

Opgave 1.

- a. Beschrijf kort, maar zo volledig mogelijk een niet-lineair model.
- b. 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 90% betrouwbaarheid voor θ_1 . Men vindt voor de eerste en de tweede data set (0.33,0.67), respectievelijk, (0.59,0.92) als het betrouwbaarheidsinterval gebaseerd op de klassieke theorie, en (0.41,0.59), respectievelijk, (0.65,0.86) als het bootstrap betrouwbaarheidsinterval. Welke conclusie(s) zou je trekken op grond van deze intervallen? Beargumenteer je antwoord.

Opgave 2.

Een gezondheidwetenschapper heeft de beschikking over het cholesterolgehalte in het bloed van een aantal mannen in drie verschillende sociaal-economische klassen gelabeld (van hoog naar laag) I, II en III. Ze wil onderzoeken of er een significant verschil bestaat tussen de verwachte cholesterolgehalten van de drie groepen. Groep I bevatte 5 personen, groep II 10 en groep III 6.

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

Source	Sum of squares	DF	Mean square	Test statistic (T)	Sign. of T (p-value)
Between groups	1202.5	2	601.2	?	0.88
Within groups	85750.5	?	4763.9		
Total	86953.0	20			

- a. Beschrijf een variantie analyse model voor dit probleem.
- b. Beschrijf een toets om na te gaan of de verwachte cholesterol gehalten van de groepen

verschillen. Vermeld hiertoe in elk geval de nulhypothese, de toetsingsgrootheid, de verdeling van de toetsingsgrootheid onder de nulhypothese en voor welke waarden van de toetsingsgrootheid de nulhypothese verworpen dient te worden.

c. Welke getallen horen op de plaatsen van de vraagtekens te staan? Beargumenteer je antwoord.

d. Voer de toets uit met onbetrouwbaarheid $\alpha = 0.05$ en geef je conclusie.

Opgave 3.

Beschrijf hoe twee geneste gegeneraliseerde lineaire modellen met elkaar kunnen worden vergeleken.

Opgave 4.

Veronderstel dat aan je gevraagd wordt een tijdreeksanalyse uit te voeren op een gegeven data set $X_1, \dots, X_n$. Beschrijf kort, maar zo volledig mogelijk, de stappen die je gaat uitvoeren om aan deze opdracht te voldoen.

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. Vermeld hierbij in elk geval wat de respons- en verklarende variabelen zijn of welke variabelen een tijdreeks vormen.

- a. Elk jaar op de dag voor koninginndag bestelt een hamburgerbakker het aantal hamburgers dat hij de volgende dag denkt nodig te hebben op grond van het aantal hamburgers dat hij het jaar daarvoor heeft verkocht. De afgelopen koninginndag was echter totaal verregend en de man vraagt zich af of hij zijn bestelling van volgend jaar op de geringe verkoop van dit jaar zou moeten baseren. Hij vraagt zich af hoe hij dit zou kunnen onderzoeken. De bestellingen en verkoopcijfers van de afgelopen 14 koninginndagen zijn opgeslagen in zijn computer.
- b. 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 deze het dieet volgt. De meetfouten zijn klein ten opzichte van het gewicht.
- c. Een verzekeringsmaatschappij wil de afhankelijkheid onderzoeken van het aantal gerapporteerde stormschades aan woningen van variabelen als ouderdom van de woning, inkomen van de bewoners, type buurt, raamoppervlakte, dakoppervlakte, hoogte van de woning.

Normering:

1.a: 4	2.a: 2	3: 3	4: 7	5.a: 3
1.b: 4	2.b: 4			5.b: 3
	2.c: 2			5.c: 3
	2.d: 1			

Eindcijfer tentamen = totaal/4 + 1