

Opgave 1.

Beschrijf kort, maar zo volledig mogelijk, een drie-factor variantie-analyse model met interacties.

Opgave 2.

- Wanneer is een tijdreeks $\{X_t\}$ stationair?
- Beschrijf kort twee modellen voor een stationaire tijdreeks.
- Wat is (zijn) het (de) essentiële verschil(len) tussen de (partiële) autocorrelaties van de twee modellen die onder b. zijn beschreven?
- In Figuur 1 op de volgende bladzijde zijn de autocorrelatie- en partiële autocorrelatiefunctie van een tijdreeks gegeven. Welk van de modellen AR, MA en ARMA past op grond van deze plaatjes volgens jou het best bij deze tijdreeks? Schat ook de orde van het proces. Beargumenteer je keuze.

Opgave 3.

Beschrijf de rol van het begrip deviantie bij een statistische analyse met behulp van een gegeneraliseerd lineair model (het gaat hier niet om het geven van een formule).

Opgave 4.

Zij

$$Y_i = f(x_i, \theta) + \varepsilon_i, \quad i = 1, \dots, n$$

een niet-lineair regressiemodel met $Y_1, \dots, Y_n$ onafhankelijke stochastische responsvariabelen, x_i een k -dimensionale vector van verklarende variabelen behorende bij Y_i ($i = 1, \dots, n$), ε_i een stochastische meetfout met verwachting 0 en variantie σ^2 behorende bij Y_i ($i = 1, \dots, n$), θ een onbekende p -dimensionale parametervector en f een niet-lineaire functie van θ .

- Hoe is de kleinste kwadratenschatter $\hat{\theta}$ van θ gedefinieerd?
- Zij $\widehat{\text{Cov}}(\hat{\theta}) = \hat{\sigma}^2(\hat{V}^T \hat{V})^{-1}$ een schatter voor de covariantiematrix van $\hat{\theta}$. Geef een, bij benadering, 95% betrouwbaarheidsinterval voor θ_l , $l = 1, \dots, p$.
- Men heeft de beschikking over twee data sets waarvan men vermoedt dat hetzelfde niet-lineaire model van toepassing is. Het idee is dat de parameterwaarden voor de beide sets hetzelfde zijn, behalve die van θ_1 . Men bepaalt voor de beide data sets betrouwbaarheidsintervallen met bij benadering 95% betrouwbaarheid voor θ_1 . Men vindt voor de

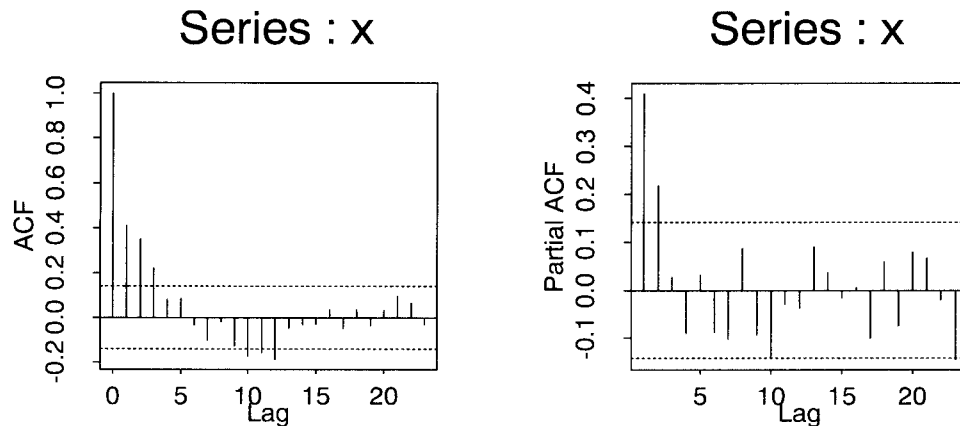


Figure 1: Autocorrelatie en partiële autocorrelatie functie van een tijdreeks.

eerste en de tweede data set (4.32,4.68), respectievelijk, (4.59,5.19) als het betrouwbaarheidsinterval gebaseerd op de klassieke theorie, en (4.32,4.68), respectievelijk, (4.59,5.19) als het bootstrap betrouwbaarheidsinterval. Welke conclusie(s) zou je trekken op grond van deze intervallen? Beargumenteer je antwoord.

Opgave 5.

Hieronder worden enkele praktijksituaties geschetst. Kies voor elke situatie één of meer geschikte modellen. Kies hierbij uit: lineair regressiemodel, variantie-analyse model, niet-lineair model, gegeneraliseerd lineair model, tijdreeksmodel. Beargumenteer je antwoord. Beschrijf het gekozen model zo volledig mogelijk.

- a. In een bepaald industrieel proces komen onregelmatig kleine fouten voor. Om de oorzaak te onderzoeken, wordt het volgende experiment gedaan: verschillende partij- en ruw materiaal worden geselecteerd en elke partij wordt in twee gedeeltes gesplitst. Voor elke partij wordt het ene deel met de standaard methode behandeld, het andere deel met een gemodificeerde behandeling waarin de temperatuur gedurende een tijdje wordt verlaagd. Vóór de behandeling wordt aan elke partij in z'n geheel de zuiverheidsgraad van het materiaal gemeten. Tijdens de behandelingen wordt genoteerd of ja dan nee de kleine fouten optreden.
- b. Twee fabrieken produceren hetzelfde soort touw via een verschillend procédé. Om te onderzoeken of er een systematisch verschil bestaat tussen de sterkte van de twee typen touw, worden van elke fabriek 8 spoelen met 3000 m kabel aselekt geselecteerd en op sterkte onderzocht. Hiertoe worden paren van naast elkaar liggende breuken

(zodat er duplicaties zijn) geïnduceerd op 500 m intervallen over de hele lengte van het touw op een spoel, dit door het touw onder steeds sterkere spanning te zetten. De data zijn dan de gemeten spanningen op de momenten van breken.

- c. Zwaarlijvige mensen die op een dieet worden gezet vertonen een exponentiële afname van hun vetweefsel gedurende de periode van het dieet. Om te onderzoeken hoe lang iemand op een bepaald dieet gezet zou moeten worden, wordt van een proefpersoon gedurende een maand dagelijks het gewichtsverlies gemeten terwijl hij het dieet volgt. De meetfouten zijn klein ten opzichte van het gewicht.

Normering:

1: 5	2.a: 2	3: 4	4.a: 1	5.a: 3
	2.b: 4		4.b: 2	5.b: 3
	2.c: 2		4.c: 4	5.c: 3
	2.d: 3			

Eindcijfer tentamen = totaal/4 + 1