INSTITUUT VOOR CULTUURTECHNIEK EN WATERHUISHOUDING

NOTA 277, d.d. 18 november 1964

Eenvoudige proefschema's ten behoeve van de Sinderhoeve met de bijbehorende analyses

(Vervolg NOTA 263)

I. G. M. Brück

Nota's van het Instituut zijn in principe interne communicatiemiddelen, dus geen officiële publikaties.
Hun inhoud varieert sterk en kan zowel betrekking hebben op een eenvoudige weergave van cijferreeksen, als op een concluderende discussie van onderzoeksresultaten. In de meeste gevallen zullen de conclusies echter van voorlopige aard zijn omdat het onderzoek nog niet is afgesloten.
Aan gebruikers buiten het Instituut wordt verzocht ze niet in publikaties te vermelden.
Bepaalde nota's komen niet voor verspreiding buiten het Instituut in aanmerking.

De twee-klassen indeling

Bij het gewarde blokkenschema wordt op andere wijze dan bij de één-klasse indeling getracht de nauwkeurigheid te verhogen. Bij veldproeven is bekend, dat de vruchtbaarheid van plaats tot plaats verschillen vertoont. Daarom wordt het veld allereerst in blokken verdeeld, daarna de blokken weer in zoveel veldjes als er behandelingen moeten worden vergeleken.

N

	Blok 1	Blok 2	Blok 3	
	a	b	c	
W	b	d	a	O
	c	a	b	
	d	c	d	

Z

Indien elke behandeling even vaak gerepresenteerd is in elk blok, wordt het verschil tussen behandelingsgemiddelden niet verstoord door het verschil tussen de blokken en zijn blokken en behandelingen zogenaamd orthogonaal.

De orthogonaliteit heeft tot gevolg, dat de waarnemingscijfers statistisch gemakkelijk te verwerken zijn.

Is in een bepaalde richting in het veld bijvoorbeeld O → W een vruchtbaarheidsverloop te constateren, dan moet de richting van de blokken loodrecht, dus N → Z, staan op de richting van het vruchtbaarheidsverloop in het veld. Het verschil in vruchtbaarheid binnen de blokken is nu dus gemiddeld kleiner dan het verschil in vruchtbaarheid tussen veldjes in verschillende blokken. Het model van de waarnemingsvector $\underline{x}$ is:

$$\underline{x} = \underline{x}_N + \underline{x}_{Blok^*} + \underline{x}_{Beh^*} + \underline{x}_T$$

Indien in het algemeen een indeling van m blokken, elk van n behandelingen wordt gebruikt, dan is de verdeling van de dimensies:

$$n \times m = 1 + (m-1) + (n-1) + (m-1)(n-1)$$

De ruimten N, Blok*, Beh* en T zijn onderling orthogonale ruimten van respectievelijk niveau, zuiver blok-,zuiver behandelingseffect en toeval. De toevalsruimte T bevat de vruchtbaarheidsverschillen voor zover deze zich binnen de blokken manifesteert.

Daar projectie van de waarnemingsvector $\underline{x}$ op elk van deze ruimten verkrijgt men de kwadraten van de lengte van de projecties, die het schema voor de variantie-analyse vormen. Door de onderlinge orthogonaliteit is de stelling van Pythagoras van toepassing en is:

$$\underline{x}^2 = \underline{x}^2_N + \underline{x}^2_{Blok^*} + \underline{x}^2_{Beh^*} + \underline{x}^2_T \qquad (1)$$

Voorbeeld:

Gegevens van drainafvoer in mm per etmaal. 19 Percelen gelegen op Schouwen, per perceel 3 drains.

Doel:

Het nagaan of op alle percelen dezelfde afvoer optreedt. Gegevens verstrekt door B. VAN DER WEERD te Goes, 1964.

De blokindeling is hier een indeling naar percelen, binnen de percelen een indeling naar drains. Het verschil tussen perceelsgemiddelden wordt niet beïnvloed door het verschil tussen de drains en er geldt:

$$\underline{x} = \underline{x}_N + \underline{x}_{Perc^*} + \underline{x}_{Drain^*} + \underline{x}_T$$

	Drain 1	Drain 2	Drain 3	Drain 1+2+3
Perceel 1	- 0.96	- 0.23	- 0.21	- 1.40
2	- 0.27	- 0.36	- 0.50	- 1.13
3	- 0.92	- 0.72	- 1.40	- 3.04
4	- 0.70	- 1.70	- 1.70	- 4.10
5	- 1.00	- 0.22	- 0.25	- 1.47
6	+ 0.09	- 0.21	+ 0.29	+ 0.17
7	- 0.68	- 0.96	- 0.59	- 2.23
8	- 0.80	- 0.72	- 0.72	- 2.24
9	+ 0.25	+ 0.39	+ 0.39	+ 1.03
10	- 0.89	- 0.68	- 1.52	- 3.09
11	- 0.80	- 0.55	- 1.30	- 2.65
12	- 1.15	- 0.96	- 0.70	- 2.81
13	- 0.51	- 1.52	- 0.52	- 2.55
14	- 0.96	- 0.72	- 0.44	- 2.12
15	- 1.40	- 0.74	- 0.35	- 2.49
16	- 0.77	- 0.59	- 0.68	- 2.04
17	- 0.11	- 0.82	- 0.15	- 1.08
18	- 0.44	- 0.02	- 0.01	- 0.47
19	+ 0.18	+ 0.24	0.05	+ 0.47
$\sum$ Percelen	-11.84	- 11.09	-10.31	-33.24

Totaal 3 x 19 = 57 gegevens

Van de oorspronkelijke gegevens werd een constant getal afgetrokken. (Dit om het rekenwerk zonder machine te vergemakkelijken) Voor de variantie maakt dit echter géén verschil.

Projectie van $\underline{x}$ op de respectievelijke ruimten:

Niveau:

$$x_N = \frac{-33.24}{57}\begin{pmatrix}1 & 1 & 1\\ \vdots & \vdots & \vdots\\ 1 & 1 & 1\end{pmatrix} \quad \text{en} \quad x_N^2 = \left(\frac{33.24}{57}\right)^2 \times 57 = 19.384$$

<u>Perceel*</u>:

$$x_{Perc^*}: \frac{-1.40}{3}\begin{pmatrix}1&1&1\\0&0&0\\\vdots&\vdots&\vdots\\0&0&0\end{pmatrix} + \frac{-1.13}{3}\begin{pmatrix}0&0&0\\1&1&1\\0&0&0\\\vdots&\vdots&\vdots\\0&0&0\end{pmatrix} + \ldots + \frac{0.47}{3}\begin{pmatrix}0&0&0\\\vdots&\vdots&\vdots\\0&0&0\\1&1&1\end{pmatrix}$$

en

$$x^2_{Perc} = (\frac{1.40}{3})^2 \times 3 + (\frac{1.13}{3})^2 \times 3 + \ldots + (\frac{0.47}{3})^2 \times 3 = 29.988$$

maar

$$x^2_{Perc^*} = x^2_{Perc} - x^2_N = 29.988 - 19.384 = 10.604$$

<u>Drain*</u>:

$$x_{Drain}: \frac{-11.84}{19}\begin{pmatrix}1&0&0\\\vdots&\vdots&\vdots\\1&0&0\end{pmatrix} + \frac{-11.09}{19}\begin{pmatrix}0&1&0\\\vdots&\vdots&\vdots\\0&1&0\end{pmatrix} + \frac{-10.31}{19}\begin{pmatrix}0&0&1\\\vdots&\vdots&\vdots\\0&0&1\end{pmatrix}$$

en

$$x^2_{Drain}: (\frac{11.84}{19})^2 \times 19 + (\frac{11.09}{19})^2 \times 19 + (\frac{10.31}{19})^2 \times 19 = 19.446$$

maar

$$x^2_{Drain^*} = x^2_{Drain} - x^2_N = 19.446 - 19.384 = 0.062$$

<u>Totaal</u>:

$$0.96^2 + 0.27^2 + \ldots + 0.05^2 = 34.522$$

Voor de berekening van de netto kwadraatsommen of ook wel kwadraatsommen gezuiverd van niveau voor percelen respectievelijk drains heeft men dus nodig de sommen per perceel respectievelijk drain en de totale som.

Na aftrekking van de kwadraatsommen verkrijgt men de som van kwadraten

voor toeval (analoog aan (1)):

$$\begin{aligned} \underline{x}_T^2 &= \underline{x}^2 - \underline{x}_N^2 - \underline{x}_{Perc}^2{}^{*} - \underline{x}_{Drain}^2{}^{*} \\ &= 34.522 - 19.384 - 10.604 - 0.062 \\ &= 4.472 \end{aligned} \qquad (2)$$

en de verdeling van de dimensies:

$$36 = 57 - 1 - 18 - 2 \qquad (3)$$

waaruit een zuivere schatting van σ^2 volgt:

$$S\ (\sigma^2) = \frac{x_T^2}{\dim T} = \frac{4.472}{36} = 0.124 = s^2$$

en

$$s = \sqrt{0.124} = 0.352$$

Bij proeven met 2 of meer indelingen is sprake van 2 of meer effecten.

Bij een gewarde blokkenproef is dus een blokeffect en een behandelingseffect, in dit voorbeeld een perceelseffect en een draineffect.

Om deze te kunnen toetsen stelt men de nulhypothese:

H_o : 'er is géén verschil tussen de percelen'

en indien H_o geldt, wordt de som van kwadraten voor zuiver perceelseffect opgevat als een som van kwadraten voor toeval met 18 dimensies.

Past men de <u>F</u> toets toe, dan is volgens (2) en (3):

$$F_{36}^{18} = \frac{10.604/18}{4.472/36} = \frac{0.589}{0.124} = 4.75$$

met een bijbehorende overschrijdingskans $P < 0.1\%$

Voor toetsing van het draineffect luidt de nulhypothese:

H_o : 'er is géén verschil tussen de drains'

en indien H_o geldt, wordt de som van kwadraten voor zuiver draineffect opgevat als een som van kwadraten voor toeval met 2 dimensies en weer uit (2) en (3):

$$F^2_{36} = \frac{0.062/\ 2}{4.472/36} = \frac{0.031}{0.124} = 0.25 \qquad \text{met} \qquad P > 25\%$$

Een overzicht van de variantie-analyse luidt:

Bron van variantie	Som van kwadraten	Dimensie	S (σ^2)	F	P
Niveau	19.384	1			
Tussen percelen	10.604	18	0.589	4.75	< 0,1%
Tussen drains	0.062	2	0.031	0.25	> 25 %
Toeval	4.472	36	0.124		
Totaal	34.522	57			

Conclusie

Tussen de drains blijkt géén significant verschil te bestaan, daarentegen tussen de percelen kan een duidelijk significant verschil worden aangetoond met een onbetrouwbaarheid < 0,1%. Dat wil zeggen dat de uitspraak gedaan kan worden, dat de afvoeren van perceel tot perceel onderling sterk zullen verschillen met een kleine kans, namelijk < 0,1%, dat deze verschillen op toeval berusten.

Het Latijnse vierkant

Een uitbreiding op de twee-klasse indeling naar een drie-klassen indeling is onder andere een proefschema met 3 orthogonale ruimten bijvoorbeeld van kolom, rij en behandeling: een Latijns vierkant. Hier hebben dus rij- èn kolomeffect géén invloed op de verschillen tussen behandelingsgemiddelden.

Een schema voor een 4 x 4 Latijns vierkant en behandelingen a, b, c en d:

Kolom

	1	2	3	4
1	a	b	c	d
Rij 2	d	a	b	c
3	c	d	a	b
4	b	c	d	a

Het model van de waarnemingsvector $\underline{x}$ is:

$$\underline{x} = \underline{x}_N + \underline{x}_{Kol^*} + \underline{x}_{Rij^*} + \underline{x}_{Beh^*} + \underline{x}_T \qquad (4)$$

waarin N, Kol^*, Rij^* en Beh^* en T onderling orthogonale ruimten zijn van respectievelijk niveau, zuiver kol-, rij-, behandelingseffect en toeval.

Voor een n x n vierkant wordt de verdeling van de dimensies behorende bij (4):

$$n \times n = 1 + (n-1) + (n-1) + (n-1) + (n-1)\,(n-2)$$

Projectie van $\underline{x}$ op elk van de ruimten levert de variantie-analyse:

$$\underline{x}^2 = \underline{x}_N^2 + \underline{x}_{Kol^*}^2 + \underline{x}_{Rij^*}^2 + \underline{x}_{Beh^*}^2 + \underline{x}_T^2$$

en voor de analyse zijn weer nodig:

1. Som van kwadraten van alle waarnemingen
2. Sommen per rij, kolom en behandeling
3. Totale som
4. Kwadraten van de sommen onder 2. en 3.

Door aftrekking verkrijgt men de som van kwadraten voor toeval:

$$\underline{x}_T^2 = \underline{x}^2 - \underline{x}_N^2 - \underline{x}_{Kol^*}^2 - x_{Rij^*}^2 - \underline{x}_{Beh^*}^2$$

Bij een 4 x 4 Latijns vierkant zijn de basis vectoren van de verschillende ruimten:

<u>Niveau</u>:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

waarmee wordt aangegeven dat alle veldjes gelijke opbrengsten zouden geven, als

er geen behandelings-, noch kolom-, nog rij-effect zou bestaan.

Een basis voor de ruimte van de kolommen:

$$\begin{pmatrix} 1&0&0&0\\ 1&0&0&0\\ 1&0&0&0\\ 1&0&0&0 \end{pmatrix}, \begin{pmatrix} 0&1&0&0\\ 0&1&0&0\\ 0&1&0&0\\ 0&1&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&1&0\\ 0&0&1&0\\ 0&0&1&0\\ 0&0&1&0 \end{pmatrix}, \begin{pmatrix} 0&0&0&1\\ 0&0&0&1\\ 0&0&0&1\\ 0&0&0&1 \end{pmatrix}$$

waarmee wordt aangegeven dat in een kolom alle veldjes gelijke opbrengsten zouden geven als er geen rij-, noch behandelingseffect zou bestaan.

De ruimte van zuiver kolomeffect staat loodrecht op de ruimte van het niveau in de ruimte van kolommen, dus

$$\left[p \begin{pmatrix} 1&0&0&0\\ 1&0&0&0\\ 1&0&0&0\\ 1&0&0&0 \end{pmatrix} + q \begin{pmatrix} 0&1&0&0\\ 0&1&0&0\\ 0&1&0&0\\ 0&1&0&0 \end{pmatrix} + r \begin{pmatrix} 0&0&1&0\\ 0&0&1&0\\ 0&0&1&0\\ 0&0&1&0 \end{pmatrix} + s \begin{pmatrix} 0&0&0&1\\ 0&0&0&1\\ 0&0&0&1\\ 0&0&0&1 \end{pmatrix} \right] \times \begin{pmatrix} 1&1&1&1\\ 1&1&1&1\\ 1&1&1&1\\ 1&1&1&1 \end{pmatrix} = 0$$

en dus

$$4p + 4q + 4r + 4s = 0$$
$$p + q + r + s = 0$$

dus

$$s = -p - q - r$$

De basis van zuiver kolomeffect: $\underline{\text{Kol}}^{*}$ ziet er nu als volgt uit:

$$\begin{pmatrix} 1&0&0&-1\\ 1&0&0&-1\\ 1&0&0&-1\\ 1&0&0&-1 \end{pmatrix}, \begin{pmatrix} 0&1&0&-1\\ 0&1&0&-1\\ 0&1&0&-1\\ 0&1&0&-1 \end{pmatrix}, \begin{pmatrix} 0&0&1&-1\\ 0&0&1&-1\\ 0&0&1&-1\\ 0&0&1&-1 \end{pmatrix} \text{ met 3 dimensies}$$

op analoge wijze:

<u>Rij*</u>:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{pmatrix}$$

<u>Beh*</u>:

$$\underset{B_1}{\begin{pmatrix} 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}}, \underset{B_2}{\begin{pmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \end{pmatrix}}, \underset{B_3}{\begin{pmatrix} 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}}$$

Indien het rij-effect van kolom tot kolom verschillen vertoont, spreekt men van interactie tussen rijen en kolommen.

De basisvectoren van deze interactie ruimte van zuiver rij x kolomeffect verkrijgt men door vermenigvuldiging van de basisvectoren van Rij* met de basisvectoren van Kol*:

Rij* x Kol*:

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} = y_1, \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} = y_2, \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} = y_3,$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} = y_4, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} = y_5, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} = y_6,$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ -1 & 0 & 0 & 1 \end{pmatrix} = y_7, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 1 \end{pmatrix} = y_8, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} = y_9$$

Nu blijkt, dat de basisvectoren van de zuivere behandelingseffect-ruimte Beh* een lineaire combinatie zijn van de basisvectoren van Rij* x Kol* en wel:

$$B_1 = y_1 - y_4 + y_5 - y_8 + y_9$$

$$B_2 = y_2 - y_4 - y_8 + y_6$$

$$B_3 = y_7 - y_4 - y_8 + y_3$$

en de zuivere behandelingseffect-ruimte is een deelruimte van de interactie-ruimte van zuiver rij x kolomeffect, of wel de ruimte van zuiver behandelingseffect is gestrengeld met de interactie-ruimte van zuiver rij x kolomeffect.

Belangrijk is, dat bij een Latijns vierkant elke effectruimte gestengeld is met de interactie-ruimte van de beide andere effecten en alléén indien géén interactie wordt verondersteld, zijn de verschillende invloeden te schatten en de toetsen, immers de restruimte van de interactie-ruimte wordt dan opgevat als toevalsruimte T.

Deze ruimte T levert dan een S (σ^2), zodat met de F toets het kolom-, rij- en behandelingseffect kan worden getoetst.

Het nut van Latijnse vierkanten is echter zeer beperkt:

Bij vierkanten kleiner dan 4 x 4 is de dimensie van de toevalsruimte al zo klein, dat de schatting van de variantie onvoldoende nauwkeurig is om andere dan zeer grote invloeden overtuigend te doen uitkomen. Vierkanten groter dan 6 x 6 worden daarentegen slechts zelden toegepast, omdat het aantal herhalingen groter is dan mogelijk of nodig.

Bovendien worden rijen en kolommen zo langgerekt, dat interactie tussen rijen en kolommen niet is uitgesloten.

Voorbeeld:

Rode bessen-proefveld 1963.
Gemiddelde struikopbrengst van 12 struiken in grammen.
Gegevens verstrekt door de Sinderhoeve.

Doel:

Het nagaan van de invloed van 4 beregeningstrappen A, B, C en D.

De nulhypothese, die moet worden getoetst luidt:

H_o: 'er is géén verschil tussen de beregeningstrappen A, B, C en D'.

Schema 1

	Kolom 1	Kolom 2	Kolom 3	Kolom 4	∑ kolom	Rijgem.	Rijgem.-niveau = rijeffect
Rij 1	1810.8^{A}	1734.6^{B}	1362.9^{C}	1518.8^{D}	6427.1	1606.8	- 17.2
2	1238.3^{D}	1394.2^{A}	1495.4^{B}	1264.2^{C}	5392.1	1348.0	-276.0
3	1382.5^{C}	1520.0^{D}	1313.3^{A}	1342.5^{B}	5558.3	1389.6	-234.4
4	1607.5^{B}	2238.3^{C}	2845.8^{D}	1915.0^{A}	8606.6	2151.6	527.6
∑ Rij	6039.1	6887.2	7017.4	6040.5	25984.1		
Kolomgemiddelde	1509.8	1721.8	1754.4	1510.1		1624.0 = Niveau	
Kolomgemiddelde-Niveau = kolomeffect	-114.2	97.8	130.4	-113.9			

Behandeling	∑	Beh.gem.	Beh.gem.-niveau = beh.effect
A	6433.3	1608.3	- 15.7
B	7161.2	1790.3	166.3
C	6247.9	1562.0	- 62.0
D	6141.7	1535.4	- 88.6
∑	25984.1		

Voor de variantie-analyse wordt berekend:

$$x^2 \quad : 1810.8^2+1238.3^2+ \ldots +1915.0^2 \qquad = 44872302.15$$

$$x_N^2 \quad : \frac{25984.1^2}{16} \qquad = 42198340.80$$

$$x_{Kol}^2 *: \frac{6039.1^2+6887.2^2+7017.4^2-6040.5^2}{4} - 42198340.80 = 210263.76$$

$$x_{Rij}^2 *: \frac{6427.1^2+5392.1^2+5558.3^2+8606.6^2}{4} - 42198340.80 = 1639314.01 \quad > 2007953.87$$

$$x_{Beh}^2 *: \frac{1608.3^2+1790.3^2+1562.0^2+1535.4^2}{4} - 42198340.80 = 158376.10$$

dus

$$x_T^2 = 44872302.15 - 42198340.80 - 2007953.87 = 666007.48$$

met dimensies 16 - 1 - 3 x 3 = 6

waaruit $S(\sigma^2) = \frac{666007.48}{6} = 111001.25$

Toets op behandelingseffect = $F_6^3 = \frac{158376.10/3}{111001.25} = 0.05$

met de bijbehorende overschrijdingskans P > 25%

Conclusie

H_o wordt dus niet verworpen en er kan géén significant verschil tussen de beregeningsgemiddelden worden aangetoond, of met andere woorden er is géén reden te veronderstellen, dat de gevonden verschillen tussen de beregeningsgemiddelden niet door toeval zouden zijn veroorzaakt.

Zijn nu in plaats van de gemiddelden van 12 struiken de opbrengsten per struik (12 struiken per veld van 36 m^2) bekend, dan wordt de verdeling van de dimensies:

Totaal	=	Niveau	+	Kol*	+	Rij*	+	Beh*	+	Rest
192	=	1	+	3	+	3	+	3	+	182

De berekening van de variantie-analyse is analoog aan het voorgaande.

Co-variantie-analyse

Verondersteld kan worden, dat de opbrengst per struik wordt beïnvloed door:

1. Het aantal trossen per struik = y_1
2. Het aantal gestel takken per struik = y_2
3. Het aantal éénjarige takken per struik = y_3

en zodanig, dat deze storende invloed van y_1, y_2 en y_3 in de toevalsruimte van x (= opbrengst per struik) ligt, met andere woorden dat van y_1, y_2 en y_3 de toevallige verschillen tussen de struiken (dus niet veroorzaakt door beregenings-

verschillen) de opbrengst per struik x beïnvloed hebben.

In dit geval wordt een co-variantie-analyse berekend teneinde te kunnen bepalen of deze invloed van y_1, y_2 en y_3 dusdanig is geweest, dat hiervoor een correctie op de opbrengstcijfers in rekening moet worden gebracht, om daarna dan de verschillen tussen de gecorrigeerde beregeningsgemiddelden op significantie te toetsen.

Evenals schema 1 voor $\underline{x}$ verkrijgt men schema 2, 3 en 4 voor respectievelijk $\underline{y}_1$, $\underline{y}_2$ en $\underline{y}_3$.

Schema 2. Gemiddelde aantal trossen per struik = $\underline{y}_1$

	Kolom 1	Kolom 2	Kolom 3	Kolom 4	∑ kolom	Rijgem.	Rijeffect
Rij 1	284.0^A	299.1^B	248.7^C	210.3^D	1042.1	260.5	39.1
2	166.8^D	217.8^A	219.2^B	223.2^C	827.0	206.8	-14.6
3	182.4^C	235.3^D	218.7^A	167.2^B	803.6	200.9	-20.5
4	181.5^B	224.3^C	240.2^D	224.0^A	870.0	217.5	- 3.9
∑ Rijen	814.7	976.5	926.8	824.7	3542.7		
Kolomgemiddelde	203.7	244.1	231.7	206.2		221.4 = Niveau	
Kolomeffect	-17.7	22.7	10.3	-15.2			

Behandeling	∑	Beh.gem.	Beh.effect
A	944.5	236.1	14.7
B	873.3	218.3	- 3.1
C	878.6	219.6	- 1.8
D	846.3	211.6	- 9.8
∑	3542.7		

Schema 3. Gemiddelde aantal gestel takken per struik = y_2

	Kolom 1	Kolom 2	Kolom 3	Kolom 4	∑ kolom	Rijgem.	Rijeffect
Rij 1	4.5 A	4.7 B	3.8 C	4.0 D	17.0	4.2	-0.3
2	4.6 D	4.7 A	4.2 B	4.2 C	17.7	4.4	-0.1
3	4.6 C	4.8 D	4.6 A	4.3 B	18.3	4.6	0.1
4	4.2 B	4.8 C	4.4 D	5.0 A	18.4	4.6	0.1
∑ Rijen	17.9	19.0	17.0	17.5	71.4		
Kolomgemiddelde	4.5	4.8	4.2	4.4		4.5 = Niveau	
Kolomeffect	0	0.3	- 0.3	- 0.1			

Behandeling	∑	Beh.gem.	Beh.effect
A	18.8	4.7	0.2
B	18.0	4.5	0
C	17.4	4.4	- 0.1
D	17.2	4.3	- 0.2
∑	71.4		

Schema 4. Gemiddelde aantal éénjarige takken per struik = y_3

	Kolom 1	Kolom 2	Kolom 3	Kolom 4	∑ kolom	Rijgem.	Rijeffect
Rij 1	18.0 A	23.7 B	22.3 C	24.1 D	88.1	22.0	-1.4
2	16.7 D	26.2 A	23.8 B	24.8 C	91.5	22.9	-0.5
3	18.8 C	24.8 D	25.8 A	22.7 B	92.1	23.0	-0.4
4	20.4 B	26.8 C	28.8 D	27.3 A	103.3	25.8	2.4
∑ Rijen	73.9	101.5	100.7	98.9	375.0		
Kolomgemiddelde	18.5	25.4	25.2	24.7		23.4 = Niveau	
Kolomeffect	- 4.9	2.0	1.8	1.3			

Behandeling	∑	Beh.gem.	Beh.effect
A	97.3	24.3	0.9
B	91.9	23.0	- 0.4
C	92.7	23.2	- 0.2
D	93.1	23.3	- 0.1
∑	375.0		

De vectoren $\underline{y}_1$, $\underline{y}_2$ en $\underline{y}_3$ moeten worden ontbonden in loodrechte componenten, één in de ruimte <u>kolom + rij + behandeling</u> en één hier loodrecht op, in de ruimte $\underline{R}$, gedefinieerd als <u>restruimte</u>:

$$\underline{y}_1 = \underline{y}_{1\,(kol + rij + beh)} + \underline{y}_{1\,R} = \underline{y}_{1\,N} + \underline{y}_{1\,Kol^*} + \underline{y}_{1\,Rij^*} + \underline{y}_{1\,Beh^*} + \underline{y}_{1\,R}$$

$$\underline{y}_2 = \underline{y}_{2\,(kol + rij + beh)} + \underline{y}_{2\,R} \qquad (1)$$

$$\underline{y}_3 = \underline{y}_{3\,(kol + rij + beh)} + \underline{y}_{3\,R}$$

Daarna wordt $\underline{x}$ geprojecteerd op de ruimte kolom + rij + behandeling en op R. Het model van de opbrengstvector $\underline{x}$ was oorspronkelijk:

$$\underline{x} = \underline{x}_N + \underline{x}_{Kol^*} + \underline{x}_{Rij^*} + \underline{x}_{Beh^*} + \underline{x}_R = \underline{x}_{(kol + rij + beh)} + \underline{x}_R \qquad (2)$$

en wordt thans:

$$\underline{x} = \underline{x}_{(kol + rij + beh)} + \beta_1\underline{y}_{1_R} + \beta_2\underline{y}_{2_R} + \beta_3\underline{y}_{3_R} + \underline{x}_T$$

waarin T nu de ruimte is van zuiver toeval.

Door projectie van $\underline{x}$ op R, dus op $\underline{y}_{1_R}$, $\underline{y}_{2_R}$ en $\underline{y}_{3_R}$ ontstaan 3 normaalvergelijkingen, waaruit de β's kunnen worden opgelost:

$$\underline{x} - (\beta_1\underline{y}_{1_R} + \beta_2\underline{y}_{2_R} + \beta_3\underline{y}_{3_R}) \perp y_{1_R}, \underline{y}_{2_R}, \underline{y}_{3_R}$$

etc.

waaruit volgt:

$$\underline{x}\,\underline{y}_{1_R} - \beta_1\underline{y}_{1_R}\underline{y}_{1_R} - \beta_2\underline{y}_{1_R}\underline{y}_{2_R} - \beta_3\underline{y}_{1_R}\underline{y}_{3_R} = 0$$

$$\underline{x}\,\underline{y}_{2_R} - \beta_1\underline{y}_{2_R}\underline{y}_{1_R} - \beta_2\underline{y}_{2_R}\underline{y}_{2_R} - \beta_3\underline{y}_{2_R}\underline{y}_{3_R} = 0$$

$$\underline{x}\,\underline{y}_{3_R} - \beta_1\underline{y}_{3_R}\underline{y}_{1_R} - \beta_2\underline{y}_{3_R}\underline{y}_{2_R} - \beta_3\underline{y}_{3_R}\underline{y}_{3_R} = 0$$

Voor het oplossen van deze normaalvergelijkingen moeten de restvectoren $\underline{y}_{1_R}$, $\underline{y}_{2_R}$ en $\underline{y}_{3_R}$ worden berekend, bovendien de restvector van $\underline{x}$, zodat uit de overeenkomstige schema's en volgens (1) en (2) de kolom-, rij- en behandelingseffecten vermeerderd met het niveau in mindering gebracht moeten worden op de oorspronkelijke vectoren.

Op deze wijze ontstaan per struik:

No	$\underline{y}_{1_R}$	$\underline{y}_{2_R}$	$\underline{y}_{3_R}$	en	$\underline{x}_R$
1	35.5	0.6	-1.0		- 136.9
2	- 3.5	-0.4	-1.0		893.1
3	15.5	-1.4	-3.0	= een matrix Y van (192x3)	653.1
4	109.5	-0.4	-2.0		1933.1
5	53.5	-0.4	4.0		1263.1
⋮	⋮	⋮	⋮		⋮
192	62.0	0.3	4.0		- 792.0

Het kwadraat van de lengte van $\underline{x}_R$ is:

$$(136.9^2+893.1^2+653.1^2+1933.1^2+1263.1^2+ \ldots +792.0^2) = 78494090 =$$

som van kwadraten voor rest met 192 dimensies.

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
⋮		⋮	⋮	⋮
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Nu is in matrix-notatie

$$Y\hat{\beta} = x_R$$

en

$$Y'Y\hat{\beta} = Y'x_R$$

noemen we nu $Y'Y = A$ = de matrix van inproducten
en $Y'x_R$ = de kolom van inproducten.

Het inproduct van bijvoorbeeld $\underline{y}_{1_R}$ en $\underline{y}_{2_R}$ is:

$$35.5x0.6+3.5x0.4-15.5x1.4-0.4x109.5+0.3x62.0 = 110.30$$

dan is weer in matrix-notatie:

$$\underbrace{\begin{pmatrix} 689702.20 & 110.30 & 20501.16 \\ 110.30 & 150.60 & 235.30 \\ 20501.16 & 235.30 & 4498.36 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix}}_{\hat{\beta}} = \underbrace{\begin{pmatrix} 2711850.74 \\ -14081.10 \\ 107627.95 \end{pmatrix}}_{Y'x_R}$$

en het rekenschema wordt:

$$\underbrace{\begin{matrix} 689702.20 & 110.30 & 20501.16 \\ & 150.60 & 235.30 \\ & & 4498.36 \end{matrix}}_{A} \quad \underbrace{\begin{matrix} 2711850.74 \\ -14081.10 \\ 107627.95 \end{matrix}}_{Y'x_R} \quad \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{I = \text{eenheidsmatrix}}$$

met de methode van choleski ontstaat hieruit:

$$\underbrace{\begin{matrix} 830.483 & 0.133 & 24.686 \\ & 12.271 & 18.908 \\ & & 59.426 \end{matrix}}_{} \quad \underbrace{\begin{matrix} 3265.390 \\ -1182.903 \\ 831.031 \end{matrix}}_{DY'x_R} \quad \underbrace{\begin{matrix} 0.0012 & & \\ 0 & 0.0815 & \\ -0.0005 & -0.0259 & 0.0168 \end{matrix}}_{D}$$

waarin

$$T'T = A$$

en

$$D'D = A^{-1}$$

dan is

$$\hat{\beta} = A^{-1}Y'x_R = D'DY'x_R$$

(nota Kamil no.140 dd.25-7-1962)

zodat

$$\hat{\beta}_1 = 0.0012x3265.390- \quad 0 \quad x1182.903-0.0005x831.031 = \quad 3.5030$$
$$\hat{\beta}_2 = \quad 0 \quad x3265.390-0.0815x1182.903-0.0259x831.031 = -117.9303$$
$$\hat{\beta}_3 = \quad 0 \quad x3265.390- \quad 0 \quad x1182.903+0.0168x831.031 = \quad 13.9613$$

Het kwadraat van de lengte van de projectie van $\underline{x}$ op $\underline{y}_{1_R} + \underline{y}_{2_R} + \underline{y}_{3_R}$ bedraagt:

$$Y'x_R D'DY'x_R = 2711850.74x3.5030+14081.10x117.9303+107627.95x13.9613 = 12662827.59$$

zodat de som van kwadraten voor zuiver toeval na regressie wordt:

$$\underline{x}_T^2 = \underline{x}_R^2 - \text{kwadraatprojectie} = 78494090-12662827.59 = 65831262.41$$

met dimensie $\quad 182 - 3 = 179$

en

$$S\ (\sigma^2) = \frac{65831262.41}{179} = 367772.42 = s^2$$

$$s = 606.44$$

De toets op de $\hat{\beta}$'s significant van '0' verschillen kan direct worden uit-

gevoerd met:

$$F^1_{179} = \frac{\left|\left[A^{-1}Y'x_R\right]\right|^2}{a_{ii}\ S\ (\sigma^2)} = \frac{\hat{\beta}_i^2}{a_{ii}\ S\ (\sigma^2)}$$

waarin a_{ii} = diagonaal element van A^{-1} (inverse matrix van A), te berekenen uit D:

$$a_{11} = 0.0012^2 + 0.0005^2 = 0.169 \quad x\ 10^{-5}$$
$$a_{22} = 0.0815^2 + 0.0259^2 = 0.00731$$
$$a_{33} = 0.0168^2 \qquad = 0.00028$$

$$F^1_{179} \text{ voor } \hat{\beta}_1 = \frac{3.5030^2}{0.169x10^{-5}xS(\sigma^2)} = 19.74 \quad \text{met} \quad P < 0.001$$
$$\hat{\beta}_2 = \frac{117.9303^2}{0.00731\ x\ S(\sigma^2)} = 5.17 \quad \text{met} \quad P < 0.05$$
$$\hat{\beta}_3 = \frac{13.9613^2}{0.00028\ x\ S(\sigma^2)} = 1.89 \quad \text{met} \quad P < 0.05$$

Het blijkt, dat $\hat{\beta}_1$ en $\hat{\beta}_2$ significant van '0' verschillen, indien $\alpha = 0.05$, dus kan op de behandelingsgemiddelden van $\underline{x}$ een correctie voor $\underline{y}_1$ en $\underline{y}_2$ in rekening worden gebracht.

De invloed van $\underline{y}_3$ (= aantal éénjarige takken) kan met een risico van 5% verwaarloosd worden.

De correctieterm wordt berekend uit:

$$- \sum_i \beta_i y_{i_{Beh}*}$$

en het behandelingsgemiddelde na correctie wordt: (zie ook schema 1, 2 en 3)

Behandeling	Gem.	β_1	$y_{1_{Beh}*}$	β_2	$y_{2_{Beh}*}$	Beh.gem.na correctie
A	1608.3	-3.5030	x14.7	+117.9303	x0.2	= 1.580 kg
B	1790.3	+3.5030	x 3.1	+117.9303	x0.0	= 1.801 kg
C	1562.0	+3.5030	x 1.8	-117.9303	x0.1	= 1.557 kg
D	1535.4	+3.5030	x 9.8	-117.9303	x0.2	= 1.546 kg

Van belang is tenslotte het verschil tussen de beregeningsgemiddelden A, B, C en D ná correctie van $\underline{y}_1$ en $\underline{y}_2$.

Om dit te toetsen wordt als nulhypothese gesteld:

H_o: 'er is géén verschil tussen de beregeningsgemiddelden'

en indien H_o geldt, verandert de ruimte van zuiver behandelingseffect in een toevalsruimte en het Latijnse vierkant gaat over in een schema met een 2 klassen indeling naar rijen en kolommen

De gehele bewerking moet nu worden herhaald voor de twee-klassen indeling dus zonder behandelingseffecten, er moeten weer normaal vergelijkingen worden opgelost om de nieuw β's te kunnen bepalen, en dus de nieuwe restvectoren te berekenen volgens:

$$\underline{x}_{R'} = \underline{x} - \underline{x}_N - \underline{x}_{Kol^*} - \underline{x}_{Rij^*} \tag{3}$$

en

$$\underline{y}_{1_{R'}} = \underline{y}_1 - \underline{y}_{1_N} - \underline{y}_{1_{Kol^*}} - \underline{y}_{1_{Rij^*}} \tag{4}$$

etc.

De vectoren $\underline{y}$ worden weer ontbonden in 2 loodrechte componenten, één in de ruimte <u>rij + kolom</u> en één in de ruimte hier loodrecht op, in $\underline{R}'$, welke nu <u>indien H_o geldt</u> een andere <u>restruimte</u> is dan $\underline{R}$.

Daarna projectie van $\underline{x}$ op deze ruimten, en

$$\underline{x} = \underline{x}_N + \underline{x}_{Kol^*} + \underline{x}_{Rij^*} + \beta'_1 \underline{y}_{1_{R'}} + \beta'_2 \underline{y}_{2_{R'}} + \beta'_3 \underline{y}_{3_{R'}} + \underline{x}_{T'}$$

De nieuwe restvectoren van $\underline{x}$ en $\underline{y}$ onstaan volgens (3) en (4) door uit de respectievelijke schema's nu de <u>kolom + rijeffecten</u> vermeerderd met het niveau

in mindering te brengen op de oorspronkelijke vectoren:

No	$\underline{y}_{1_{R'}}$	$\underline{y}_{2_{R'}}$	$\underline{y}_{3_{R'}}$	$\underline{x}_{R'}$
1	50.2	0.8	-0.1	- 152.6
2	11.2	-0.2	-0.1	877.4
3	30.2	-1.2	-2.1	637.4
4	124.2	-0.2	-1.1	1917.4
5	68.2	-0.2	4.9	1247.4
⋮	⋮	⋮	⋮	⋮
192	76.7	0.1	4.9	- 807.7

Het rekenschema wordt nu

$$\underbrace{\begin{matrix} 705285.36 & 237.46 & 21317.86 \\ & 151.09 & 246.32 \\ & & 4548.27 \end{matrix}}_{A} \quad \underbrace{\begin{matrix} 2723186.06 \\ -12518.58 \\ 104566.12 \end{matrix}}_{Y'x_{R'}} \quad \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{I}$$

$$\underbrace{\begin{matrix} 839.813 & 0.283 & 25.384 \\ & 12.289 & 19.459 \\ & & 59.374 \end{matrix}}_{T} \quad \underbrace{\begin{matrix} 3242.610 \\ -1093.355 \\ 733.171 \end{matrix}}_{DY'x_{R'}} \quad \underbrace{\begin{matrix} 0.0012 & & \\ 0 & 0.0814 & \\ -0.0005 & -0.0267 & 0.0168 \end{matrix}}_{D}$$

en

$$\hat{\beta}'_1 = 0.0012 \times 3242.610 - 0 \times 1093.355 - 0.0005 \times 733.171 = 3.5245$$

$$\hat{\beta}'_2 = 0 \times 3242.610 - 0.0814 \times 1093.355 - 0.0267 \times 733.171 = -108.5748$$

$$\hat{\beta}'_3 = 0 \times 3242.610 - 0 \times 1093.355 + 0.0168 \times 733.171 = 12.3173$$

De bijbehorende kwadraatprojectie:

$$3.5245 \times 3242.610 + 108.5748 \times 1093.355 + 12.3173 \times 733.171 = 12245045.76$$

en $$x^2_{R'} = 80394836.28$$

in mindering te brengen op de <u>oorspronkelijke</u> vectoren:

No	$\underline{y}_{1_{R'}}$	$\underline{y}_{2_{R'}}$	$\underline{y}_{3_{R'}}$	$\underline{x}_{R'}$
1	50.2	0.8	-0.1	- 152.6
2	11.2	-0.2	-0.1	877.4
3	30.2	-1.2	-2.1	637.4
4	124.2	-0.2	-1.1	1917.4
5	68.2	-0.2	4.9	1247.4
⋮	⋮	⋮	⋮	⋮
192	76.7	0.1	4.9	- 807.7

Het rekenschema wordt nu

$$\underbrace{\begin{matrix} 705285.36 & 237.46 & 21317.86 \\ & 151.09 & 246.32 \\ & & 4548.27 \end{matrix}}_{A} \quad \underbrace{\begin{matrix} 2723186.06 \\ -12518.58 \\ 104566.12 \end{matrix}}_{Y'x_{R'}} \quad \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{I}$$

$$\underbrace{\begin{matrix} 839.813 & 0.283 & 25.384 \\ & 12.289 & 19.459 \\ & & 59.374 \end{matrix}}_{T} \quad \underbrace{\begin{matrix} 3242.610 \\ -1093.355 \\ 733.171 \end{matrix}}_{DY'x_{R'}} \quad \underbrace{\begin{matrix} 0.0012 & & \\ 0 & 0.0814 & \\ -0.0005 & -0.0267 & 0.0168 \end{matrix}}_{D}$$

en

$$\hat{\beta}'_1 = 0.0012 \times 3242.610 - 0 \times 1093.355 - 0.0005 \times 733.171 = 3.5245$$

$$\hat{\beta}'_2 = 0 \times 3242.610 - 0.0814 \times 1093.355 - 0.0267 \times 733.171 = -108.5748$$

$$\hat{\beta}'_3 = 0 \times 3242.610 - 0 \times 1093.355 + 0.0168 \times 733.171 = 12.3173$$

De bijbehorende kwadraatprojectie:

$$3.5245 \times 3242.610 + 108.5748 \times 1093.355 + 12.3173 \times 733.171 = 12245045.76$$

en $$x^2_{R'} = 80394836.28$$

Zodat, indien H_o juist is, de som van kwadraten voor toeval na regressie wordt:

$$80394836.28-12245045.76 = 68149790.52 = x^2_{T'}$$

met dimensie: 182

Aangezien de som van kwadraten voor zuiver toeval x^2_T in het voorgaande berekend werd: 65831262.41 met dimensie 179, moet, indien H_o geldt, $x^2_{T'} - x^2_T$ ook een som van kwadraten voor toeval zijn met dimensie 3.

Uit de F toets:

$$\frac{(x^2_{T'} - x^2_T)/3}{x^2_T/179} = \frac{2318528.11/3}{367772.42} = 2.10$$

met een overschrijdingskans van ongeveer 10% volgt:

Conclusie:

De nulhypothese: 'er is géén verschil tussen de beregeningsgemiddelden' wordt niet verworpen, dat wil zeggen na correctie van de beregeningsgemiddelden op aantal trossen en aantal gestel takken per struik kan tussen de trappen A, B, C en D géén significant verschil worden aangetoond.

De variatie coëfficiënt bedraagt tenslotte:

$$\frac{606.44}{16.24} = 37.34$$

In tegenstelling met het rode-bessen proefveld van 1962, waarbij de variatie coëfficiënt op 19.69 werd berekend, ligt deze aan de hoge kant. Dit kan eventueel worden toegeschreven aan onder andere het snoeien, invloed van de zeer strenge winter 1962/1963 en heterogeniteit van de ontwikkeling van de struik.