Statens Forsøgsvirksomhed i Plantekultur

Forsøgsstationen ved Hornum (I. Groven)

Værdital beregnet som produkt af skønnede tal

Values figures calculated as a produkt of estimated figures

Ragnhild Christensen og Poul Erik Brander

Indledning

I forsøgsafgrøder, hvor kvalitative egenskaber spiller en afgørende rolle, for eksempel blomsterkulturer, er det almindeligt at måle de forskellige egenskaber ved hjælp af skønnede tal, karakterer. Her i landet gives karakterer som regel i talområdet 0-10 (sjældnere 1-10), hvor karakteren 0 (henholdsvis 1) er den laveste, medens 10 er den højeste karakter, der gives for en egenskab.

Ved et forsøgs afslutning vil det tit være ønskeligt at fastlægge forsøgsleddenes rækkefølge entydigt, men med hensyntagen til alle bedømte egenskaber. Dette kan for eksempel gøres ved at udregne et aritmetisk (simpelt) gennemsnit, i hvilket man lader karaktererne for de enkelte egenskaber indgå som addender, men med større eller mindre vægt, afhængig af den vurderede egenskabs betydning for pågældende kulturs dyrkningsværdi. Indgår der mange karakterer i et aritmetisk gennemsnit, vil én enkelt lav karakter, selv om den er givet for en meget betydningsfuld egenskab, altså selv om den er tillagt stor vægt, imidlertid ikke påvirke gennemsnittet nævneværdigt. Udregnes værditallet derimod som et produkt, i hvilket de enkelte karakterer og deres eventuelle vægte indgår som faktorer (*Hammarlund*, 1969, side 26), vil én enkelt lav karakter få langt større indvirkning på værditallet.

Formålet med nærværende afhandling er at definere en produktformel til beregning af værdital ud fra karakterer. Anvendeligheden af det udledte udtryk illustreres ved hjælp af eksempler.

Definition

Først skal defineres nogle af de begreber, der arbejdes med.

Vægt:
En beregningsmæssig forskrift for, hvorledes en egenskab skal indgå i værditallet (*Hammarlund*, 1969, side 25). At en egenskab tildeles en vægt på 25% vil sige, at karakteren for denne egenskab påvirker værditallet med 25%.

S:
Den største talværdi, en karakter når op på i produktformlen, når der er taget hensyn til den vægt, hvormed pågældende egenskab må påvirke værditallet (fig. A).

Vægtfaktor:
S ÷ lavest mulige karakter (fig. A).

I nedenfor angivne produktformel indgår ud over *divisoren*, der er lig med forskellen mellem højeste og laveste karakter, 2 typer af faktorer, nemlig:

1) en *karakterfaktor*, der er den opnåede karakter ÷ lavest mulige karakter, og
2) en *vægtfaktor*, det vil sige den faktor, som ovenfor nævnte karakterfaktor skal multipliceres med, for at den pågældende egenskab kan komme til at påvirke værditallet med den ønskede styrke.

Produktformel:
En formel, ved hjælp af hvilken man kan beregne et værdital, Y_p, ud fra et vilkårligt antal karakterer givet i et defineret talområde, således at de enkelte karakterer indgår med den ønskede vægt,

Produktformel:

$$Y_p = \left[\frac{(k_1-1)\cdot vf_1}{9}+1\right]\cdot\left[\frac{(k_2-1)\cdot vf_2}{9}+1\right]\cdot\ldots\cdot\left[\frac{(k_n-1)\cdot vf_n}{9}+1\right] = \prod_{i=1}^{i=n}\left[\frac{(k_i-1)\cdot vf_i}{9}+1\right]$$

k_i: = den i'nte karakter

vf_i: = vægtfaktoren for den i'nte karakter og sådan at Y_p holdes indenfor talområdet. Dette sidste er navnlig af betydning, hvis der senere skal regnes videre på Y_p.

Ad empirisk vej er fundet frem til nedenstående pro 'uktformel, der gælder for talområdet 1-10.

I produktformlen er der en kantet parentes for hver karakter, der skal indgå i Y_p. Hver af disse parenteser består af 2 led. Første led er et produkt af karakterfaktoren og vægtfaktoren divideret med antal intervaller i det anvendte talområde. Andet led er lavest mulige karakter, der har været subtraheret under multiplikationen og nu må adderes igen. Det, der sker med karaktererne ved indsætning i produktformlen, er i realiteten en transformation af de nulpunktforskudte karakterer, ka_i, til den geometriske kurve $Y = X^n$, hvor n = det antal karakterer, der indgår i værditallet, og $X_i = \left[\frac{ka_i\ vf_i}{9} + 1\right]$.

Af tabel 1 fremgår det, at man ved at gå ud fra logaritmen til den højest mulige karakter let kan finde frem til de ønskede vægtfaktorer, og at værditallet kan beregnes med en nøjagtighed, der kun begrænses af antallet af cifre i den anvendte logaritmetabel.

Tabel 1. Vægtfaktorer (karakterskala 1-10)

log 10	Vægt i %	log S = log 10 · vægt	S	Vægtfaktor = S ÷ 1
A	B	C	D	E
1,0000	100	1,00000	10,0000	9,0000
	50	0,50000	3,1623	2,1623
	25	0,25000	1,7783	0,7783
	20	0,20000	1,5849	0,5849
	12,5	0,12500	1,3335	0,3335
	10	0,10000	1,2589	0,2589
	5	0,05000	1,1220	0,1220
	4	0,04000	1,0965	0,0965
	0	0,00000	1,0000	0,0000

Fremgangsmåden ved udarbejdelsen af tabellen er følgende:

Logaritmen til S i kolonne C findes ved at multiplicere logaritmen i kolonne A med vægten $\left[\frac{\text{kol. B}}{100}\right]$. S i kolonne D findes som antilogaritme til værdien i kolonne C. Vægtfaktoren i kolonne E er værdien i kolonne D ÷ lavest mulige karakter.

Produktformlen kan meget let gøres generel, så den kan bruges i andre talområder end 1-10. (Ved udarbejdelsen af tabellen erstattes logaritmen til 10 i kolonne A da med logaritmen til højeste tal i det nye talområde, og værdierne i kolonne C beregnes ved multiplikation, værdierne i kolonne D findes som antilogaritme til kolonne C, og værdierne i kolonne E som differencen mellem kolonne D og det nye talområdes laveste tal, der skal være større end 0).

I fig. A er givet en grafisk fremstilling af vægtfaktorer.

Når man har fundet frem til vægtfaktorerne, indsættes disse sammen med de foreliggende karakterer i produktformlen, og beregning af Y_p foretages. En manuel beregning er ret tidskrævende, men udregningen er let at udføre ved hjælp af EDB. Der er udarbejdet et EDB-program, som står til rådighed for eventuelle interesserede.

Nedenfor er vist eksempler på brug af produkt-

Eks. 1.

Karakterer:

10, 1, 9, 2, 8, 3, 7, 4

Vægte:

25%, 25%, 10%, 10%, 10% 10%, 5%, 5% = i alt 100%

$$Y_p = \left[\frac{(10\text{-}1)\cdot 0{,}7783}{9} + 1\right] \cdot \left[\frac{(1\text{-}1)\cdot 0{,}7783}{9} + 1\right]$$
$$\cdot \left[\frac{(9\text{-}1)\cdot 0{,}2589}{9} + 1\right] \cdot \left[\frac{(2\text{-}1)\cdot 0{,}2589}{9} + 1\right]$$
$$\cdot \left[\frac{(8\text{-}1)\cdot 0{,}2589}{9} + 1\right] \cdot \left[\frac{(3\text{-}1)\cdot 0{,}2589}{9} + 1\right]$$
$$\cdot \left[\frac{(7\text{-}1)\cdot 0{,}1220}{9} + 1\right] \cdot \left[\frac{(4\text{-}1)\cdot 0{,}1220}{9} + 1\right]$$
$$= 1{,}7783 \cdot 1 \cdot 1{,}2301 \cdot 1{,}0283 \cdot 1{,}2014 \cdot 1{,}0575 \cdot 1{,}0813 \cdot 1{,}0407$$
$$\log Y_p = 0{,}2500 + 0 + 0{,}08995 + 0{,}01233 + 0{,}07968 + 0{,}03395 + 0{,}02428 + 0{,}01732$$
$$= 0{,}50751$$
$$Y_p = 3{,}217 \simeq 3$$

formlen. For oversigtens skyld arbejdes der kun med 8 egenskaber og 3 forskellige vægte, nemlig *vægtene 25%, 25%, 10%, 10%, 10%, 10%, 5% og 5%*, i alt 100%. I eks. 3 vægtene 50%, 20%, 20% og 10%.

Eksempel 1 viser det almindelige tilfælde med karakterer spredt ud over talområdet. I eksemplet arbejdes med karaktererne 10, 1, 9, 2, 8, 3, 7 og 4.

Efter at de kantede parenteser er beregnet, findes log Y_p = 0,50751, d.v.s. at værditallet bliver 3.

Eks. 2. *Alle karakterer = 10*

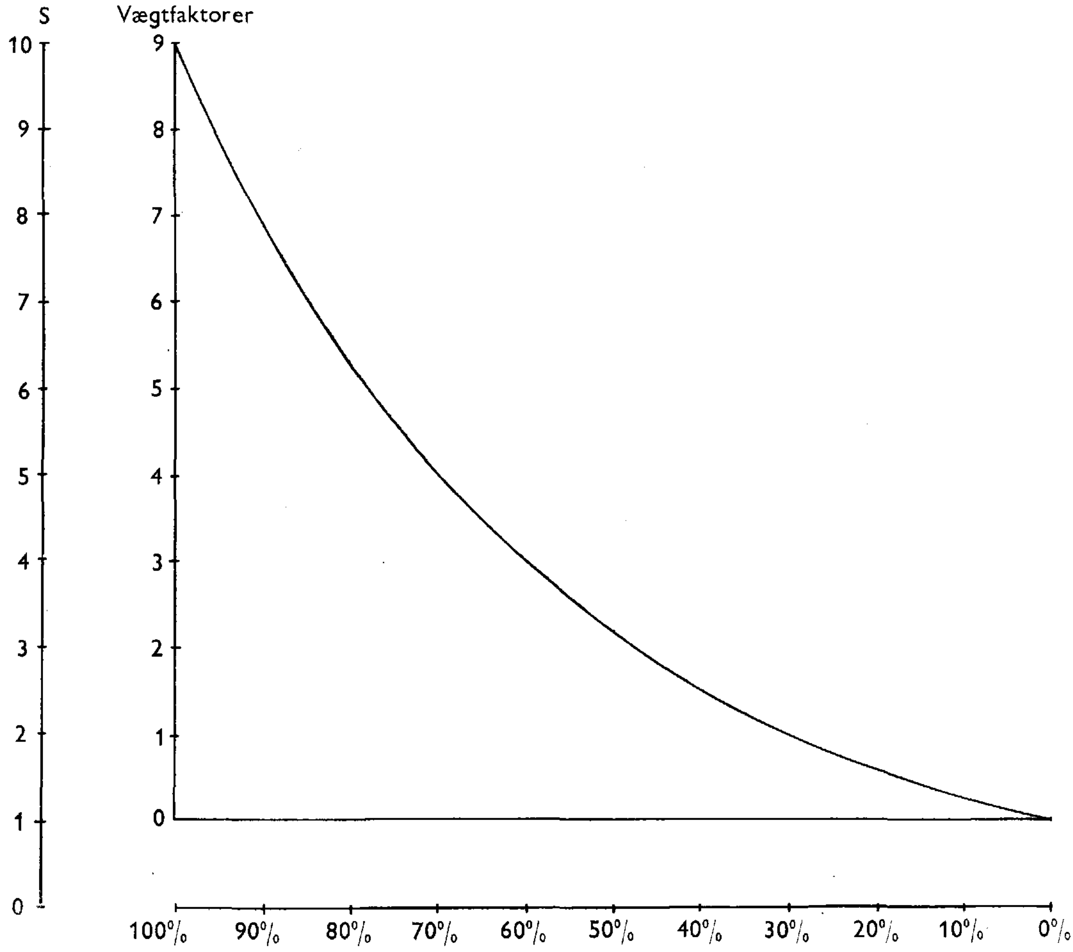
Vægte: se eks. 1.

$$Y_p = 1{,}7783^2 \cdot 1{,}2589^4 \cdot 1{,}1220^2$$

$$\log Y_p = 2 \cdot 0{,}25 + 4 \cdot 0{,}10 \cdot 2 - 0{,}05 = 1{,}0$$

$$Y_p = 10$$

I eksempel 2, hvor alle karakterer er 10, bliver størrelserne i de kantede parenteser netop lig S i tabellens kolonne D, fordi karakterfaktoren (10-1) elimineres af divisoren. Summen og logaritmerne bliver her = summen af vægtene, d.v.s. 1. Følgelig bliver Y_p = 10, d.v.s. højest mulige karakter.



Figur A. Vægtfaktorer i produktformel (karakterskala 1:10)
Figure A. Weight factors in the product formula (scale 1:10)

Eks. 3. Karakterer: *7, mgl., 9, mgl., 2, mgl., 4, mgl.* (4 karakterer ikke givet)
Vægte: 50%, 0%, 20%, 0%, 20%, 0%, 10%, 0% = i alt 100%

$$Y_p = \left[\frac{(7\text{-}1)\cdot 2{,}1623}{9} + 1\right] \cdot \left[\frac{(9\text{-}1)\cdot 0{,}5849}{9} + 1\right]$$
$$\cdot \left[\frac{(2\text{-}1)\cdot 0{,}5849}{9} + 1\right] \cdot \left[\frac{(4\text{-}1)\cdot 0{,}2589}{9} + 1\right]$$
$$\cdot \left[\frac{(0\text{-}1)\cdot 0{,}0000}{9} + 1\right]^4 = 4{,}293 \sim 4$$

Hvis enkelte karakterer mangler, for eksempel fordi man på et forsøgssted har undladt at bedømme en eller flere egenskaber, kan der alligevel beregnes et Y_p. Dette er illustreret i eksempel 3, hvor det er forudsat, at halvdelen af karaktererne mangler. Vægtfaktorerne for disse egenskaber sættes lig 0, men da karaktererne indgår i Y_p med forskellige forudbestemte vægte, må de resterende vægte forhøjes, således at summen af vægtene stadig bliver 100%. Ved beregningen bliver første led i de kantede parenteser, der hører til de manglende egenskaber, 0, fordi vægtfaktoren er 0, og parentesens samlede sum bliver altså 1. Det vil sige, at Y_p teoretisk forbliver uændret. I praksis har det vist sig, at en kurve over Y_p beregnet som et produkt af mange karakterer har en lidt stærkere krumning end en kurve over Y_p beregnet som et produkt af få karakterer (se fig. B).

Diskussion

Ovenfor er omtalt værdital, beregnet som det aritmetiske gennemsnit af de foreliggende karakterer, d.v.s. som

$$Y_a = \frac{\sum_{i=1}^{i=n} k_i v_i}{\sum_{i=1}^{i=n} v_i}$$

k_i = den i'nte karakter
v_i = den vægt, hvormed den i'nte karakter indgår i værditallet.

Desuden er beskrevet en metode til beregning af værdital som et produkt af nævnte karakterer, Y_p (se produktformlen side 129).

Værditallet kunne tænkes beregnet på mange måder. For eksempel ville det være nærliggende at prøve at beregne det som det geometriske gennemsnit af karaktererne, d.v.s. som

$$Y_g = \sqrt[\sum_{i=1}^{i=n} v_i]{\prod_{i=1}^{i=n} k_i^{v_i}}$$

Nedenfor er gennem eksempler og grafiske fremstillinger søgt illustreret, hvordan værditallet påvirkes af, hvilken beregningsmetode der er anvendt (jvf. oversigten i sammendraget). I alle de anførte eksempler er værditallene beregnet ud fra 21 karakterer, hvoraf den første vejer det femdobbelte af hver af de 20 sidste, der – for at lette oversigten og beregningsarbejdet – er ens.

Eks. 4. *Karakterer: 1 givet en gang (vejer femdobbelt) + karakteren 7 givet 20 gange* (jvf. fig. C1).

$$Y_a = \frac{1\cdot(1\cdot 5) + 20\cdot(7\cdot 1)}{25} = 5{,}8$$

$$Y_g = \sqrt[25]{1^5\cdot(7^1)^{20}} = \sqrt[25]{1\cdot 7^{20}} = 4{,}7$$

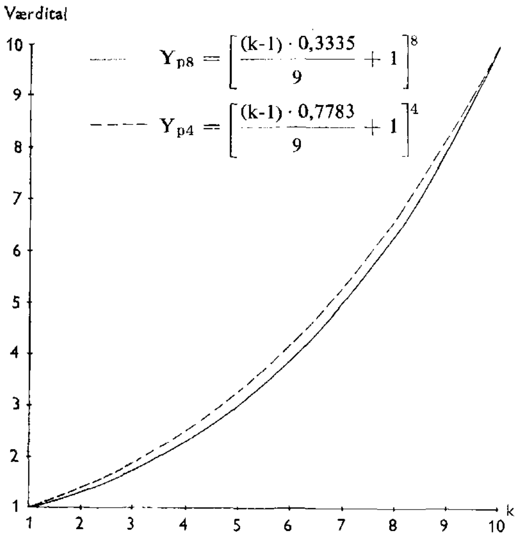
$$Y_p = \left[\frac{(1-1)\cdot 0{,}5849}{9} + 1\right] \cdot \left[\frac{(7\text{-}1)\cdot 0{,}0965}{9} + 1\right]^{20}$$
$$= 1\cdot 3{,}5 = 3{,}5$$

Eks. 5. *Karakterer: 10 givet en gang (vejer femdobbelt) + karakteren 7 givet 20 gange* (jvf. fig. C2).

$$Y_a = \frac{1\cdot(10\cdot 5) + 20\cdot(7\cdot 1)}{25} = 7{,}6$$

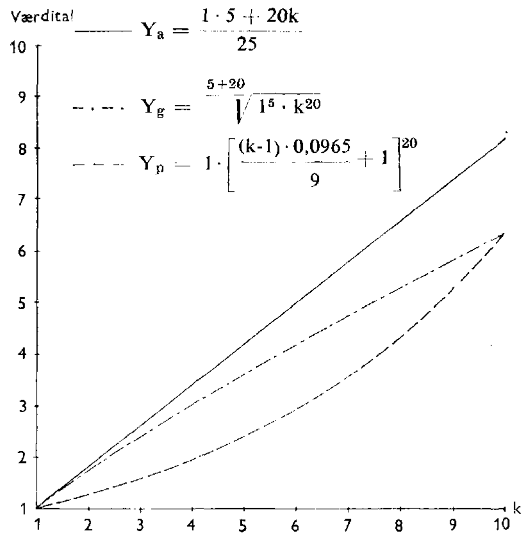
$$Y_g = \sqrt[25]{10^5\cdot 7^{20}} = 7{,}5$$

$$Y_p = \left[\frac{(10-1)\cdot 0{,}5849}{9} + 1\right] \cdot \left[\frac{(7-1)\cdot 0{,}0965}{9} + 1\right]^{20}$$
$$= 5{,}5$$



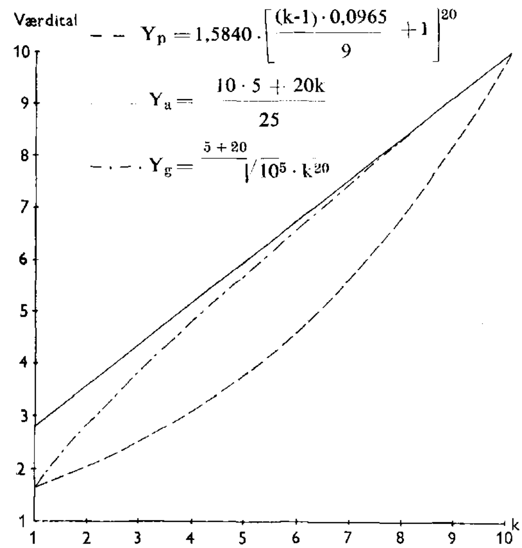
Figur B. Kurver over værdital udregnet ud fra 4 karakterer, der hver vejer 25 %, og ud fra 8 karakterer, der hver vejer 12,5 %.

Figure B. Curves for value figures calculated from 4 estimates, each weighing 25 %, and from 8 estimates, each weighing 12.5 %.



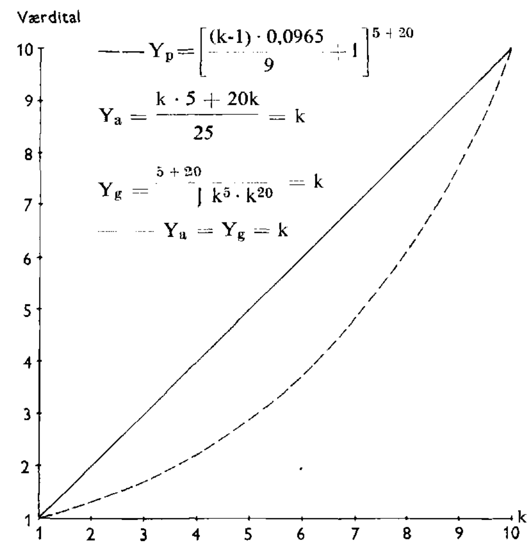
Figur C1. Kurver over værdital udregnet ud fra 21 karakterer, hvoraf karakteren 1 vejer 5-dobbelt, mens de øvrige er ens og gennemløber talområdet 1-10.

Figure C1. Curves for value figures calculated from 21 estimates, of which the estimate 1 has five times the weight of each of the 20 others, which are all the same and run through the interval 1-10.



Figur C2. Kurver over værdital udregnet ud fra 21 karakterer, hvoraf karakteren 10 vejer 5-dobbelt, mens de øvrige er ens og gennemløber talområdet 1-10.

Figure C2. Curves for value figures calculated from 21 estimates, of which the estimate 10 has five times the weight of each of the 20 others which are all the same and run through the interval 1-10.



Figur C3. Kurver over værdital udregnet ud fra 21 karakterer, hvoraf en karakter vejer femdobbelt; alle karakterer er ens og gennemløber talområdet 1-10. $Y_a = Y_g$.

Figure C3. Curves for value figures calculated from 21 estimates, of which one has five times the weight of each of the others. The estimates are all the same and run through the interval 1-10. $Y_a = Y_g$.

Eks. 6. *Alle karakterer ens* = *k* (jvf. fig. C3).

$$Y_a = \frac{k \cdot 5 + 20 \cdot k}{25} = k$$

$$Y_g = \sqrt[25]{k^5 \cdot k^{20}} = k$$

$$Y_p = \left[\frac{(k-1) \cdot 0{,}0965}{9} + 1\right]^{25}$$

$k = 7;\ Y_p = 4{,}8$

$k = 6;\ Y_p = 3{,}7$

Af fig. C1 og C2 fremgår det, at kurven for det aritmetiske gennemsnit, Y_a, er en ret linie, medens kurven for det geometriske gennemsnit, Y_g, er en parabel, og kurven for den beskrevne produktformel, Y_p, er en hyperbel. Følgelig er de 3 kurver ikke parallelle, og fordelen ved at bruge en af de omtalte metoder frem for de andre ved beregning af gennemsnit er større eller mindre, afhængig af hvor på kurven man befinder sig, og af om der indgår mange lave (fig. C1) eller mange høje (fig. C2) karakterer i gennemsnittet. I eksempel 6 og fig. C3 er vist det grænsetilfælde, hvor alle karakterer er ens. Her bliver kurven for det geometriske gennemsnit til en ret linie, det vil sige, at det aritmetiske og det geometriske gennemsnit falder sammen, og her har man maksimal fordel af at anvende produktformlen, under forudsætning af at man ønsker et værdital, der favoriserer høje karakterer og trækker nedad, når alle karakterer er på det jævne.

Omregning til andre talområder

Hvis man ønsker resultaterne, t, udtrykt i et andet talområde end 1–10, kan de ved indsætning i en simpel formel,

$$T = (t - \text{nedre grænse}) \cdot \frac{\text{ny øvre grænse} - \text{ny nedre grænse}}{\text{øvre grænse} - \text{nedre grænse}} \tag{4}$$

let omregnes til et hvilket som helst talområde i talrækken. Se eksempel 7.

Eks. 7. *Tal fra talområdet 1–10 omregnet til talområdet 0–200.*

$$\text{Formel: } T = (t - 1) \cdot \frac{200 - 0}{10 - 1}$$

t	T
10	200
1	0
4	67

Omregning af værdital til forholdstal ud fra gennemsnit og standardafvigelse

I løbende forsøg med blomsterkulturer, der vurderes ved karakterer, og hvorfra der jævnligt offentliggøres resultater, er det ønskeligt at have en konkret størrelse at gå ud fra, når resultaterne skal sammenlignes fra forsøg til forsøg; men det kan af mange årsager være vanskeligt at finde et acceptabelt sammenligningsgrundlag.

Går man imidlertid ud fra, at de tal, der ønskes sammenlignet, er fremkommet som resultat af mange målinger af mange egenskaber i et stort antal sorter, må det antages, at de er normaltfordelte. Det skulle følgelig være lovligt at sammenligne dem ud fra deres gennemsnit,

$$\bar{y} = \frac{\Sigma y}{n} \tag{5}$$

og deres standardafvigelse,

$$s = \sqrt{\frac{\Sigma y^2 - \frac{(\Sigma y)^2}{n}}{n - 1}} \tag{6}$$

Bruges gennemsnittet, $\bar{y}$, af værditallene fra et forsøg således som udgangspunkt, og sættes der en øvre og en nedre grænse lig med $\bar{y} \pm 3$ s (d.v.s. der foretages nulpunktforskydning og skalaændring), kan der ud fra disse størrelser beregnes et tal, V, for hver sort, der kan bruges som sammenligningsgrundlag forsøgene imellem.

Hvis V ønskes i talområdet 0–200, kan $\bar{y}$ sættes lig med 100, og man får V som et forholdstal med gennemsnit = 100. Den talværdi, der skal svare til øvre grænse, 200, bliver da $\bar{y} + 3$ s, og den nedre grænse, 0, svarer til $\bar{y} - 3$ s. Da tallene er normaltfordelte (kontrolberegninger af forhåndenværende materiale har bekræftet dette), vil – ifølge statistikkens love – ca. 68% af samtlige V placere sig mellem talværdierne 67 og 133 ($\bar{y} \pm s$), og kun få sorter, ca. 5%, vil få tillagt et V, der ligger under 33 og over 167 ($\bar{y} \pm 2\,s$). Kun helt undtagelsesvis, i knapt 3 tilfælde ud af 1000, vil der forekomme V under 0 og over 200 ($\bar{y} \pm 3\,s$).

Nedenfor gives en formel til og et eksempel på beregning af V.

Formel til beregning af V i et vilkårligt talområde:

$$V_{sort} = \frac{Y_{sort} - (\bar{y} - 3s)}{s} \cdot$$

$$\frac{\text{øvre grænse} - \text{nedre grænse}}{6} \qquad (7)$$

Tabel 2. Oversigt over de i teksten omtalte eksempler (jvf. fig. C_{1-3})
Table 2. Survey over the examples given in the text (see also Fig. C_{1-3})

	Karakter estimate figure	Antal bedømmelser number of repeated estimations	Vægt i% weight in%	Vægtfaktor weight factor	Y_a	Y_g	Y_p
Eks. 1	10	1	25	0.7783			
	1	1	25	0.7783			
	9	1	10	0.2589			
	2	1	10	0.2589			
	8	1	10	0.2589	(5.5)*	(3.9)	3.2
	3	1	10	0.2589			
	7	1	5	0.1220			
	4	1	5	0.1220			
Eks. 2	10	2	25	0.7783			
	10	2	10	0.2589	(10)	(10)	10
	10	2	10	0.2589			
	10	2	5	0.1220			
Eks. 3	7	1	50	2.1623			
	9	1	20	0.5849	(6.1)	(5.4)	4,3
	2	1	20	0.5849			
	4	1	10	0.2589			
Eks. 4 (fig. C1)	1	1	20	0.5849	5,8	4.7	3.5
	7	20	4	0.0965			
Eks. 5 (fig. C2)	10	1	20	0.5849	7.6	7.5	5.5
	7	20	4	0.0965			
Eks. 6 (fig. C3)	7	1	20	0.5849	7.0	7.0	4.8
	7	20	4	0.0965			

* De i parentes angivne værdier har ikke været omtalt i eks. 1-3, men er medtaget for fuldstændighedens skyld.
Values in brackets are not found in examples 1-3. They are included for the order af completeness.

Eks. 8. *Beregning af V i talområdet 0–200*

$$\text{Formel: } V_{sort} = \frac{Y_{sort} - (\bar{y} - 3s)}{s} \cdot \frac{200}{6} \qquad (8)$$

s: 1,20
$\bar{y}$: 5,50
$\bar{y} - 3s = 1{,}90$
$\bar{y} + 3s = 9{,}10$

Y_{sort} 6,25 $\qquad V_{sort} = \frac{6{,}25 - 1{,}90}{1{,}20} \cdot 33{,}33 = \mathit{121}$

Y_{sort}: 9,10 $\qquad V_{sort} = \mathit{200}$

Et EDB-program, der beregner $\bar{y}$ og V, står til rådighed for eventuelle interesserede.

Sammendrag

I denne afhandling er søgt defineret en produktformel til beregning af værdital ud fra karakterer, der har fået tillagt større eller mindre vægt. Anvendeligheden af det udledte udtryk er illustreret ved hjælp af eksempler, og der er draget en sammenligning mellem værdital udregnet som det aritmetiske gennemsnit, Y_a, som det geometriske gennemsnit, Y_g, og som et produkt, Y_p.

Forskellen i resultaterne fra de tre metoder fremgår af tabel 2.

Desuden omtales omregning af resultaterne til andre talområder, samt omregning af værdital til forholdstal med gennemsnit af samtlige værdital som udgangspunkt og øvre og nedre grænse lig med gennemsnittet $\pm$ 3 gange standardafvigelsen.

Summary

»*Values figures calculated as a produkt of estimated figures*«

In this paper a product formula (1) for calculating value figures, Y, from weighted estimate figures, k, is defined. The use of the formula is shown in examples 1-3.

Comparison is carried out between value figures calculated as an arithmetic mean, Y_a, as a geometric mean, Y_g, and as a product, Y_p, see examples 4-6.

Summary of the results is given in Table 2.

In example 7 the use of a formula (4) for converting figures into another scale is demonstrated.

At last a formula (7) for transgenerating value figures, Y, into proportion figures, V, from their average, $\bar{y}$ (5), and their standard deviation, s (6), is given — upper and lower limits put to $\bar{y} \pm 3s$. In example 8 the version of the formula (8) for the interval 0-200 is given together with a few practical examples of calculations of V.

Litteraturhenvisning

Hammarlund, Lars (1969): Skøn (karakter) som målemetode i forsøg. Duplikeret.