Om Variationsanalyse.

Af R. K. Kristensen.

I »Nordisk Jordbrugsforskning«, 1932, Side 69—83, har *Olof Tedin*, Svalöf, meddelt Resultatet af nogle Undersøgelser, der er udført i Tilknytning til *R. A. Fishers* »Variationsanalyse« (se Referatet Side 540). De forskellige Kategorier af ensidige Afvigelser, som jeg ved mine tidligere beskrevne Fremgangsmaader har elimineret ved Bestemmelsen af den egentlige Forsøgsfejl[1]), bliver ved denne Analyse skilt ud som selvstændige Fraktioner, repræsenteret ved deres Afvigelseskvadratsum, og Divisorerne i de tilsvarende Middelfejlsformler udtages gennem et System af »Frihedsgrader«[2]). *Olof Tedin* skriver herom:

»Var och en, som räknat med *Kristensens* metod, vet nu, att när man skall bestämma standardavvikelse och medelfel på grundval av de reducerade Sd^2, så får man icke utan vidare dividera med $n \div 1$, utan olika korrektionsfaktorer måste komma till användning. Även i detta fall förefinnes fullkomlig principiell överensstämmelse mellan *Kristensens* och *Fishers* metoder och resultaten bli desamma, endast utförandet är olika. . . . *R. K. Kristensens* metoder för nämnarminskning äro väl kända, *Fisher* använder sig här av begreppet frihetsgrader, degrees of freedom (D. F.). Vad frihetsgrader verkligen är, deras matematiska innebörd, torde för alltid stanna utom räckhåll för de flesta icke-matematici. . . . Grundregeln är den, att antalet frihetsgrader i en variationsserie uppgår till antalet varianter, minskat med ett.«

I *Fishers* Fremstilling er det ikke vist matematisk, at de opstillede Frihedsgrader er identiske med de søgte Divisorer[3]),

[1]) Bestemmelse af Middelfejlen ved Kombinationer af ensidige og tilfældige Afvigelser. Tidsskrift for Planteavl, 28. Bind (1922), og senere Arbejder.

[2]) Statistical Methods for Research Workers. Edinburgh 1925.

[3]) Man savner en Definition af, hvad *Fisher* forstaar ved Frihedsgrader. Begrebet Frihedsgrad, der nærmest forekommer i angelsachsisk Litteratur, er klart nok, saa længe det drejer sig om en almindelig, simpel Fejlberegning, men ikke, naar Forholdene bliver mere sammensatte.

derimod har *N. P. Johansen* fremstillet og bevist den almindelige Formel (se Tidsskrift for Planteavl, 35. Bind (1929), Side 657):

$$m^2 = \frac{[v_s^2]}{(n \div 1)\,(n \div (s \div 1))},$$

der gælder for et vilkaarligt Antal (s) Fraktioner (en speciel Formel for $s = 2$ var allerede offentliggjort i det før nævnte Arbejde fra 1922). I det følgende forelægges en lille Undersøgelse over Forbindelsen mellem de søgte Divisorer og de omtalte Fraktioners Kvadratsummer.

Grundlaget for Variationsanalysen findes allerede i *E. Lindhards* og *V. H. O. Madsens* Undersøgelser over »det matematiske Grundlag for Dyrkningsforsøg paa Agermark« (Tidsskrift for Landbrugets Planteavl, 16. Bind (1909), Side 346), og det er angivet i *Tedins* her omhandlede Arbejde (Side 71). Det kan formuleres saaledes: pr Elementer e deles i p Grupper med r Elementer i hver. For hver Gruppe dannes Middeltal, g, af de r Elementer og dernæst Differenser, $e \div g$, mellem dette Middeltal og de enkelte Elementer. Endvidere dannes Differenser, $g \div G$, mellem de enkelte Middeltal og Middeltallet af alle pr Elementer samt mellem dette og de enkelte Elementer. Man har da for et vilkaarlig valgt Element:

$$e \div G = (e \div g) + (g \div G),$$
$$(e \div G)^2 = (e \div g)^2 + 2\,(e \div g)\,(g \div G) + (g \div G)^2.$$

Kvadratsummen for den Gruppe, som vedkommende Element hører til, bliver da (idet $g \div G$ er konstant for samme Gruppe):

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

eller, da $[(e \div g)]$ er 0:

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

og for alle p Grupper:

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

Betegnes disse tre Kvadratsummer henholdsvis som S, S_2 og S_1, har man:

$$S = S_2 + S_1.$$

Den totale Variation, udtrykt som Kvadratsum, er altsaa lig Gruppernes Variation plus Elementernes Variation indenfor Grupperne.

Vi kan nu eliminere Gruppernes Variation, enten ved at korrigere Elementerne, saa de faar samme Gennemsnitsværdi i hver Gruppe, eller simpelthen ved at benytte Differenserne $e \div g$ (jvf. Tidsskrift for Planteavl, 28. Bind, Side 96—97, og 35. Bind, Side 625). De saaledes fremkomne ny (eller afledede) Elementer kan atter deles i et Antal Grupper og de ny Gruppers Variation elimineres o. s. v. Da den før anførte Formel er absolut, gælder den uden Hensyn til efter hvilke Love, Elementerne varierer, og det er derfor ligegyldigt efter hvilke Regler, Grupperne dannes; en bestemt Kombination af visse Elementer kan gaa igen fra een Gruppedannelse til en anden. Man kan i alle Tilfælde sætte

$$S = S_1 + S_2, \; S_2 = S_3 + S_4, \; S_4 = S_5 + S_6 \text{ o. s. v.}$$

eller ved s Gruppedannelser:

$$S = S_1 + S_3 + S_5 + \ldots\ldots + S_{2s \div 1} + S_{2s},$$

idet Kvadratsummerne med ulige Indeks udtrykker Gruppernes Variation paa den før angivne Maade, medens de lige Indeks betegner Elementernes Variation inden for Grupperne.

Derimod kan man ikke ved et simpelt System af Frihedsgrader finde det Tal, hvormed S_{2s} skal divideres for at give den for ensidige Afvigelser befriede Middelfejl, naar man er gaaet videre end til 1. Gruppedannelse, og Elementerne udtages vilkaarligt. Divisor i Middelfejlsformlen kan ikke paa nogen simpel Maade udledes af Varianternes Antal.

Helt anderledes forholder det sig, naar Grupperne dannes paa den Maade, at to Elementer, der een Gang har været sammen i en Gruppe, ikke kombineres tiere (Tidsskrift for Planteavl, 35. Bind, Side 621—22). 2. Gruppedannelse vil da ske saaledes, at der til hver Gruppe udtages et Element fra hver af de forrige Grupper; næste Gang dannes en Gruppe ved at udtage et Element fra hver Gruppe ved baade 1. og 2. Gruppedannelse o. s. v. Heraf følger, at hver Gang man danner en ny Gruppe, vil dens Elementers Gennemsnitsværdi være upaavirket af de forrige Gruppedannelser med tilhørende Eliminationer; Gennemsnitsværdien vil være lig Gennemsnitsværdien af de tilsvarende oprindelige Elementer. Dette indses let, naar man f. Eks. tænker sig Grupperne ved 1. Gruppedannelse bestaaende af de »lodrette« Parcelrækker i et Markforsøg, ved 2. Gruppedannelse af de »vandrette« Rækker.

De ved 1. Gruppedannelse udførte Korrektioner vil da ophæve hverandre (+ og ÷ vil dække hinanden) i hver af de vandrette Rækker. Naar Grupperne fortsat skal dannes paa denne Maade, maa deres Antal være konstant og lig Antallet af Elementer i hver Gruppe.

En Fejlberegning kan nu gennemføres ud fra den Forudsætning, at alle ensidige Afvigelser er fjærnet gennem de udførte Gruppedannelser med tilhørende Eliminationer. Vi kan da se bort fra de ensidige Variationer og betragte Elementerne som kun behæftet med rent tilfældige Afvigelser, der alle hører til samme Fejlsystem og følger den almindelige Fejllov.

Ved 1. Gruppedannelse har vi — med Tilnærmelse — ifølge det foregaaende og ifølge den almindelige Fejlteori, naar n^2 er Elementernes Antal og m^2 er Kvadratet paa Middelfejlen, som udtrykker den tilfældige Variation:

Kvadratsum	Divisor	Middelfejlskvadrat
S	$n^2 \div 1$	m^2
S_1	$n \div 1$	m^2
$S_2 = S \div S_1$	$n^2 \div 1 \div (n \div 1)$	m^2

Ved 2. Gruppedannelse kan vi nu tage S_2 med tilhørende Divisor til Udgangspunkt. Divisor til S_3 fremgaar af, at Elementernes Gennemsnitsværdi i hver Gruppe er upaavirket af den udførte Korrektion. Gruppernes Variation er uforandret, følgelig er ogsaa Kvadratsum og Divisor uforandret, og den sidste er altsaa $n \div 1$. Divisor til S_4 er givet ved, at m^2 skal være konstant, hvoraf følger, at Kvadratsum og Divisor skal være proportionale; da S_4 er Forskellen mellem S_2 og S_3, maa Divisor til S_4 være lig Forskellen mellem Divisorerne til S_2 og S_3:

Kvadratsum	Divisor	Middelfejlskvadrat
S_2	$n^2 \div 1 \div (n \div 1)$	m^2
S_3	$n \div 1$	m^2
$S_4 = S_2 \div S_3$	$n^2 \div 1 \div 2\,(n \div 1)$	m^2

Saaledes kan nu fortsættes, idet Divisor til Kvadratsummerne med de ulige Indeks stadig er $n \div 1$, og den almindelige Formel for Middelfejlen bliver da:

$$m^2 = \frac{S_{2s}}{n^2 \div s\,(n \div 1) \div 1},$$

der er sammenfaldende med *N. P. Johansens* Formel, medens de fundne Divisorer er lig de af *Fisher* opstillede Frihedsgrader. Divisorernes Bestemmelse ad denne Vej er altsaa knyttet til den ganske bestemte Maade, som Grupperne, der bærer de forskellige Fraktioner af ensidige Afvigelser, er dannet paa. Der kan — paa de angivne Betingelser — udføres $n + 1$ Gruppedannelser (et bestemt Element kan $n + 1$ Gange kombineres med $n \div 1$ ny Elementer; $(n^2 \div 1) : (n \div 1) = n + 1$), men ved den sidste er alle Afvigelser elimineret, Kvadratsum og Divisor bliver Nul, og Middelfejlen kan ikke bestemmes (se Tidsskrift for Planteavl, 35. Bind, Side 657).

Efterskrift.

I en for nylig fremkommen Beskrivelse af *R. A. Fishers* Variationsanalyse[1]) skriver *Ivar Bachér* følgende om Begrebet Frihedsgrad: »Utan att här närmare ingå på den matematiska betydelsen av dette begrepp kan följande regel, avsedd att tjäna som vägledning vid beräkning av det mot en viss kvadratsumma svarande antalet frihetsgrader, uppställas: Antalet frihetsgrader är lika med antalet termer, som ingå i kvadratsumman, minskat med antalet av varandra oberoende relationer mellan dessa termer« (de her i Betragtning kommende Relationer er, at Summen af Afvigelserne fra et vist Gennemsnitstal er Nul: $[v] = 0$). Som det ses, overflødiggør denne Difinition ikke den i det foregaaende anførte Undersøgelse over Forholdet mellem Middelfejlen og de opstillede Frihedsgrader.

Om *Fishers* Ræsonnement (at naar ensidige Afvigelser er elimineret til en vis Grad, maa den tilbageværende Rest af ensidig Variation uskadeliggøres gennem en delvis tilfældig Spredning af Parcellerne — af Hensyn til Fejlberegningen) mener Forfatteren, at »dess tillämpning i vissa fall kan leda till praktiska svårigheter.« Senere hedder det, at de Fejl, der begaas ved at anvende de paagældende Formler paa Forsøg med systematisk Parcelfordeling, er saa smaa (ifølge *Tedins* Undersøgelser), »att de i flesta fall sakna praktisk betydelse«.

Forfatteren fremhæver, at Markforsøgets store Betydning kræver, »att det utföres efter en väl genomtänkt plan och på ett sådant sätt, att man har möjlighet på ett riktigt och fullt objektivt sätt bedöma tillförlitligheten hos de erhållna försöksresultaten. . . . I våra dagar torde försöksresultat, som ej genom angivande av »arbetsfel« eller på annat sätt giva möjlighet till en bedömning av tillförlitligheten, lämnas utan större avseende«.

---

[1]) »Moderna synpunkter på fältförsökets metodik och den statistiska analysen av försöksresultatet«, Nordisk Jordbrugsforskning, 1933, Side 220—239.