



Tentamen Statistische Modellen

30 januari 1998

Opgave 1.

Beschrijf kort, maar zo volledig mogelijk, een gegeneraliseerd lineair model.

Opgave 2.

Hieronder worden enkele praktijk situaties geschetst. Kies voor elke situatie één of meer geschikte modellen. Kies hierbij uit: lineair regressie model, variantie analyse model, gegeneraliseerd lineair model of een tijdreeks model. Beargumenteer je antwoord. Beschrijf het gekozen model zo volledig mogelijk.

1. Een tandarts doet onderzoek naar het tijdstip dat tanden bij zuigelingen doorbreken. Daartoe vraagt hij een consultatiebureau om gedurende een week bij alle zuigelingen die voor controle komen te noteren of er al tanden aanwezig zijn of niet en hoe oud (in weken) het kind is op het moment van het bezoek. De tandarts is geïnteresseerd in de fractie van het aantal zuigelingen waarbij minstens één tand is doorgebroken in de loop der tijd. Stel een model voor dit probleem op.
2. In een onderzoek naar de verschillen in gemiddelde diastolische bloeddruk tussen de drie economische klassen (arm/modaal/rijk), worden van 250 mannen de bloeddruk opgemeten en wordt er bepaald tot welke sociale klasse de persoon behoort. Omdat gewicht invloed heeft op de bloeddruk wordt deze ook opgemeten. Stel een model op om de invloed van de sociale klasse op de bloeddruk te bepalen.
3. Een fabrikant van kattenvoer wil nagaan of het toevoegen van bepaalde smaakstoffen de eetlust van katten vergroot. Hij heeft keuze uit de smaakstoffen A, B, C en D. De onderzoeksgroep van 40 katten wordt standaard hoeveelheden van kattenvoer met de smaken A, B, C en D aangeboden. De volgorde van de smaken wordt door het toeval bepaald. Elke kat krijgt dus vier keer een schoteltje kattenvoer. Elke keer wordt genoteerd hoelang de kat nodig heeft om zijn schoteltje leeg te eten. Stel een model op om te onderzoeken welke smaakstof het meest in trek is bij de katten.

Opgave 3. De Wailuku Sugar Company doet onderzoek naar de invloed van drie verschillende soorten kunstmest op drie soorten rietsuiker. Elke combinatie is vier keer uit geprobeerd op stukken grond van gelijke grootte. De opbrengst voor alle 36 stukjes grond wordt genoteerd. Het doel is een zo hoog mogelijke opbrengst te krijgen.

De waarnemingen worden geanalyseerd met het software pakket S-plus. De uitvoer is hieronder weergegeven. Enkele getallen hebben we door vraagtekens vervangen.

Source	sum of squares	DF	Mean Square	Test Statistic (T)	Sig. of T (p-value)
Main Effects					
suikersoort	296.674	2	148.337	?	0.0366
mestsoort	50.491	?	?	0.638	0.5360
Interaction					
suiker-mest	613.538	4	153.384	3.877	0.0129
Residual	1068.107	27	39.560		

- Beschrijf een variantie analyse model voor dit probleem.
- Welke getallen horen op de plaatsen van de vraagtekens te staan? Beargumenteer je antwoord.
- Beschrijf een toets om na te gaan of een additief model voldoet. Geef de nulhypothese, de toetsingsgrootte, de verdeling van de toetsingsgrootte onder de nulhypothese en voor welke waarden van de toetsingsgrootte we de toets moeten verwerpen. Voer de toets uit met onbetrouwbaarheid $\alpha = 0.05$.
- Toets of de hoofdeffecten in het model moeten worden opgenomen. Geef opnieuw de nulhypothese, de toetsingsgrootte, de verdelingen van de toetsingsgrootte onder de nulhypothese en voor welke waarden van de toetsingsgrootte we de toetsen verwerpen. Voer de toetsen uit met onbetrouwbaarheid $\alpha = 0.05$.
- Geef je conclusie.

Opgave 4.

Beschrijf een AR(p) proces. Leid een recursieformule voor de autocovarianties af.

Opgave 5.

Hieronder zijn plaatjes van een van de autocorrelaties en partiële autocovarianties van een



MA of een AR proces weergegeven. Welk model past volgens jou het beste, een AR of een MA? Schat ook de orde van het proces. Beargumenteer je keuze.

Series : tijdreeks

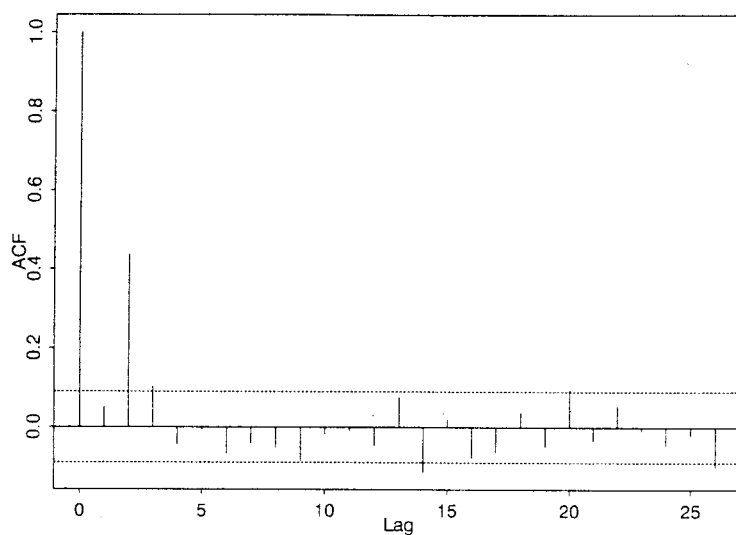


Figure 1: autocorrelaties

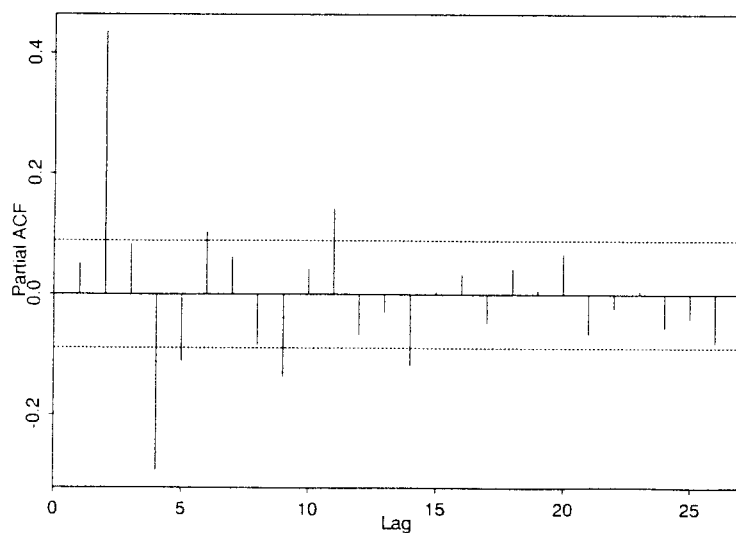


Figure 2: partiële autocorrelaties

Normering:

1: 4	2.1: 3	3a: 2	4: 5
	2.2: 3	3b: 2	5: 5
	2.3: 3	3c: 3	
		3d: 3	
		3e: 3	

Eindcijfer tentamen = totaal/4 + 1