Om Variationsanalyse. II.

Af R. K. Kristensen.

I Tidsskrift for Planteavl, 39. Bind, Side 536—539 (»Om Variationsanalyse«), blev det vist, at naar en Samling Elementer, der varierer i Størrelse, deles i et Antal (lige store) Grupper, vil Elementernes totale Variation, udtrykt som Afvigelseskvadratsum, være lig Gruppernes Variation plus Elementernes Variation inden for Grupperne. Er G Middeltallet af samtlige Elementer e, medens g er Middeltallet for en enkelt Gruppe, har man, naar der er r Elementer i hver Gruppe:

$$[(e \div G)^2] = \left[[(e \div g)^2]\right] + [r\,(g \div G)^2].$$

Fjærnes Gruppernes Variation, enten ved at korrigere Elementerne, saa de faar samme Gennemsnitsværdi i hver Gruppe, eller — simplere — ved at benytte Differenserne $e \div g$, der skyldes Elementernes Variation inden for Grupperne, og dannes ny Grupper, der ikke falder sammen med de forrige, vil den totale Variation være lig Gruppernes samlede Variation ved første og anden Gruppedannelse plus Variationen inden for Grupperne ved den sidste Gruppedannelse. Danner man Grupper (med paafølgende Elimination af Gruppernes Variation) s Gange, og betegnes Gruppernes Afvigelseskvadratsum — udtrykt som ovenfor — ved de fortsatte Gruppedannelser som A_1, A_2, A_3 o. s. v., medens Summen af de kvadrerede Afvigelser inden for Grupperne ved den sidste Gruppering betegnes som a_s, og A uden Indeks er den totale Variations Kvadratsum, har man:

$$A = A_1 + A_2 + \ldots\ldots\ldots A_s + a_s,$$

og dette gælder uden Hensyn til, efter hvilket System Grupperne dannes, eller paa hvilken Maade de griber ind i hverandre ved de fortsatte Gruppedannelser.

Eksempel 1. Af et Prøvedyrkningsmateriale er udtaget de i Fig. 1 opførte Afgrøder (jvf. Tidsskrift for Planteavl, 35. Bind, Side 625), medens Fig. 2 viser en almindelig Springertræksfordeling ved et femdelt Forsøg.

Fig. 1.

					Sum	Gsn., g
243	237	230	251	242	1203	241
251	228	238	248	256	1221	244
238	238	233	243	248	1200	240
212	229	228	245	252	1166	233
215	224	229	230	231	1129	226
				Gsn.	1184	237

Fig. 2.

1	2	3	4	5
4	5	1	2	3
2	3	4	5	1
5	1	2	3	4
3	4	5	1	2

Lader man de 5 »vandrette« Parcelrækker danne Grupper, faar man de ved Fig. 1 opførte Værdier af g, og Kvadratsummen $[r\,(g \div G)^2]$ bliver 1067 (udregnet ved at benytte Sumtallene i Stedet for g-Værdierne og dividere med $r = 5$ i Stedet for at multiplicere med 5). Derefter fjærnes Rækkernes Variation ved at trække g-Værdierne fra de enkelte e-Værdier, se Fig. 3; Kvadratsummen, der udtrykker Søjlernes (de »lodrette« Rækkers) Variation, udregnes ligesom før og bliver 1040. Nu elimineres ogsaa denne Variation, og man faar Tallene i Fig. 4.

Fig. 3.

	2	−4	−11	10	1
	7	−16	−6	4	12
	−2	−2	−7	3	8
	−21	−4	−5	12	19
	−11	−2	3	4	5
Sum	−25	−28	−26	33	45
Gsn.	−5	−6	−5	7	9

Fig. 4.

7	2	−6	3	−8
12	−10	−1	−3	3
3	4	−2	−4	−1
−16	2	0	5	10
−6	4	8	−3	−4

Heraf dannes det tredje Sæt Grupper ved at samle de Parceller, der hører til samme Forsøgsled, se Fig. 2. Disse Grupper giver Kvadratsummen 330, medens Variationen inden for Grupperne (Fællesparcellernes Afvigelse fra deres Middeltal, Forsøgsfejlen) giver 665, og alle 25 Parcellers Afvigelser fra Hovedgennemsnitstallet 237 giver Kvadratsummen 3102. Tillige udregnes Kvadratsummerne for Rækker, Søjler og Forsøgsled ved at benytte de oprindelige Afgrødetal i Fig. 1 uden at foretage nogen Elimination af Gruppernes Variation, og Forsøgsfejlens Kvadratsum bestemmes som Differens. Resultaterne er følgende:

	Med Elimination	Uden Elimination
Rækker	$A_1 = 1067$	1067
Søjler	$A_2 = 1040$	1040
Forsøgsled	$A_3 = 330$	330
Forsøgsfejl	$a_3 = 665$	(665)
Sum	3102	
Total Variation	3102	3102

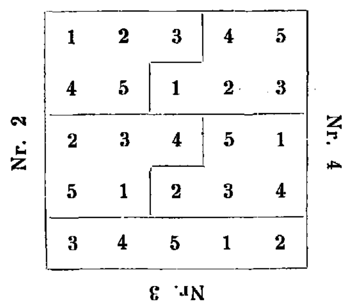
Som det ses, stemmer begge Fremgangsmaader overens, og Forsøgsfejlen kan beregnes efter Formlen:

$$m^2 = \frac{a_3}{n \div s(p \div 1) \div 1}, \qquad \text{(a)}$$

naar der er n Elementer, og p er Gruppernes Antal ved hver Gruppedannelse (se »Om Variationsanalyse«, Side 538, hvor den samme Formel — med lidt andre Betegnelser — er udviklet).

Eksempel 2. I Stedet for at lade Rækker og Søjler danne Grupper, benyttes de i Fig. 5 viste Afdelinger, hvor Forsøgsplanen er den samme som i Fig. 2. Samtlige Forsøgsled har en Parcel i hver Afdeling, og dette vil stadig være Tilfældet, naar Inddelingen fortsat drejes $^1/_4$ Omgang ret om (Numrene i Fig. 5 viser Afdelingernes Udseende ved de 4 Gruppedannelser, idet Numrene vendes opad), medens Forsøgsplanen og Afgrøderne ligger fast. Der foretages de samme Beregninger som før, idet de 4 Inddelinger træder i Stedet for de to forrige (Rækker og Søjler), og Beregningerne gentages med Benyttelse af selve Afgrødetallene uden Eliminering af Gruppernes Variation:

Fig. 5. Nr. 1



	Med Elimination	Uden Elimination
Nr. 1	$A_1 = 1566$	1566
» 2	$A_2 = 246$	1251
» 3	$A_3 = 483$	1633
» 4	$A_4 = 144$	1517
Forsøgsled	$A_5 = 330$	330
Forsøgsfejl	$a_5 = 333$	?
Sum	3102	
Total Variation	3102	3102

Ved Eliminationsmetoden stemmer de sammenlagte Kvadratsummer ligesom før med den totale Variations Kvadratsum, derimod er der ingen Overensstemmelse, naar Gruppernes Variation ikke bortskaffes efterhaanden. Forklaringen er følgende: I Eksempel 1 foregik en Gruppedannelse paa den Maade, at der til hver Gruppe udtoges et Element fra hver af de tidligere Grupper; et Forsøgsled havde en

Parcel i hver Række og een i hver Søjle. Men Summen af de Afvigelser, der skyldtes Rækkernes og Søjlernes Variation, var Nul, idet Plus og Minus ophævede hinanden, og en Bortskaffelse af denne Variation fik derfor ingen Indflydelse paa en Gruppes (et Forsøgsleds) Gennemsnitsværdi. Det var ligegyldigt, om man beregnede Søjlernes eller Forsøgsleddenes Variation af de oprindelige eller af de korrigerede Afgrødetal.

Ved de fortsatte Gruppedannelser efter Fig. 5 udtages derimod ikke en Parcel fra hver af de forrige Afdelinger men et skiftende Antal Parceller fra et skiftende Antal Afdelinger, fordi Afdelingerne ved de forskellige Gruppedannelser griber ind i hverandre paa forskellige Maader. Plus og Minus vil derfor ikke ophæve hinanden som før, Kvadratsummerne bliver forskellige, efter som Gruppernes Variation elimineres eller ikke[1], og i sidste Tilfælde vil de sammenlagte Kvadratsummer ikke stemme med den totale Variations Kvadratsum. Udføres Eliminationen, stemmer Kvadratsummerne, men Forsøgsfejlen kan ikke beregnes af Kvadratsummen 333 efter den før anførte Formel, som netop hviler paa, at en Gruppes Gennemsnitsværdi ikke paavirkes af de foretagne Eliminationer (Tidsskrift for Planteavl, 39. Bind, Side 538). De vilkaarlige Gruppedannelser kan fortsættes meget længe, inden a_s bliver 0, medens Divisor bliver 0 allerede efter 6 Gruppedannelser.

Eksempel 3. Fig. 6 viser en Forsøgsplan med 5 Forsøgsled og 10 Fællesparceller. Et Forsøgsled har to Parceller i hver Række og een i hver Søjle. Lader man Rækker, Søjler og Forsøgsled danne Grupper ligesom i Eksempel 1, vil der ved 2. Gruppedannelse udtages en Parcel af hver Gruppe fra 1. Gruppedannelse; tredje Gang udtages to Parceller af hver Gruppe fra 1. Gruppedannelse og een Parcel af hver Gruppe fra 2. Gruppedannelse. Ogsaa her vil Plus og Minus ophæve hinanden, fordi der udtages lige mange Parceller af hver Gruppe for hver enkelt Grupperings Vedkommende[2]. Kvadratsummerne bliver ens med og uden Elimination, og Forsøgsfejlen kan beregnes efter Formlen:

Fig. 6.

1	2	3	4	5	1	2	3	4	5
4	5	1	2	3	4	5	1	2	3
2	3	4	5	1	2	3	4	5	1
5	1	2	3	4	5	1	2	3	4
3	4	5	1	2	3	4	5	1	2

[1]) Dette gælder dog ikke Forsøgsleddenes Kvadratsum, der ikke paavirkes af de udførte Eliminationer, da Plus og Minus dækker hinanden her.

[2]) Ved det femdelte Forsøg (Eksempel 1) eller Forsøg af samme Type (kvadratiske Forsøg) udtages een Parcel af hver Gruppe fra hver Gruppedannelse, hvad der er ensbetydende med, at der ikke i en Gruppe kommer

$$m^2 = \frac{a_3}{n \div (p_1 \div 1) \div (p_2 \div 1) \div (p_3 \div 1) \div 1}, \quad \text{(b)}$$

hvor p_1, p_2 og p_3 betyder Antallet af Rækker, Søjler og Forsøgsled ($p_1 = p_3$. Formlen kan udvikles paa samme Maade som Formel a og er — ligesom denne — i Overensstemmelse med *Fishers* Variationsskema).

Eksempel 4. Fig. 7 har 10 Forsøgsled og 5 Fællesparceller. Et Forsøgsled har en Parcel i hver Række men kun een i hveranden Søjle, to og to Søjler maa derfor slaas sammen til en Afdeling (Gruppe), for at denne kan indeholde én Parcel af hvert Forsøgsled. Ved 2. Gruppedannelse udtages to Parceller fra hver af de første Grupper (Rækker), tredje Gang udtages en Parcel fra hver Gruppe ved de to forrige Gruppedannelser. Ogsaa her udtages lige mange Parceller af hver Gruppe fra hver Gruppedannelse, og man faar de samme Kvadratsummer, enten Grupperne elimineres eller ikke. Forsøgsfejlen kan dog ikke beregnes efter Formel b (hvor p_2 altsaa maatte betyde Antallet af Afdelinger), fordi ensidige Variationer griber forstyrrende ind. Forsøgsled Nr. 1 ligger f. Eks. stadig i venstre Side af Afdelingen, Nr. 2 i højre Side, og dette vil — hvis der er en Skraaplansvariation til Stede — forøge Fejlen paa Forsøgsleddenes Afgrøder, medens Forskellen mellem Fællesparcellerne og den af a_3 beregnede Middelfejl bliver forholdsvis lille. Kan den ensidige Variation afbildes ved to modsat vendende Skraaplaner, vil — som paavist i tidligere Arbejder — det modsatte være Tilfældet.

Fig. 7.

1	2	3	4	5	6	7	8	9	10
5	6	7	8	9	10	1	2	3	4
9	10	1	2	3	4	5	6	7	8
3	4	5	6	7	8	9	10	1	2
7	8	9	10	1	2	3	4	5	6

R. A. Fisher fordeler som bekendt Parcellerne tilfældigt, dog med den Begrænsning, at hvert Forsøgsled — ved et Forsøg som det sidst omtalte — skal have en Parcel i hver Række (Længderække) og een i hver Afdeling (Blok). Herved forvandles den ensidige Variation inden for Afdelingen til en tilfældig Variation, og Forsøgsfejlen kan beregnes, uden at den ensidige Variation griber forstyrrende ind. En saadan Fordeling maa anses for bedre end den i Fig. 7 viste (naar Forsøget gøres op ved en simpel Sammentælling af Fællesparcellernes Afgrøder — og man ser bort fra de praktiske Ulemper ved Parcellernes uregelmæssige Beliggenhed), fordi Virkningen af den ensidige

---

to eller flere Parceller, der har været sammen i en tidligere dannet Gruppe. Ved det rektangulære Forsøg falder denne Begrænsning bort, men Plus og Minus skal — ligesom ved det kvadratiske Forsøg — dække hinanden paa den angivne Maade.

Variation svækkes ved den tilfældige Fordeling af Parcellerne, men den maa anses for ringere end Fordelingen i Fig. 7, hvis denne kombineres med en Udjævning, der ophæver Virkningen af den ensidige Variation, da der kan ventes et bedre Resultat af at fjærne den ensidige Variation end af at forvandle den til en tilfældig.

Ved Forsøg af samme Type som det femdelte, hvor hvert Forsøgsled har en Parcel i hver Række og een i hver Søjle, ligger Sagen anderledes. Her er den tilfældige Fordeling i Almindelighed ringere end den systematiske, fordi den sidstnævnte er den bedst mulige, idet den giver den størst mulige Afstand mellem sammenhørende Parceller (Fællesparceller). Her anvender *Fisher* den tilfældige Fordeling af Hensyn til Fejlberegningen.

E k s e m p e l 5. I Fig. 8 er der 9 Forsøgsled og 6 Fællesparceller. Her gaar 1½ Søjle paa hver Afdeling, og ved 2. Gruppering dannes en Gruppe ved at udtage dels een, dels to Parceller fra hver af de første Grupper. Der udtages altsaa ikke lige mange Parceller fra hver Række, Plus og Minus vil ikke ophæve hinanden, Kvadratsummerne vil ikke stemme overens ved de to Beregningsmaader, og Formel b gælder derfor ikke. Men den ensidige Variation kan — ligesom i forrige Eksempel — fjærnes ved en eller anden Udjævningsmetode og Middelfejlen beregnes i Overensstemmelse med denne, eventuelt ved Hjælp af en Tilnærmelsesformel.

Fig. 8.

1	2	3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6
2	3	4	5	6	7	8	9	1
5	6	7	8	9	1	2	3	4
8	9	1	2	3	4	5	6	7

I *Fishers* Variationsskema bruges — som tidligere omtalt — Betegnelsen Frihedsgrader (degrees of freedom) om Divisorerne til de forskellige Kvadratsummer. Ved en almindelig simpel Fejlberegning med n Gentagelser er der $n \div 1$ Frihedsgrader eller, som det ogsaa undertiden udtrykkes, Frihedsgraden er $\frac{n \div 1}{n}$. Divisor til den Kvadratsum, der repræsenterer Forsøgsfejlen i Variationsskemaet, udtages — som Differens — af de opstillede Frihedsgrader. I en nylig udkommen Vejledning[1]) gives følgende Forklaring af Begrebet Frihedsgrad, der »has proved to be one of the most puzzling in *Fishers* works«, idet der benyttes et Eksempel med 60 Observationsresultater, hvis Afvigelser fra Middeltallet udregnes: »Of the 60 deviations thus derived, only 59 are quite independent. This means that if the values of 59 deviations are chosen arbitrarily, the sixtieth is fixed by the fact that the sum of the deviations must be zero.« Af denne Definition kan man ikke — lige saa lidt som af tidligere anførte — se, hvornaar Divisor i Middelfejlsformlen kan bestemmes

[1]) »Calculation and Interpretation of Analysis of Variance and Covariance« af *George W. Snedecor*, Ames, Iowa 1934.

paa den nævnte simple Maade (som Differens) og hvornaar ikke. Det maa afgøres ved en Undersøgelse af den Maade, som Grupperne dannes paa. I alle Tilfælde gælder det, at Divisor findes gennem en Analyse af de Beregninger, der har ført til den paagældende Kvadratsum. Frihedsgraden kan da udtrykkes ved en Brøk, hvor den fundne Divisor er Tæller, og Elementernes Antal er Nævner, idet Frihedsgraden er den fri, synlige Del af Fejlkvadratsummen (jvf. *N. P. Johansens* Redegørelse i Tidsskrift for Planteavl, 35. Bind, Side 657—658).

I et nyere Arbejde har *J. O. Irwin*[1]) paavist, at de forskellige Kvadratsummer og respektive Middelfejlsværdier i *Fishers* Variationsskema ikke er rene Udtryk for de forskellige Variationer, da disse fremtræder i Kombination med Forsøgsfejlen (eller de tilfældige Afvigelser). Det samme fremgaar af følgende Betragtning: Kaldes Rækkesummerne i Fig. 1 R og deres Gennemsnitsværdi R_g, er, som allerede nævnt paa Side 522,

$$[r\,(g = G)^2] = \frac{[(R \div R_g)^2]}{r},$$

idet r er Antallet af Elementer i hver Række. Ifølge Variationsskemaet har man, naar der er p Rækker,

$$m^2 = \frac{(R \div R_g)^2}{r\,(p \div 1)}.$$

Tænker vi os nu, at Materialet er uden ensidige Afvigelser, og at Elementerne kun er behæftet med tilfældige Afvigelser, der alle hører til samme Fejlsystem og følger den almindelige Fejllov, vil m være den almindelige Middelfejl paa Elementerne (jvf. Tidsskrift for Planteavl, 39. Bind, Side 538), og kaldes Middelfejlen paa Rækkesummerne M, er

$$M^2 = r\,m^2 \text{ eller } M^2 = \frac{[(R \div R_g)^2]}{p \div 1}.$$

Hvis nu Rækkerne yderligere behæftes med en selvstændig Variation, der er uafhængig af den tilfældige og karakteriseres ved Middelfejlen M_R, har man:
$$M^2 + M_R^2 = M_T^2 \text{ eller } M_R^2 = M_T^2 \div M^2,$$

hvor M_T er Middelfejlen, der betegner Rækkernes (Rækkesummernes) totale Variation og beregnes af denne paa sædvanlig Maade, medens $M = \sqrt{r}\,m$ er givet ved Formel a (eller b).

Variationsbestemmelser paa Grundlag af saa faa Gentagelser (i Eksemplet 5) er dog meget usikre og af ringe Værdi, hvis de skal benyttes paa anden Maade end som et Hjælpemiddel ved Bestemmelse af Forsøgsfejlen.

[1]) »Mathematical Theorems involved in the Analysis of Variance«. Journal of the Royal Statistical Society, 94. Bind, Side 284—300. London 1931.