

Ruimtelijke Analyse en GIS 1 (AW2-2006); EINDTOETS 2004
Bacheloropleiding Aardwetenschappen

8 april 2004 Tijd: 09.00 - 12.00.

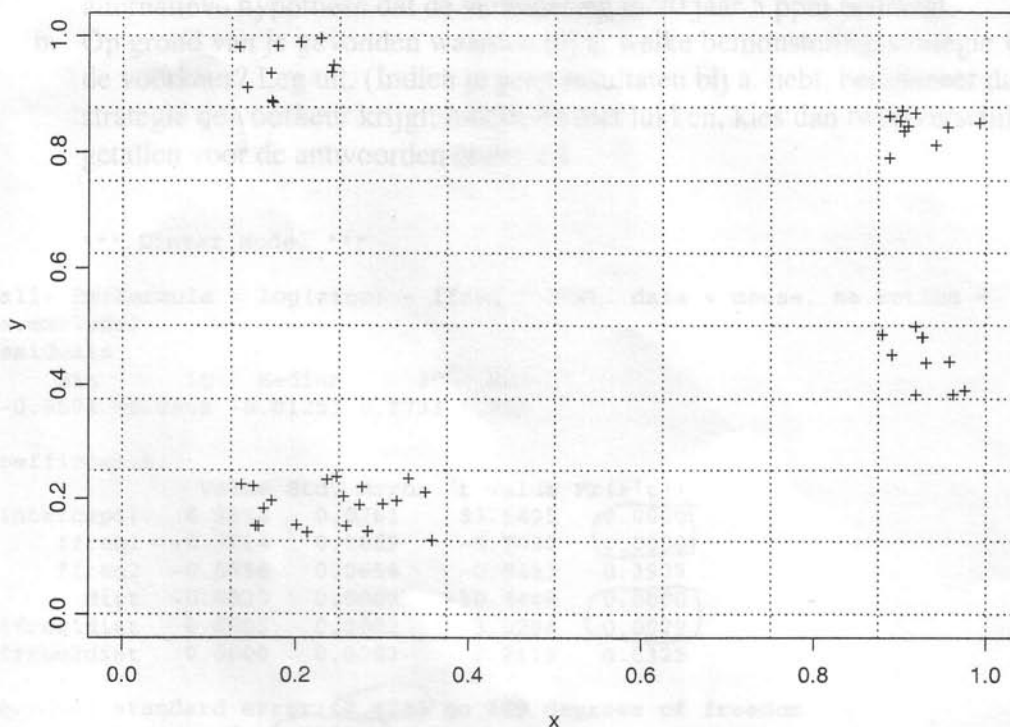
Nota bene: U moet alle vragen beantwoorden.

Tekeningen en grafieken kunnen uw antwoorden verhelderen.

1. Beschrijf in het kort de basis principes van de drie belangrijkste typen rekenoperaties zoals bij Map Algebra/PCRaster worden gebruikt. Nb. U hoeft niet alle rekenfuncties afzonderlijk te beschrijven – alleen voorbeelden van de verschillende basis typen!
2. Leg uit, met gebruik van een eenvoudig flowchart, hoe U met PCRaster:
 - a) een kaart van West Nederland zou maken waarop de plaatsen en de omvang van zwakke veendijken staan.
 - b) bepaalt waar, in een heuvelachtige maar toch gedeeltelijk bewoond gebied, langs een spoorlijn waarlangs regelmatig gifwagens rijden, de risicozones zijn voor verontreiniging in geval van een treinongeluk, als een functie van bevolkingsdichtheid, windrichting, hydrologische toestand en afstand tot de spoorlijn.

Geef in beide gevallen een lijst van de benodigde basisgegevens op.

3. Beschrijf het volgende:
 - a) vier factoren die van invloed zijn op de kwaliteit van de ruimtelijke informatie die uit digitale hoogte modellen kan worden afgeleid (b.v. helling).
 - b) Perkal's epsilon error band en het "punt in polygoon" probleem



4. Bovenstaand bemonsteringsschema is een zogenaamd hybride bemonsteringsschema: het bevat elementen van twee elementaire strategieën.
- Welke twee elementaire strategieën herken je hierin?
 - Geef een voor- en een nadeel op van de hierboven geschetste strategie.
 - Zal de standaardfout voor het gemiddelde voor deze strategie hoger dan, lager dan of gelijk zijn aan de standaardfout verkregen bij volledig aselekt bemonsteren met eenzelfde steekproefgrootte?

5. In 1984 is op een 16-tal aselekt gekozen locaties langs snelwegen, op een afstand van 15 meter van de weg, loodconcentratie in de bodem gemeten. De standaarddeviatie van de metingen in 1984 bedroeg 20 ppm. Om te kijken of er over een 20-tal jaren significante verschillen voordoen, wordt voorgesteld om dit jaar opnieuw loodconcentratie te gaan meten. Er circuleren twee voorstellen: (A) zoek dezelfde 16 locaties op (gepaarde waarnemingen), en meet daar opnieuw, (B) zoek 16 opnieuw, aselekt gekozen locaties (op 15 meter van een snelweg) op, en bemonster daar (onafhankelijke waarnemingen). Uit eerder onderzoek mag worden verondersteld dat de standaarddeviatie van veranderingen in (op identieke locaties gemeten) loodconcentratie over 20 jaar 5 ppm bedraagt.

- bereken, voor beide voorstellen A en B, het onderscheidend vermogen ($1-\beta$) bij de alternatieve hypothese dat de verandering in 20 jaar 5 ppm bedraagt.
- Op grond van je gevonden waarden bij a: welke bemonsteringsstrategie verdient de voorkeur? Leg uit. (Indien je geen resultaten bij a. hebt, beredeneer dan welke strategie de voorkeur krijgt; mocht dit niet lukken, kies dan twee verschillende getallen voor de antwoorden onder a.)

*** Linear Model ***

Call: `lm(formula = log(zinc) ~ ffreq * dist, data = meuse, na.action = na.exclude)`

Residuals:

| Min | 1Q | Median | 3Q | Max |
|---------|---------|----------|--------|-------|
| -0.9574 | -0.2968 | -0.01257 | 0.2733 | 1.789 |

Coefficients:

| | Value | Std. Error | t value | Pr(> t) |
|-------------|---------|------------|----------|----------|
| (Intercept) | 6.3594 | 0.0761 | 83.5495 | 0.0000 |
| ffreq1 | -0.3814 | 0.0669 | -5.7030 | 0.0000 |
| ffreq2 | -0.0556 | 0.0656 | -0.8481 | 0.3977 |
| dist | -0.0020 | 0.0002 | -10.3449 | 0.0000 |
| ffreq1dist | 0.0005 | 0.0002 | 3.0286 | 0.0029 |
| ffreq2dist | 0.0000 | 0.0002 | 0.2119 | 0.8325 |

Residual standard error: 0.4286 on 149 degrees of freedom

Multiple R-Squared: 0.6589

F-statistic: 57.56 on 5 and 149 degrees of freedom, the p-value is 0

Analysis of Variance Table

Response: log(zinc)

Terms added sequentially (first to last)

| | Df | Sum of Sq | Mean Sq | F Value | Pr(F) |
|------------|-----|-----------|----------|----------|------------|
| ffreq | 2 | 20.30984 | 10.15492 | 55.2702 | 0.00000000 |
| dist | 1 | 30.86334 | 30.86334 | 167.9800 | 0.00000000 |
| ffreq:dist | 2 | 1.70200 | 0.85100 | 4.6317 | 0.01118197 |
| Residuals | 149 | 27.37611 | 0.18373 | | |

6. Bovenstaande uitvoer werd gevonden door in de maas data set log(zinc) lineair te relateren aan overstromingsfrequentie (ffreq, een factor of nominale variabele) en afstand tot de rivier (dist, een numerieke variabele).

- wat is, in dit model, de gemiddelde verandering in log(zinc) wanneer afstand tot de rivier een eenheid toeneemt? Is dit een toe- of afname?
- Is de interactie tussen overstromingsfrequentie en afstand significant? Geef aan waar je dit afleest
- Is het een goed idee om variabelen behorende bij niet-significante coëfficiënten in bovenstaand model te verwijderen? Licht kort toe.
- Leg in woorden uit hoe we de interactie tussen ffreq en dist moeten begrijpen, zo mogelijk in termen van een achterliggend fysisch proces.