Oppgave 1

Ser her på AR(1)- modellen

$$x_t = a + \phi x_{t-1} + w_t,$$

a) Vi antar at modellen er kausal. Hvilket intervall må ϕ være i da? Vi antar også at forventninga ikke nødvendigvis er 0. Hva er forventninga, $\mathbb{E}[x_t]$?

b) Hva er beste lineære m-stegs prediktor for en kausal AR(1)-prosess?

Hint: Iterer differens-ligninga til en AR(1) m steg bakover.

c) Hva er variansen til prediktorfeilen?

d) Finn grensa til BLP og prediktorfeilen når $m \to \infty$.

e) Prøv å forstå hva som menes med beste prediktor og beste lineære prediktor ved å lese starten på kap. 3.4. Anta en observasjon x_1 og vi ønsker å predikere x_{m+1}. Verifiser at

$$\hat{x}_{m+1} = a \sum_{k=0}^{m-1} \phi^k + \phi^m x_1$$

er beste lineære prediktor (BLP) ved bruk av prediksjonsligningene, Property 3.3 side 103.

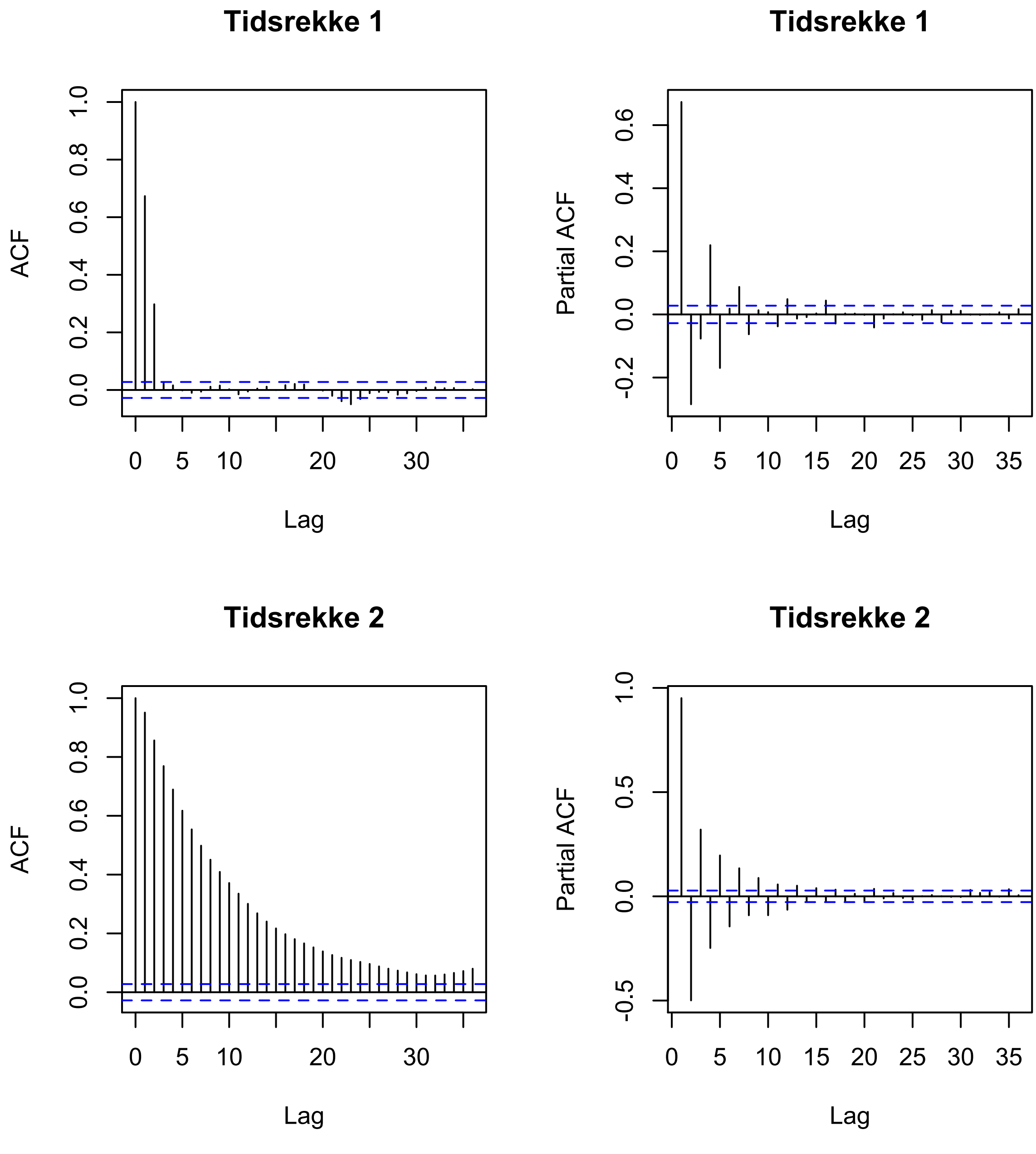
f) La oss se på situasjonen der vi har flere enn et datapunkt, dvs vi har data $x_1, \ldots, x_n$. Vis at $\hat{x}_{n+1} = a + \phi x_n$ er BLP for x_{n+1}.
Dermed avhenger BLP kun av siste observasjon. For en AR(2) vil BLP kun avhenge av de to siste datapunktene.

g) Simuler en Gaussisk AR(1) med $n = 100$. Velg parametre, eksempelvis a=1.1 og $\phi = 0.9$. Beregn BLP for $m = 1, 2, \ldots, M$ (valgfri M). Plott tidsrekka og punkt-prediksjonene.

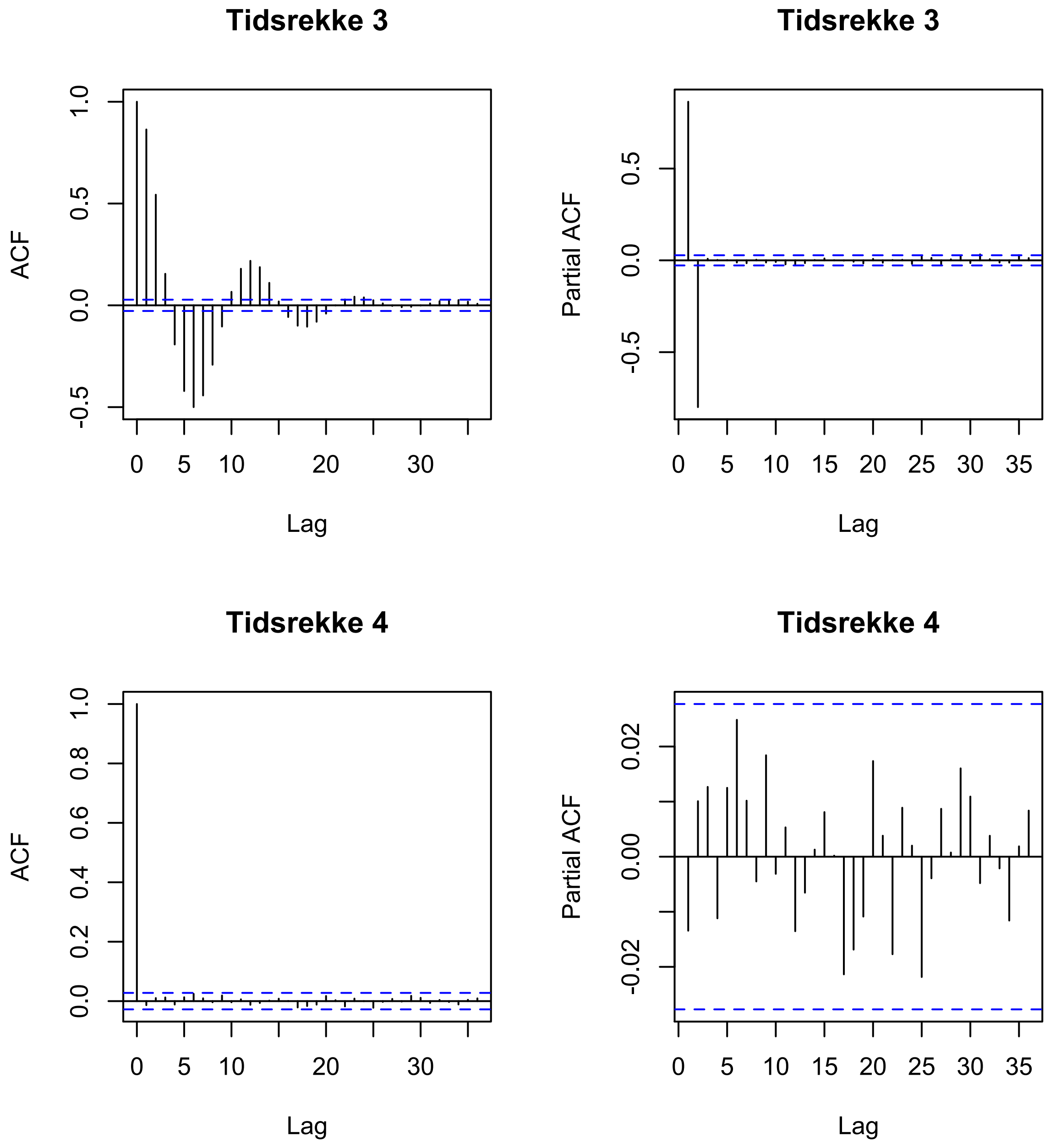
h) Fortsettelse fra c): Lag et 95% prediksjonsintervall for x_{101}. Her kan du bruke at $x_{101}|x_{100}$ er normalfordelt med forventning lik BLP og varians lik prediktorfeilen.

Oppgave 2

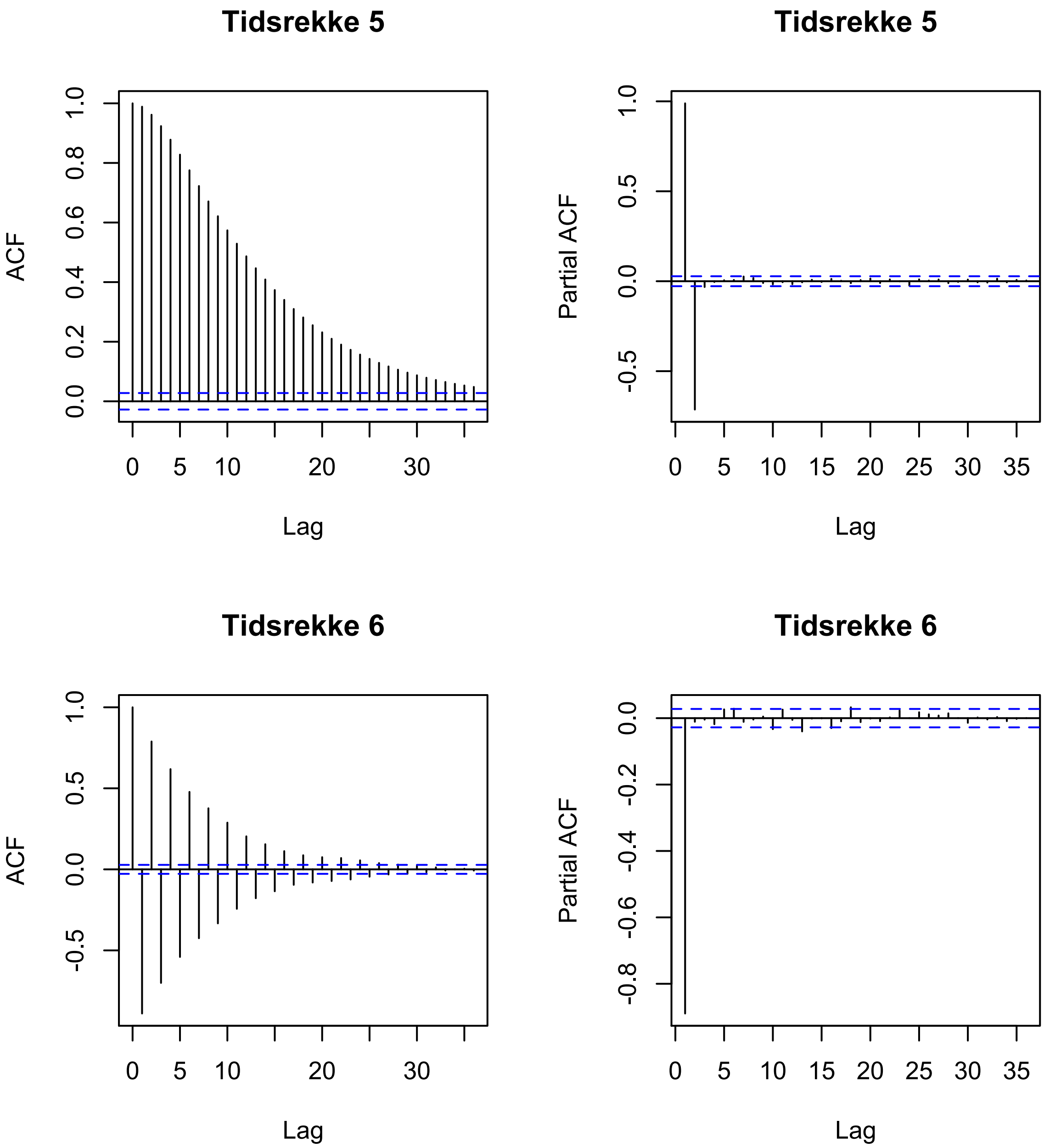
Marit har analysert 6 tidsrekker. Hun er sikker på at disse skal modelleres som ARMA(p, q) der $p \leq 2$ og $q \leq 2$. Basert på figur 1-3, hjelp Marit med å velge p og q.



Figur 1: Sample ACF and sample PACF. Legg merke til at første lag er 0 på ACF og 1 på PACF.



Figur 2: Sample ACF and sample PACF. Legg merke til at første lag er 0 på ACF og 1 på PACF



Figur 3: Sample ACF and sample PACF. Legg merke til at første lag er 0 på ACF og 1 på PACF.

Oppgave 3
La

$$x_t = a + \phi_1 x_{t-1} + \phi_2 x_{t-2} + w_t, \tag{1}$$

der vi velger parametrene ϕ_1 og ϕ_2 slik at x_t er pseudo-periodisk med periode 17. Så vi lar $\phi_1 = 2\cos(2\pi/17)k$ og $\phi_2 = -k^2$, der k bestemmer dempinga i ACF:

$$\rho(h) = ak^h \cos\left(\frac{2\pi}{17}h + b\right), \quad h \geq 0$$

Du velger selv k, f.eks. $k = 9/10$.

a) Sjekk at x_t er kausal. Bestem verdien til a slik at forventninga er $\mu = \mathbb{E}x_t = 5$. Vi setter $\sigma_w^2 = 1$. Hva er variansen til x_t?
Hint: For å beregne varians (og ACF, acvf) kan man sette $a = 0$ (hvorfor?).

b) Anta at vi har data $x_1, \ldots, x_n$ med $n > 1$. BLP $\hat{x}_{n+1}$ for x_{n+1} er

$$\hat{x}_{n+1} = a + \phi_1 x_n + \phi_2 x_{n-1}. \tag{2}$$

Man kan vise at $\hat{x}_{n+1}$ er BLP ved å verifisere at prediksjonsligningene holder (samme framgangsmåte som for AR(1)). Alternativt kan man istedet bruke Durbin-Levinson algoritmen: Anta her at konstant-leddet er rett, sett så a=0 og vis ved Durbin-Levinson. Hva er PACF? LF i eksempel 3.21 i boka.

c) Ser på to-stegs prediksjon. Vi har

$$x_{t+2} = a + \phi_1 x_{t+1} + \phi_2 x_t + w_{t+2}.$$

Det kan vises at 2-stegs BLP er

$$\hat{x}_{n+2} = a + \phi_1 \hat{x}_{n+1} + \phi_2 x_n \tag{3}$$

der $\hat{x}_{n+1}$ er gitt i ligning (2). For $m > 2$ er BLP

$$\hat{x}_{n+m} = a + \phi_1 \hat{x}_{n+m-1} + \phi_2 \hat{x}_{n+m-2}$$

Simuler fra AR(2) modellen. Beregn BLP for $m = 1, 2, \ldots, M$ (valgfri M). Plot tidsrekka og punkt-prediksjonene. Så når $M \to \infty$ vil $\hat{x}_m \to \ldots$?

d) Variansen til prediksjons-feilen (ligning 3.86 i boka) er gitt ved

$$P_{n+m} = \sum_{j=0}^{m-1} \sigma_w^2 \psi_j^2$$

der ψ_j er definert ved

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} \tag{4}$$

I Python/Matlab finner vi ψ_j med `arma2ma`. I Python finnes denne funksjonen i statsmodels.tsa.arima_process. I R finner vi ψ_j med `ARMAtoMA`. Test med en AR(1) med $\phi = 0.9$. Da er $\psi_j = 0.9^j$. Merk at $\psi_0 = 1$ ikke returneres i disse funksjonene.

i) Beregn P_{n+m} og observer at denne går mot var(x_t) når $m \to \infty$.
ii) Sett opp et 95 % prediksjonsintervall, med valgfri steglengde m. .

Oppgave 4 (Bygger på kap. 3.5: Estimering)
I denne oppgaven skal vi modellere tidsrekka 'rec.txt' ved en ARMA(p,q) modell. Velg orden p og q basert på sample ACF og PACF. Estimer parametrene i valgt modell. Merk: Forventning eller intercept er en av parametrene i modellen.

Gjør en residualsanalyse. Residualene er $r_t = x_t - \hat{x}_t$ (e.g, for en AR(1) er $\hat{x}_t = \hat{a} + \hat{\phi} x_{t-1}$). Er residualene normalfordelt? Er sample ACF og PACF av residualene konsistent med en hvit støy?

Basert på valgt modell: Plot $1, 2, \ldots, M$-stegs BLP, og lag et 95% prediksjons-intervall for neste obervasjon.