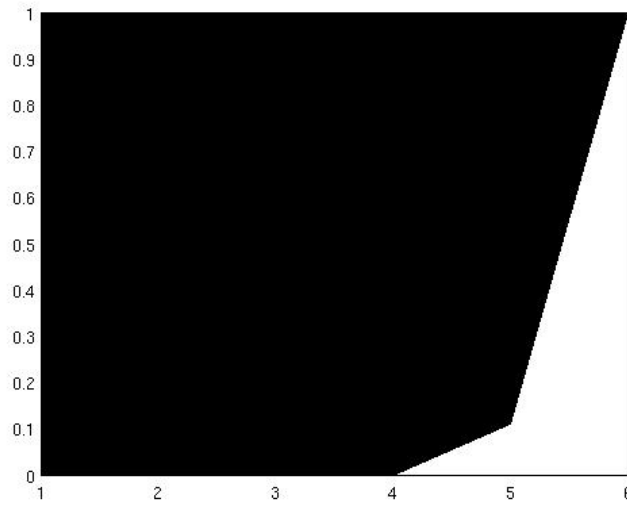
Figur 15: Informationsfönster för modellen $X_t = \epsilon_t + 0.5\epsilon_{t-1} + 0.5\epsilon_{t-2}$ med arean 0.1527.

Tabellen nedan jämför prediktionsindex som erhöles av den simulerade datan med den teoretiska.

I_l	$I_{l,S}$	$I_{l,T}$
1	0.0010	0
2	0.0010	0
3	0	0
4	0.0157	0.0833
5	0.2462	0.2500
6	1	1

Vi ser att prediktionsindexen överensstämmer relativt bra. De fel som slumpen orsakar är små. Detta säger oss att informationsfönstret ger en förväntad bild av prediktionernas kvalite.

Vi fortsätter med att se på ett informationsfönster med persistensmetoden.



Figur 16: Informationsfönster för modellen $X_t = \epsilon_t + 0.5\epsilon_{t-1} + 0.5\epsilon_{t-2}$ med arean 0.1222.

Informationsfönstrena för moving average modellerna ovan, säger oss att de olika modellerna av grad 1 och 2 är olämpliga och att prediktionerna av X_{2000} är i princip informationslösa. Detta beror troligtvis på att moving average processen har ett kort minne.

Vi såg tidigare att alla konstanta prediktioner av X_{2000} som var simulerade via en MA(1) process, var helt informationslösa. Den senaste prediktionen av X_{2000} bör bidra med lite information, eftersom det finns en korrelation mellan X_{2000} och X_{1999} . Även för MA(2)-processen förväntar man sig att den fjärde prediktionen av X_{2000} ska vara bättre.

4 Utan replikat

Låt oss nu se på en tidsserie som är simulerad enligt det andra sättet som står beskrivet i avsnitt *Beskrivning av data och prediktionsmetoder*.

Som vi nämnde tidigare är inte s_L^2 en väntevärdesriktig skattning av det sanna värdets varians. Vi kommer nu undersöka det närmare.

$$E[s_L^2] = \frac{1}{T-1} \sum_{i=1}^T E(X_i - \bar{X})^2 = \frac{1}{T-1} \frac{1}{2T} \sum_{i=1}^T \sum_{j=1}^T E(X_i - X_j)^2 \quad (46)$$

$$\begin{aligned} \sum_{i=1}^T \sum_{j=1}^T E[(X_i - X_j)]^2 &= \sum_{i=1}^T \sum_{j=1}^T \text{Var}(X_i) + \text{Var}(X_j) - 2\text{Cov}(X_i, X_j) = \\ &= \sum_{i=1}^T \sum_{j=1}^T 2\sigma^2 - 2\psi(i-j) \end{aligned} \quad (47)$$

Eftersom X_t inte är oberoende får vi en kovarians som är större än noll. Låt oss titta på AR(1)-processens kovarians.

AR(1)-processen defineras på samma sätt som (13). Låt $k \geq 0$. Vi multiplicerar det autoregressiva sambandet med X_{t-k} och får då

$$X_{t-k}X_t = \phi_1 X_{t-k}X_{t-1} + X_{t-k}\epsilon_t \quad (48)$$

Genom att bilda väntevärdet i de båda leden får vi:

$$E[X_{t-k}X_t] = E[\phi_1 X_{t-k}X_{t-1}] + E[X_{t-k}\epsilon_t] \quad (49)$$

X_{t-k} beror endast på $\epsilon_{t-k}, \epsilon_{t-1-k}, \dots$ och är därmed okorrelerad med ϵ_t .

För $k > 0$ fås därför att $E[X_{t-k}\epsilon_t] = E[X_{t-k}]E[\epsilon_t] = 0$, vilket då ger oss den rekursiva kovariansen

$$\psi_X(k) = \phi_1 \psi_X(k-1). \quad (50)$$

Vi ser att $\psi_X(1) = \phi_1 \psi_X(0)$ där $\psi_X(0) = \sigma_X^2$, ger oss sambandet

$$\psi(k) = \sigma_X^2 \phi_1^{|k|}, k = 0, \pm 1, \pm 2, \dots \quad (51)$$

$$E[s_L^2] = \frac{T}{T-1} \sigma^2 - \frac{1}{T-1} \frac{1}{T} \sum_{i=1}^T \sum_{j=1}^T \phi_1^{|i-j|} = \frac{\sigma^2}{T-1} \left(T - \frac{1}{T} \sum_{i=1}^T \sum_{j=1}^T \phi_1^{|i-j|} \right). \quad (52)$$

$$\sum_{i=1}^T \sum_{j=1}^T \phi_1^{|i-j|} = T + 2(T-1)\phi_1 + 2(T-2)\phi_1^2 + 2(T-3)\phi_1^3 + \dots + 2(T-(T-1))\phi_1^{T-1} = \quad (53)$$

$$T(1 + 2\phi_1 + 2\phi_1^2 + 2\phi_1^3 + \dots + 2\phi_1^{T-1}) - 2(\phi_1 + 2\phi_1^2 + 3\phi_1^3 + \dots + (T-1)\phi_1^{T-1}). \quad (54)$$

Vi inleder med att summera den andra summan i (54)

$$2(\phi_1 + 2\phi_1^2 + 3\phi_1^3 + \dots + (T-1)\phi_1^{T-1}) = 2\phi_1(1 + 2\phi_1 + 3\phi_1^2 + \dots + (T-1)\phi_1^{T-2}). \quad (55)$$

Vi ser att uttrycket i parantesen ovan är derivatan av

$$1 + \phi_1 + \phi_1^2 + \phi_1^3 + \dots + \phi_1^{T-1}. \quad (56)$$

Detta är en geometrisk summa och vi får:

$$1 + \phi_1 + \phi_1^2 + \phi_1^3 + \dots + \phi_1^{T-1} = \frac{(1 - \phi_1^T)}{1 - \phi_1}. \quad (57)$$

Om vi deriverar båda sidorna får vi

$$1 + 2\phi_1 + 3\phi_1^2 + \dots + (T-1)\phi_1^{T-2} = \frac{(T-1)\phi_1^T + 1 - T\phi_1^{T-1}}{(1 - \phi_1)^2}. \quad (58)$$

Vi multiplicerar sedan in ϕ_1 och får slutligen summan till

$$\frac{(T-1)\phi_1^{T+1} + \phi_1 - T\phi_1^T}{(1 - \phi_1)^2}. \quad (59)$$

Nu får vi enkelt den första summan i (54) till

$$\frac{2T(1 - \phi_1^T)}{1 - \phi_1} - T. \quad (60)$$

Vi tar båda summorna minus varandra

$$\frac{T(1 - \phi_1^T)}{1 - \phi_1} - \frac{(T-1)\phi_1^{T+1} + \phi_1 - T\phi_1^T}{(1 - \phi_1)^2}. \quad (61)$$

Gör om det till gemensam nämnare

$$\frac{T - T\phi_1 - T\phi_1^T + T\phi_1^{T+1} - T\phi_1^{T+1} + \phi_1^{T+1} - \phi_1 + T\phi_1^T}{(1 - \phi_1)^2} = \frac{T - \phi_1(T+1) + \phi_1^{T+1}}{(1 - \phi_1)^2}. \quad (62)$$

Vi gångrar in 2:an och tar uttrycket minus T

$$\frac{2T - 2\phi_1(T + 1) + 2\phi_1^{T+1}}{(1 - \phi_1)^2} - \frac{(T + T\phi_1^2 - 2T\phi_1)}{(1 - \phi_1)^2} = \quad (63)$$

$$\frac{T - T\phi_1^2 - 2\phi_1 + 2\phi_1^{T+1}}{(1 - \phi_1)^2} = \quad (64)$$

$$\frac{T(1 - \phi_1^2)}{(1 - \phi_1)^2} + \frac{2\phi_1(\phi_1^T - 1)}{(1 - \phi_1)^2} = \quad (65)$$

$$\frac{T(1 + \phi_1)}{1 - \phi_1} + \frac{2\phi_1(\phi_1^T - 1)}{(1 - \phi_1)^2}. \quad (66)$$

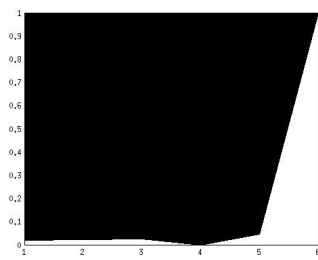
Slutligen får vi väntevärdet av s_L^2 till

$$E[s_L^2] = \frac{\sigma^2}{T - 1} \left(T - \frac{1}{T} \sum_{i=1}^T \sum_{j=1}^T \phi_1^{|i-j|} \right) = \quad (67)$$

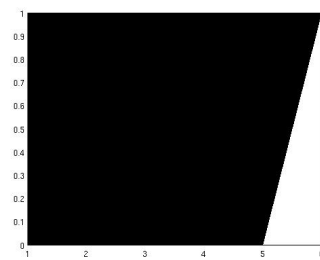
$$\frac{\sigma^2}{T - 1} \left(T - \frac{1 + \phi_1}{1 - \phi_1} - \frac{2\phi_1(\phi_1^T - 1)}{T(1 - \phi_1)^2} \right) \quad (68)$$

vilket alltså inte är en väntevärdesriktig skattning av variansen.

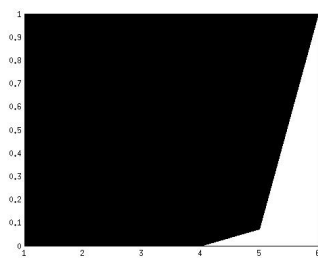
På nästa sida illustreras exempel på fyra informationsfönster av koefficienten $\phi_1 = 0.5$ och $\phi_1 = 0.99$. Prediktionerna är gjorda med MSE-metoden.



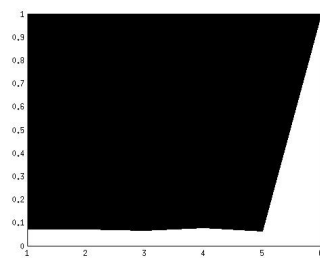
I_p är 0.1201.



I_p är 0.1.

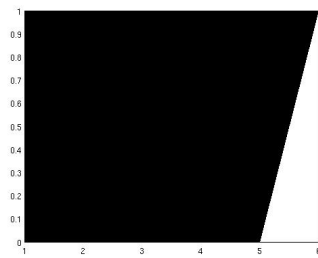


I_p är 0.1142.

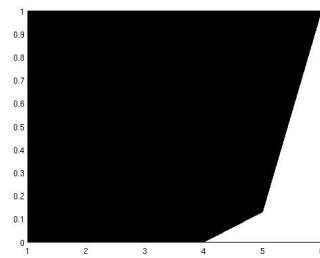


I_p är 0.1624

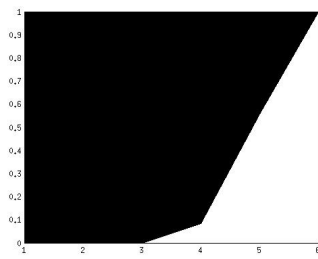
Figur 17: Informationsfönster för modellen $X_t = 0.5X_{t-1} + \epsilon_t$



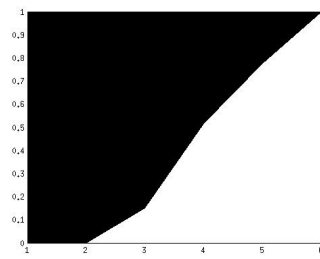
I_p är 0.1.



I_p är 0.1260.



I_p är 0.2282.



I_p är 0.3880.

Figur 18: Informationsfönster för modellen $X_t = 0.99X_{t-1} + \epsilon_t$

Inledningsvis ser vi att informationsfönstrena skiljer sig åt för varje koefficient och att fönstrena är mörkare i jämförelse med informationsfönstrena vi såg i sektionen *Oberoende replikat*. Förklaringen till ett mörkare och ostabilare informationsfönster är troligtvis den positiva korrelationen. Den bidrar dels till att s_L^2 blir mindre än den väntevärdesriktiga variansen som erhålls när vi har oberoende X_t och dels till en ostabil s_L^2 . Ett lägre s_L^2 leder till ett mindre prediktionsindex och därmed till ett mörkare fönster och en ostabil s_L^2 förklarar varför det uppstår skillnader mellan informationsfönstrena för varje koefficient.

Tabellen nedan illustrerar den reducerade $E[s_L^2]$ för koefficienterna $\phi_1 = 0.5$ och $\phi_1 = 0.99$. Vi ser att ju större koefficienten ϕ_1 är, desto mindre blir den förväntade standardavvikelsen. Detta beror på att korrelationen mellan $X_{1991}, \dots, X_{2000}$ blir starkare ju större ϕ_1 .

ϕ_1	$E[s_L^2]$
0.5	$0.8222\sigma^2$
0.99	$0.0359\sigma^2$

5 Diskussion/Slutsats

Ovan har vi presenterat och illustrerat det nya begreppet informationsfönster, där vi har gjort en principiell analys med hjälp av Autoregressiv- samt Moving average process. Vi har studerat två olika sätt, dels när vi simulerar oberoende tidsserier och dels en ensam tidsserie, dvs utan replikat.

Låt oss inleda med det första sättet som är beskrivet i avsnitt *Beskrivning av data och prediktionsmetoder*, dvs sektionen *Oberoende replikat*.

Vi har utgått ifrån två prediktionsmetoder, dels när vi gör en serie prediktioner av X_{2000} genom att betinga väntevärdet av X_{2000} , givet tidigare observerade värden och även konstanta prediktioner.

Informationsfönstrena visar tydligt att den första prediktionsmetoden är en bättre metod, vilket även var väntat. Målet är att få ett högt värde på prediktionsindex, I_l och för att uppnå detta gäller det att minimera MSE (mean square error). Den första prediktionsmetoden använder MMSE, som står för minimum mean square error, där MMSE estimatorn är det betingade väntevärdet av, i det här fallet, X_L , givet kända observerade värden. Teorin säger oss alltså att den första prediktionsmetoden ger upphov till en mer

lämplig tidsserie av prediktioner och därmed är statistikan inte i lika stor behov av revidering.

Vi jämförde de teoretiska prediktionsindexen med den simulerade tidsseriens prediktionsindex. I samtliga modeller såg vi att teorin överensstämde väl med prediktionsindexen som vi fick av dess simulerade tidsserie. Detta säger oss att informationsfönsterna ger den förväntade samt den korrekta bilden av prediktionerna.

Om man är säker på att en dataserie följer en Autoregressiv eller en Moving average process är det i teorin möjligt att hitta en modell med hjälp av informationsfönster. Detta är dock aldrig fallet i praktiken. På så sätt kan vi konstatera att ett informationsfönster inte kan hjälpa oss att hitta en lämplig modell för en given dataserie.

Idag används Autoregressiv samt Moving average flitigt inom inte minst ekonomin. Dessa processer används vid framställning av tidsserier som beskriver predikterade framtida värden. I informationsfönsterna ovan såg vi att olika modeller av moving average och autoregression ger olika bra prediktioner. Exempelvis, informationsfönstret illustrerade att en AR(1)-process med ett högt värde på koefficienten ϕ_1 , ger en rimlig serie prediktioner av det sanna värdet. I avsnittet *Härledning av prediktionsformler för en Autoregressiv process av grad 1* såg vi att

$$I_{L-k} = \phi_1^{2k}$$

med det stationära villkoret, $|\phi_1| < 1$. Vi ser nu tydligt att ju närmare ϕ_1 är 1, desto större blir I_{L-k} , för alla $k > 0$.

Som vi såg i avsnittet *Härledning av prediktionsformler för en Moving average process av grad 1* gäller

$$I_{L-1} = \frac{\theta_1^2}{\theta_1^2 + 1}$$

I formeln ser vi att för $|\theta_1| \leq 1$ ökar I_{L-1} och för $|\theta_1| \geq 1$ minskar I_{L-1} . Detta illustrerar även gränsvärdet nedan.

$$\lim_{\theta_1 \rightarrow \infty} I_{L-1} = \frac{\theta_1^2}{(\theta_1^2 + 1)^2} \rightarrow 0. \quad (69)$$

En väsentlig nackdel med informationsfönster är att den inte kan illustrera kvaliteten på icke-stationära processer såsom ARCH och GARCH, som är viktiga modeller inom bland annat makroekonomi. Men trots detta har vi lyckats visa att informationsfönstret ger en god illustration av prediktionernas kvalite. Man kan nu med gott samvete dra slutsatsen att man har

funnit ett rimligt sätt att undersöka hur bra stationära tidsserier, ger oss det vi förväntar oss av dem, och detta kan göras med hjälp av informationsfönster.

Vi fortsätter med det andra sättet som är beskrivet på sidan 11, dvs sektionen Utan replikat.

Enligt härledningen vi såg på sidorna 36-38 är s_L^2 inte en väntevärdesriktig skattning av den stationära processens varians. Därför ifrågasätter man vad s_L^2 egentligen är en skattning av och om det är rimligt att standardisera med den. För att det finns en positiv korrelation mellan tidpunkterna i den stationära tidsserien, blir s_L^2 alltför ostabil samt bidrar till ett lägre prediktionsindex. Den ostabila s_L^2 leder till ett känsligt informationsfönster, vilket skillnaden mellan informationsfönstrena (se Figur 17 och 18) visar. Detta innebär att det inte är rimligt att standardisera med s_L^2 och att informationsfönstret är därmed, inte tillförlitligt och kan tyvärr inte heller hjälpa oss att hitta en lämplig modell för en dataserie.

Vi ser att Öller och Teterukovskys teori är anpassad efter det första sättet, dvs oberoende replikat, men tyvärr är den verkliga datan inte byggd på det sättet. Undersökningen av *Utan replikat* tillåter oss att konstatera att baserat på verklig data, är informationsfönster inte tillförlitligt och inte säkert ger en god illustration av prediktionernas kvalitet.

6 Referenser

Öller, L. -E., & Teterukovsky, A., (2007). Quantifying the quality of macroeconomic variables. *International Journal of Forecasting*, 217.

Brockwell, P., & Davis, (1996). Introduction to Time Series and Forecasting. *Springer-Verlag GmbH*, 420.

Öller L. -E., & Hansson, K. -G., (2004). Revisions of Swedish national accounts 1980–1998. *Journal of Business Cycle Measurement and Analysis*.

Ross, S. -M., (2010). Introduction to Probability Models. Academic Press, 9th edition. *Academic Press is an imprint of Elsevier*, 731.