Puslerier med gennemsnit.

Af Børge Barfod[1]).

I.

Hvis jeg spørger en tilfældig mand på gaden, hvad gennemsnittet er af 2 og 8, vil han sandsynligvis svare 5 med en hovedrysten over, at jeg kan stille et så banalt spørgsmål. Hvis jeg derpå siger til ham, at gennemsnittet af 2 og 8 „lige så godt" kan være 4 eller $3^1/_3$ eller et hvilkensomhelst andet tal et sted mellem 2 og 8 vil han sandsynligvis vende sig bort og ikke mere snakke med sådan en tåbe!

Denne mands forestilling om gennemsnitsbegrebet var akkurat så primitiv, som den kan blive. Måske vil han alligevel på sit kontor eller på sit værksted flere gange om året tale om gennemsnittet af dette eller hint, og måske vil han ydermere drage konklusioner eller foretage dispositioner på grundlag af gennemsnittet dannet på den simpleste af alle måder som summen af en række tal divideret med deres antal.

Men måske kunne jeg have været en smule heldigere og stødt på en anden tilfældig mand, der godt ville godkende min påstand om, at f. eks. 4 kunne blive resultatet af en gennemsnitsberegning på 2 og 8.

Han ville sige, vi får 4 frem som gennemsnit, hvis 2 tillægges dobbelt så stor betydning eller vægt som 8. Tallet 2 ganges med vægten 2 og lægges til 8 ganget med vægten 1. Det giver 12 og dette tal divideres med summen af vægtene 2 og 1. Det giver så 4 som et *vejet gennemsnit* af 2 og 8. I fagteknisk sprogbrug: Et vejet *aritmetisk* gennemsnit.

Men hvis man til denne anden mand sagde, at gennemsnittet af 2 og 8 efter omstændighederne med god mening kan siges at være 4 *også,* hvis vi tillægger de to tal *samme* betydning eller vægt, ville han måske give op og bede om en nærmere forklaring. Og her er den. *Først* må vi forlange oplyst, hvad de to tal *konkret* dækker over. Uden en sådan *konkretisering* har det strengt taget slet ingen mening at ville udtale sig om,

[1]) Profesor, leder af Institutet for Driftsøkonomi og Statistik ved Handelshøjskolen i Aalborg.

hvad gennemsnittet af 2 og 8 er. Lad os som et eksempel tænke os, at en stue er 2 m på den ene led og 8m på den anden. Vi har nu et konkret indhold vedrørende de to tal. Men det er ikke *nok* til at give en entydig bestemmelse af tallenes gennemsnit. Den omstændighed at en stue er 2 m på den ene led og 8 m på den anden har *virkninger* i flere retninger. De to måske mest nærliggende virkninger er, at en sådan stue har en *omkreds* på 20 m og et *fladeindhold* på 16 m^2. Det er enten en smagssag eller kommer an på de nærmere omstændigheder, hvilken af de to virkninger, man vil anse for afgørende.

I alle tilfælde tvinges man til et *valg* til fremskaffelse af et grundlag for gennemsnitsberegningen. Lad os først vælge *omkredsen.* Med dette valg truffet kan vi definere og dermed også beregne det søgte gennemsnit på følgende måde:

Gennemsnittet af 2 og 8 er længden af væggen i en kvadratisk stue, hvis *omkreds er den samme* som stuens på 2×8 m, nemlig 20.

Det fører til det simple aritmetiske gennemsnit 5.

Vi kunne imidlertid i stedet have lagt afgørende vægt på den anden nævnte virkning: *fladeindholdet* og var da kommet til denne definition: Gennemsnittet af 2 og 8 er længden af væggen i en kvadratisk stue, hvis *fladeindhold er det samme* som i stuen på 2×8, nemlig 16.

Det fører til gennemsnittet 4, nemlig kvadratroden af 2 gange 8 og kaldes som bekendt det *geometriske* gennemsnit.

Det geometriske gennemsnit af 3 tal er den tredie rod af tallenes produkt o. s. v. Et geometrisk gennemsnit kan efter omstændighederne vejes. Tillægges 2 dobbelt så stor betydning som 8 bliver det geometriske gennemsnit $\sqrt[3]{2^2 \cdot 8}$. Rodexponenten er summen af vægtene, der tillige indgår under rodtegnet som eksponenter til de respektive tal. Det geomtriske gennemsnit er altid mindre end det aritmetiske.

Forskellen imellem de to definitioner ovenfor er den *virkning*, man har valgt som den afgørende. *Fælles* for de to definitioner er den betingelse, som udtryktes med ordene „den samme".

For at have et slagord, kunne man sige, at gennemsnittet her er defineret ved anvendelse af *„idemprincippet"* (idem = det samme). Dette skete som fjerde etape. De tre forudgående etaper på vej til en entydig definition var (1) *Konkretiseringen,* (2) *Specifikation* af tallenes virkninger og (3) *valg* af virkning[2]).

[2]) Disse synspunkter m. h. t. det ikke-stochastiske gennemsnitsbegreb er her i landet særligt pointeret af O. Strange Petersen, Nogle Bemærkninger om Gennemsnit, NTTØ 1940.

II.

Jeg mødte nu en mand, som jeg kendte. Han var en dygtig motorfører, men havde ikke dybere indsigt i gennemsnit. Jeg spurgte ham, hvad er gennemsnittet af 50 og100. Han mente, det måtte være 75. Jeg ved, sagde jeg til ham, at du forleden spurtede afsted med 100 km i timen for at nå færgen. På tilbagevejen havde du din svigermoder på bagsædet, og hun insisterede på, at du kun måtte køre med 50 kilometers fart. Fastholder du, at din gennemsnitsfart var 75 km i timen? Det ville nemlig være langt *rimeligere* at sige, at den var knapt 67 km/t.!

Lad os anvende etapemetoden fra før. Det konkrete er, at 50 og 100 er kilometer pr. time. Den mest *rimelige* virkning at se på her, er, at manden, der kørte en distance på 200 km ud og 200 km hjem, brugte 2 timer ud og 4 timer hjem, altså ialt kørte 400 km på 6 timer. Idemprincippet fører nu til definitionen: Gennemsnittet af 50 og 100 er den *ens* hastighed ud og hjem, som manden skulle have kørt med for at have tilbagelagt den *samme* samlede strækning på den *samme* samlede tid.

Dette fører til gennemsnittet $\frac{400}{6}$ = ca. 67 km/t. Sat på formel er beregningen:

$$\frac{\text{Strækning ud} + \text{Strækning hjem}}{\frac{\text{strækning ud}}{\text{hastighed ud}} + \frac{\text{strækning hjem}}{\text{hastighed hjem}}}$$

Hele nævneren er brugt tid.

Da frem og tilbage er lige langt, kan „strækning ud“ og „strækning hjem“ divideres bort, hvorved ses, at det ikke havde været nødvendigt at vide, hvor langt manden kørte for at udføre gennemsnitsberegningen. Efter denne „bortdivision“ får vi $\frac{2}{\frac{1}{50} + \frac{1}{100}}$ = ca. 67 og kommer derved til den type gennemsnit, der kaldes det *harmoniske*. Det er for 2 tal 2, for 3 talt 3, o. s. v. divideret med summen af de reciproke værdier af de tal, men skal finde gennemsnitet af. (Den reciproke værdi af 50 er 1 divideret med 50, den reciproke værdi af $\frac{2}{3}$ er $\frac{3}{2}$ o. s. v.).

Et vejet harmonisk gennemsnit er summen af vægtene divideret med summen af tallenes reciproke værdier efter multiplikation med deres respektive vægte; på formel således

$$\frac{V_1 + V_2 + \ldots\ldots}{\frac{V_1}{x_1} + \frac{V_2}{x_2} + \ldots\ldots}$$

hvor V er vægtene og x tallene, som vi tager gennemsnit af.

Det harmoniske gennemsnit er altid mindre end det geometriske, der som nævnt altid er mindre igen end det aritmetiske.

III.

Vi tager nu et lidt outreret og stiliseret eksempel.

I en fabrik, hvis produktion er underkastet stærke sæsonsvingninger, er produktionen 2000 ton pr. måned i årets 6 måneder og 14.000 ton pr. måned i årets andre seks måneder. Hvor stor er den gennemsnitlige produktion pr. måned?

Hvis fabrikens driftsleder svarer, at den naturligvis er 8000 t, kan man for så vidt ikke anfægte hans svar. Den definition, som han vel nærmest ubevidst lægger til grund for dette resultat, er, at den gennemsnitlige produktion pr. måned må være af en sådan størrelse, at – hvis produktionen faktisk havde haft dette lige store omfang pr. måned – *årsproduktionen var blevet den samme* som under den svingende produktion.

Men ligesom i de tidligere exempler er problemet, hvad man „dybest inde" ønsker svar på ved sin gennemsnitsberegning.

Lad os sige, at denne fabrik arbejder med progressive omkostninger. Man er derfor klar over, at en jævnere produktion ville lønne sig, og man stiller sig nu det spørgsmål, hvor meget kunne produktionen pr. år sættes i vejret med samme indsats af omkostninger som nu under den svingende produktion, hvis det var muligt at gennemføre fuldstændig jævn produktion måned for måned hele året?

For at besvare dette spørgsmål må man vide, hvordan omkostningerne varierer med produktionen pr. måned. Lad os antage, at omkostningerne pr. måned vokser med kvadratet på produktionen pr. måned. På formel

$$K = x^2$$

hvor K er omkostninger pr. måned i 1000 kr. og x produktion pr. måned i 1000 ton.

I lavsæsonen er omkostningerne pr. måned derfor

$$K = 2^2 = 4$$

og i hele lavsæsonen $6 \cdot 4 = 24$,
og i højsæsonen

$$K = 14^2 = 196 \text{ pr. måned;}$$

i hele højsæsonen $6 \cdot 196 = 1176$.

Årsomkostningerne bliver således

$$24 + 1176 = 1200$$

Ved en hel jævn produktion bliver omkostningsudviklingen pr. måned også helt jævn og 1200 : 12 = 100 pr. måned ved en sådan produktion, der medfører de samme årsomkostninger, som vi havde ved den jævne produktion.

Da omkostningerne pr. måned iflg. forudsætningen varierer med kvadratet på produktionen pr. måned, vil produktionen pr. måned kunne beregnes som kvadratroden af de pr. måned indsatte omkostninger. Når disse sidste er 100 kr. bliver derfor produktionen pr. måned 10 ton, årsproduktionen bliver 120 mod 96 ved den ujævne, og svaret på vort spørgsmål bliver, at årsproduktionen ved total sæsonudjævning og samme indsats af omkostninger, kan øges med 25 %. Dette svar kunne også og lettere findes ved en gennemsnitsberegning.

Tallet 10 ton pr. måned kan efter anførte forudsætninger tolkes som gennemsnittet af 2 og 14 beregnet således

$$\sqrt{\frac{2^2 + 14^2}{2}} = 10$$

Dette gennemsnit hører til blandt de kendte og navngivne typer og kaldes det *kvadratiske*.

Det kvadratiske gennemsnit er måske særlig kendt i tilknytning til beregning af spredningen i en statistisk iagttagelsesrække, hvor spredningen jo bestemmes som det kvadratiske gennemsnit af enkeltiagttagelsernes afvigelser fra deres aritmetiske gennemsnit.

IV.

I produktionsexemplet kom vi frem til det kvadratiske gennemsnit, fordi vi traf det valg at lade omkostningsvirkningerne være det afgørende moment i forbindelse med den specielt valgte *kvadratiske* form for omkostningsfunktionen. Hvis omkostningsfunktionen bevarer sin kvadratiske form, men skrives således

$$K = ax^2 + b$$

hvor a og b er konstante talværdier – „konkret“ er b de „faste“ omkostninger pr. måned, bliver gennemsnitsproduktionen ud fra samme problemstilling som ovenfor fortsat at beregne som et kvadratisk gennemsnit. Gennemsnittets talværdi og problemets løsning *påvirkes ikke af, hvilke talværdier a og b har.*

Skrives omkostningsfunktionen derimod således

$$K = ax^k + b$$

hvor exponenten k er positiv og angiver elasticiteten af de variable omkostninger vil ud fra iøvrigt samme problemstilling gennemsnittets *type* ændres, og typen vil blive bestemt af den talmæssige værdi af k. Lad os sige, at produktionen svinger måned for måned igennem året. Den gennemsnitlige produktion pr. måned defineret som ovenfor skal nu beregnes således

$$\bar{x} = \left(\frac{x_1^k + x_2^k + \cdots\cdots + x_{12}^k}{12} \right)^{\frac{1}{k}}$$

der specielt – som vi så – for omkostningselasticiteten 2 bliver det kvadratiske gennemsnit. For en lineær omkostningsfunktion d. v. s. $k = 1$, bliver $\bar{x}$ det almindelige artimetiske middeltal.

Hvis omkostningerne er degressive, d. v. s. elasticiteten k mindre ned 1, er der ingen fordel at hente ved at søge fremkaldt en jævnere produktion. Tværtimod ville årsproduktionen blive mindre ved samme indsats af omkostninger. Det giver sig udtryk i, at gennemsnittet af den månedlige produktion under den degressive forudsætning vil blive mindre end det aritmetiske middeltal. Lad f. eks. k kære 0,5 og månedsproduktionen svinge mellem 2 og 14 som i exemplet ovenfor.

Gennemsnittet af 2 og 14 bliver nu

$$\bar{x} = \left(\frac{\sqrt{2} + \sqrt{14}}{2} \right)^2$$

der udregnet bliver knapt 7. Under sidst anførte forudsætninger ville ved overgang til jævn produktion med samme indsats af omkostninger årsproduktionen falde i forholdet 8:7, idet 8 er det aritmetiske middeltal af 2 og 14. Sidstnævnte gennemsnit, der eventuelt kunne kaldes *kvadratrodsgennemsnittet,* har vistnok ikke noget almindeligt fastslået navn og beregnes altså ved at uddrage kvadratroden af tallene, lægge dem sammen og dividere med deres antal og til sidst opløfte det udkomne til anden potens. Værdien af et kvadratrods gennemsnit skyder sig ind imellem værdien af det geometriske og det aritmetiske.

V.

En mand køber et fjernsynsapparat på afbetaling. Restsummen, efter at den kontante udbetaling er fratrukket prisen, er 2.000 kr. Det aftales, at dette beløb skal afdrages med 100 kr. pr. måned. I samme forretning køber han samtidig et køleskab – også på afbetaling. Restkøbesummen er her 1.200 kr., der skal afvikles med 200 kr. om måneden. Sælgeren foreslår nu at sammenfatte de 2 køb i een købekontrakt med samlet restkøbesum 3.200 kr. Hvor meget skal køberen da ialt afdrage pr. måned? Dette tilsyneladende måske meget lette problem kræver nærmere eftertanke. Det er et gennemsnitsproblem og et gennemsnitsproblem af den art, som kan løses efter de retningslinier, som har været omtalt foran.

Exemplets tal er skematisk opstillet:

Restkøbesum	Rate pr. måned	Kredittid i mdr.
$p_1 = 1200$	$a_1 = 200$	$n_1 = 6$
$p_2 = 2000$	$a_2 = 100$	$n_2 = 20$

Sælgeren foreslog, at den samlede månedsrate ganske enkelt skulle være 300 kr. Så måtte efter hans mening såvel han selv som køberen blive stillet akkurat som, hvis de 2 køb var blevet afviklet separat.

Køberen mente, at han burde slippe med et samlet månedligt afdrag på mindre end 300 kr. Efter sælgerens forslag ville fælleskontraktens kredittid blive

$$\frac{3200}{300} = 10\,\frac{2}{3} \text{ måneder}$$

Køberen mente åbenbart, at afdragstiden burde være længere end disse ca. 11 mdr. Hvem havde ret? Og kan man overhovedet afgøre spørgsmålet på det foreliggende grundlag? Svaret på det sidste er nej. Man ser straks, at spørgsmålet om ratens størrelse i den sammenlagte kontrakt flyttes til at være et proglem om at finde „det rigtige" gennemsnit af 6 og 20. Sælgerens forslag, at finde fælleskontraktens løbetid ved at dividere summen af de separate rater op i den sammenlagte købesum, er ensbetydende med at ville beregne gennemsnittet af 6 og 20 som et med raterne vejet aritmetisk middeltal af de separate kredittider. Men om netop dette gennemsnit er „rigtigt" eller „galt" kan ikke, ganske som det har været illustreret ved tidligere exempler, afgøres uden at man får præciseret, hvad det er man „dybest inde" ønsker svar på ved gennemstitsberegningen. Baggrunden for at fremdrage dette exempel er en igangværende undersøgelse af afbetalingsforhold i Danmark. Et af målene er at belyse forbrugernes samlede afbetalingsgæld. Det er ensbetydende med at finde, hvor stor en kapital, sælgerne ialt har investeret i denne kreditform. Man har statistiske data for varegrupper og for forskellige sociale grupper og indkomstgrupper i form af restkøbesummer og tilsvarende månedlige afdrag ganske som i det lille exempel ovenfor.

Når afbetalingsgældens størrelse – eller set fra sælgernes synspunkt – størrelsen af det af afbetalingskreditten fremkaldte kapitalbehov – trækkes frem som en *væsentlig virkning* af afbetalingskøbenes størrelse og kredittidernes længde, har man dermed truffet det „valg af virkning", som er nøglen til en definition af det søgte gennemsnit, der nu kan formuleres således: Ved den gennemsnitlige kredittid af en række afbetalingskøb af forskellig størrelse med varierende kredittid forstås den kredittid, der i forbindelse med de samme samlede afbetalingsbeløb ville afføde det samme kapitalbehov, som affødes ialt af de separate køb.

Vi gør nu et lille sidespring og tænker os, at en sælger hver måned sælger for 9000 kr. restkøbesum på afbetaling over 3 måneder. Hvor stor kapital vil han så efter nogle måneders forløb have bundet i denne kreditgivning?

Det kan udregnes således:

	Salg pr. md.	Summeret salg	Rateindgang	Summeret rateindgang	Bunden kapital
	(1)	(2)	(3)	(4)	(5)
1. md.	9000	9000	0	0	9000
2. md.	9000	18000	3000	3000	15000
3. md.	9000	27000	6000	9000	18000
4. md.	9000	36000	9000	18000	18000
5. md.	9000	45000	9000	27000	18000

Kol (1) angiver månedligt salg efter fradrag for kontant udbetaling. Kapitalbehovet fremkommer i sidste kolonne, der er differencen imellem kolonne (2), der angiver, hvad han op til ethvert tidspunkt ialt ville have fået ind (udover de kontante udbetalinger), hvis han ikke havde ydet kredit og kolonne (3), der viser, hvad han faktisk har fået ind.

Kapitalbehovet bliver stationært 18000 fra den tredie måned. Havde kredittiden ikke være 3 måneder, men for exempel 17 md., var det stationære kapitalbehov (men naturligvis af en anden størrelse) indtrådt efter 17 måneder.

Nu kan dette stationære kapitalbehov imidlertid findes langt lettere efter følgende enkle formel[3]).

[3]) Om denne formel mv., se nærværende forfatters Om afbetalingssystemets virkninger på kapitalbehov og omsætning, H.vidensk. Tidsskr. 1954 og Formelsystem til brug for praktiske beregninger af kapitalprognoser (stencil.) 1952.

$$G = \frac{R}{2}(n + 1)$$

hvor G er det stationære kapitalbehov og identisk med forbrugernes afbetalingsgæld, R betyder restkøbesum pr. måned, og n er kredittidens længde i måneder. Ved indsættelse af $R = 9000$ og $n = 3$ fra exemplet finder vi $G = 18000$ som i tabellen.

Efter dette sidespring vender vi tilbage til sælgeren af fjernsynsapparatet og køleskabet, og, idet vi gør den forudsætning, at disse salg gentager sig stationært måned for måned, kan vi efter formlen ovenfor beregne kapitalbehovet for køleskabssalget til

$$\frac{1200}{2}(6 + 1) = 4200$$

og kapitalbehovet for fjernsynssalget til

$$\frac{2000}{2}(20 + 1) = 21000$$

Det samlede kapitalbehov bliver således 25.200. Efter anførte definition på den gennemsnitlige kredittid skal denne bestemmes som den kredittid, der i forbindelse med det samme samlede salg pr. måned, nemlig 3200, medfører det samme kapitalbehov, nemlig 25.000.

Formlen ovenfor kan nu bruges „baglæns", idet vi for G indsætter det samlede kapitalbehov og for R det samlede salg pr. måned og betragter n som ligningens ubekendte, nemlig den søgte gennemsnitlige kredittid. Vi får da

$$25200 = \frac{3200}{2}(\bar{n} + 1)$$

hvoraf findes $\bar{n} = 14\frac{3}{4}$ måned.

Med bogstaver ser beregningen således ud:

$$\frac{p_1}{2}(n_1 + 1) + \frac{p_2}{2}(n_2 + 1) = \frac{p_1 + p_2}{2}(\bar{n} + 1)$$

som løst m. h. t. $\bar{n}$ giver

$$n = \frac{n_1 p_1 + n_2 p_2}{p_1 + p_2}$$

hvorved ses, at den gennemsnitlige løbetid ud fra anførte forudsætninger, skal beregnes som et vejet aritmetisk gennemsnit, men ikke med raterne som vægte (som forhandleren foreslog), men med restkøbesummerne som vægte.

Raten i exemplet i den „sammenlagte" kontrakt bliver herefter 3200:14¾ = 217 kr. mod den simple sum på 300 kr.

Iøvrigt kan det let hænde, at raten i den „sammenlagte" kontrakt kan blive mindre end den største af de separate rater. Hvis vi ændrer tallene i exemplet til

Restkøbesum	Rate pr. md.	Kredittid i mdr.
1200	200	6
2000	80	25

bliver det med restkøbesummerne vejede gennemsnit af kredittiderne knapt 18 mdr. og raten i den „sammenlagte" kontrakt ca. 179 kr. og således mindre end den største af separatraterne på 200 kr.

VI.

En landbrugsvare til konsum er underkastet ugentlig prisnotering. Statistik vedrørende forbruget af samme vare findes kun tilgængeligt opgjort for kvartaler. Man stiller sig den opgave at belyse forbrugets afhængighed af prisen. Man står da til at begynde med overfor den opgave at skulle beregne en gennemsnitlig pris for hvert kvartal på grundlag af 12 eller 13 ugentlige noteringer. Hvorledes skal dette gennemsnit beregnes?[4]).

Lad os forenkle problemet til et exempel, hvor man kun har 2 prisnoteringer for hver periode, for hvilken der foreligger eet samlet tal for forbruget:

Periode I	Forbrug 250 ton	Prisnotering 1	8 kr.
		„ „ 2	2 „
Periode II	Forbrug 84,6 ton	Prisnotering 1	4 kr.
		„ „ 2	6 „

Vi forudsætter, at varens efterspørgselskurve har ligget fast i de 2 perioder, hvilket jo indebærer, at en ændring i forbruget fra den ene periode til den anden kun kan forklares ved, at prisen har ændret sig. Tager vi nu et almindeligt aritmetisk gennemsnit af prisnoteringerne, ses, at gennemsnitsprisen bliver 5 i begge perioder. Prisen bliver altså den samme, medens forbruget er faldet fra 250 til 84,6. Fastholdes forudsætningen om den fastliggende efterspørgselskurve, bliver konklusionen tilsyneladende, at varens priselasticitet må være uendelig stor. Lad os imidlertid sige, at man fra tidligere analyser af denne vare ved, at dette såre

[4]) Dette problem er nævnt af O. Strange Petersen, cit. sted.

langt fra er tilfældet, men at priselasticiteten ligger i nabolaget af ÷ 2. Følgelig må der være noget galt med den udførte gennemsnitsberegning. Og det er da heller ikke det aritmetiske middeltal, der skal anvendes her. I analogi med de i tidligere afsnit omtalte synspunkter, bliver den relevante definition af gennemsnittet i det foreliggende tilfælde: Gennemsnittet af 2 og 8 er den ens pris, varen skulle have kostet, for at forbruget skulle være blevet det samme, nemlig 250 ton. Og tilsvarende naturligvis for periode II.

Lad os antage, at priselasticiteten er konstant og lig med e. Efterspørgselsfunktionen har da som bekendt denne form

$$x = c \cdot p^e$$

hvor x er forbruget, p er prisen og c en konstant.

Lad x i formlen angive forbruget pr. *prisnoteringsperiode*.

For periode I med priserne 2 og 8 må derfor gælde, at

$$c \cdot 2^e + c \cdot 8^e = 250$$

Betegnes den søgte gennemsnitlige pris $\overline{p}$, må ifølge anførte definition på gennemsnittet tillige gælde, at

$$c \cdot \overline{p}^e + c \cdot \overline{p}^e = 250$$

Sammenholdes disse 2 ligninger, findes

$$\overline{p} = \left(\frac{2^e + 8^e}{2}\right)^{\frac{1}{e}}$$

hvorved ses, at *typen* af det gennemsnit, som skal anvendes på prisnoteringerne, afhænger af priselasticiteten.

Som *hypotese* – f. ex. på grundlag af tidligere analyse af varen – kan man nu prøve med e = ÷ 2.

Man får da

$$\overline{p}_{I} = \sqrt{\frac{2}{\frac{1}{2^2} + \frac{1}{8^2}}} = 2,74$$

og for den anden periode

$$\overline{p}_{II} = \sqrt{\frac{2}{\frac{1}{4^2} + \frac{1}{6^2}}} = 4,71$$

og finder således nu en højere pris i den anden periode, hvor forbruget også var mindre.

For elasticiteten ÷ 2 findes gennemsnittet altså som kvadratroden af antallet af priser divideret med summen af de reciproke værdier af kvadratet på de enkelte priser.

Denne type kunne eventuelt kaldes et *harmonisk gennemsnit af anden orden.*

Hvis forudsætningerne om fastliggende efterspørgselskurve og konstant priselasticitet er rigtige og forudsat de statistiske iagttagelser ikke er behæftet med fejl, kan hypotesen om e = ÷ 2 nu efterprøves. Vi fandt

	$\bar{p}$	x
Periode I	2,74	250
Periode II	4,71	84,6

Efter forudsætningen om konstant elasticitet finder man nu

$$e = \frac{\log 250 - \log 84{,}6}{\log 2{,}74 - \log 4{,}71}$$

$$e = \frac{2{,}397 - 1{,}927}{0{,}438 - 0{,}673} = -\frac{0{,}470}{0{,}235} = -2$$

Således som forbrugstallene var valgt til exemplet, blev hypotesen bekræftet. *Men havde man ved den logaritmiske udregning fundet en fra ÷ 2 afvigende værdi af e, havde man måttet prøve med en ny hypotese og fortsat hermed indtil overensstemmelse imellem den valgte e-værdi til fastlæggelse af gennemsnittets type og den e-værdi, der efterfølgende findes ved den logaritmiske beregning.*

Er priselasticiteten specielt ÷ 1, går anførte formel over i det harmoniske gennemsnit. Formlen, der er den samme, som vi kom frem til i produktionsexemplet ovenfor, går, som vi så, over i det aritmetiske gennemsnit for e = plus een. Anvendelsen af et aritmetisk middeltal af priserne i nærværende forbrugsexempel, betinger derfor, enten at priselasticiteten er plus een (hvilket jo normalt er ganske urealistisk) eller, at efterspørgselskurven er en ret *faldende* linie på det relevante stykke.

VII.

Da mr. Nixon besøgte Sovjetunionen, skulle Khrustjov ifølge referater have bemærket, at det amerikanske bruttonationalprodukt (i det følgende forkortet BNP) var underkastet en årlig vækst på omkring 2 %. Nixon skulle have repliceret, at væxten var væsentlig større end disse 2 %, der ikke androg stort mere end den årlige befolkningstilvæxt. Ifølge den amerikanske vicepræsident var væxten omkring det dobbelte af Khrustjovs tal, og Nixon udtalte til anden side sin forundring over den sovjetiske ministerpræsidents komplette uvidenhed om den amerikanske økonomi!

Da de to statsmænd har adgang til nøjagtig de samme informationskilder vedrørende den økonomiske udvikling i USA, må man, forudsat referatet ovenfor nogenlunde dækker, hvad der blev sagt, antage, at deres uenighed på dette punkt må hænge sammen med noget mere *formelt*. Måske noget, der har lidt med gennemsnit at gøre.

I omstående tabel er i kolonne (1) anført udviklingen i det amerikanske BNP 1946/59 i milliard dollars i faste 1958-priser, taget fra *Economic Report of the President*, januar 1959. Tallet for 1959 er dog tilføjet af os som et skøn baseret på senere publicerede kvartalstal.

Ville man nu bære sig så forkert af, som overhovedet muligt, nøjedes man med at se på tallene 312 (for 1946) og 480 (for 1959) og udregnede en procentisk stigning på 54 % over disse 13 år, hvorefter man ville sige, at den gennemsnitlige tilvæxtprocent var 4,2 %, nemlig det aritmetiske middeltal 54:13.

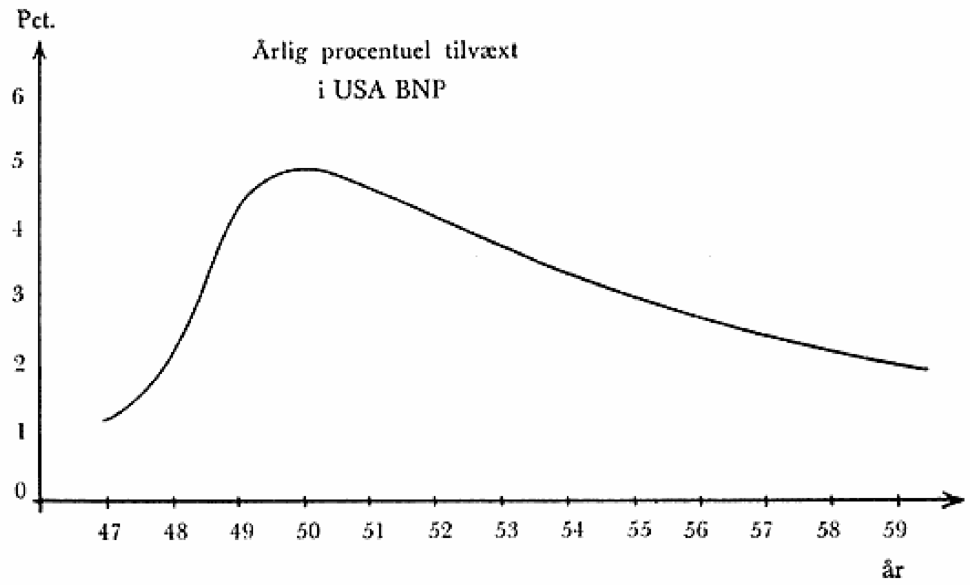
Ved denne lette fremgangsmåde begår man i virkeligheden 3 fejl samtidig. Den første fejl er, at man bruger et aritmetisk gennemsnit i stedet for, som vi senere skal se, et geometrisk. Den anden fejl, der ikke har noget med gennemsnitsformler at gøre, er, at sammenligningen mellem tallene ovenfor ikke tager hensyn til, at udviklingen af BNP er påvirket ikke blot af en langtidstendens, hovedsagelig bestemt af befolkningstilvæxten og den teknologiske udvikling, men også af den rytmiske konjunkturudvikling. Problemet her er at belyse væxten i langtidskræfterne. Ved at sammenligne 1946, der i konjunkturmæssig henseende nærmest var et normalår, med 1959, der er præget af en højkonjunktur, får man derfor en for høj tilvæxtprocent i forhold til det, man ønsker at belyse.

Den tredie fejl (der måske skulle sættes i citationstegn) er, at man bare hæfter sig ved den *gennemsnitlige* stigning i stedet for også at se på udviklingen af tilvæxtprocenten år for år. Kun det sidste kan bidrage til et fingerpeg om udviklingen videre frem.

For at „rette“ fejl nr. 2 har vi nu tallene i kolonne 2, der er fremgået af en grafisk udjævning på tallene i første kolonne. Kolonne 2 angiver langtidstendensen i BNP renset for cykliske komponenter. Kolonne 3 viser den årlige tilvæxtprocent på grundlag af tallene i kolonne 2. Kolonne 3 er vist hosstående i kurveform. Man ser, at tilvæxtprocenten toppede i 1950 og har siden været faldende, dog fra 1952 med aftagende fart og ligger nu i 1959 på 2 % – hr. Khrustjovs tal!

USA bruttonationalprodukt.

	(1)	(2)	(3)	(4)
1946	312	309	–	–
47	312	313	1,3 %	1.013
48	324	320	2.2 %	1.022
49	324	334	4.4 %	1.044
50	352	351	5.1 %	1.051
51	380	368	4.8 %	1.048
52	393	384	4.3 %	1.043
53	411	399	3.9 %	1.039
54	403	413	3.5 %	1.035
55	435	426	3.1 %	1.031
56	446	438	2.8 %	1.028
57	451	449	2.5 %	1.025
58	437	459	2.2 %	1.022
59	480	468	2.0 %	1.020



Hvad er nu den rette beregning af den *gennemsnitlige* årlige tilvæxtprocent? Vi får i analogi med de tidligere omtalte exempler definitionen: Den søgte gennemsnitsprocent er den – lige store – årlige procentiske tilvæxt, som BNP skulle have haft for gennem de 13 år at vokse fra netop 309 til netop 468. Betegnes det søgte gennemsnit, $\bar{q}$, findes $\bar{q}$ af den velkendte rentes rente formel

$$468 = 309\,(1 + q)^{13}$$

eller

$$\bar{q} = \sqrt[13]{1.53846} \div 1$$

$$\bar{q} = 3.2\ \%$$

Man finder således gennemsnittet uden at der i beregningen behøver indgå de enkelte tal, som man tager gennemsnittet af!

Beregningen er *identisk* med det geometriske gennemsnit af tallene i kolonne (4) fradraget 1. Betegnes de årlige tilvæxtprocenter $q_1\ q_2\ \ldots\ q_{13}$, bliver tallene i kolonne 4: $q_1 + 1,\ q_2 + 1,\ \ldots\ q_{13} + 1$ og vi får gennemsnittet

$$\bar{q} = \sqrt[13]{(q_1 + 1)\,(q_2 + 1)\ \ldots\ (q_{13} + 1)} \div 1$$
$$= 3.2\ \%$$

Den sidste formel må bruges i opgaver, hvor man kun kender tilvæxtprocenterne, men ikke det absolutte begyndelses- og sluttal svarende her i exemplet til BNP 1946 og 1959.

Den korrekte gennemsnitlige stigningsprocent er altså i det foreliggende tilfælde 3.2 %. Stigningen bliver overvurderet, hvis den beregnes som

$$\frac{1}{13} \cdot \frac{468 \div 309}{309}$$

hvilket giver 4 % – mr. Nixons tal!

Ifølge United Nations Statistical Yearbook 1957 var den gennemsnitlige årlige befolkningstilvæxt i USA over årene 1953/56 1.8 %.

VIII.

Til sidst et exempel, der teoretisk ligger på en anden linie end alle foregående. Gennemsnitsbegrebet som det defineredes i forbindelse med de hidtil omtalte tilfælde havde ingen tilknytning til det centrale område, der opfyldes af sandsynlighedsteorien og statistiske fordelinger.

En fabrik foretog et „produktivitetsforsøg“ med 28 arbejdere over 4 arbejdsdage med sammenlagt effektiv arbejdstid 30 timer. De 28 arbejdere udførte individuelt et og samme slags arbejde, og resultatet blev:

2 arbejdere	udførte	hver	10 stk.	ialt	20	stk.
7 „	„	„	12 „	„	84	„
10 „	„	„	15 „	„	150	„
7 „	„	„	20 „	„	140	„
2 „	„	„	30 „	„	60	„
28 arbejdere	udførte ialt:				454	stk.

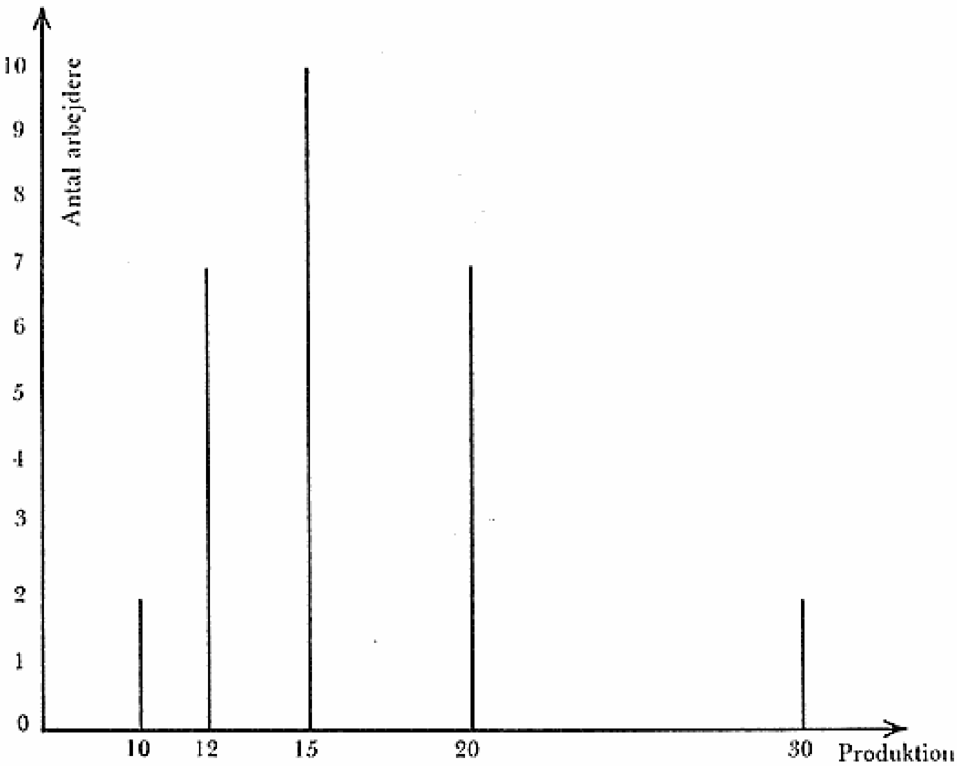
Vi har her at gøre med en *statistisk fordeling*. Og der rejser sig det sædvanlige problem at finde fordelingens middeltal – eller konkret: Hvor mange stykker udførte hver arbejder i gennemsnit i løbet af de 30 timer?

Udregnes dette som et vejet aritmetisk gennemsnit, fås

$$\frac{454 \text{ stk.}}{28} = 16\tfrac{3}{4} \text{ stk.}$$

Tegner vi fordelingen op som vist hosstående, ses, at vi har at gøre med en *skæv (usymmetrisk) fordeling*.

Sådanne skæve fordelinger er langt fra usædvanlige, og når man træffer på dem, indeholder de først og fremmest en opfordring til opløsning af iagttagelserne i grupper, der hver har sine specielle egenskaber. Hvis en sådan deling af materialet enten ikke er mulig eller ikke fører frem til symmetriske eller „næsten-symmetriske“ fordelinger for hver af de enkelte grupper, kan man forsøge at finde frem til en sådan passende *omformning* (transformation) af den statistiske variable (i exemplet antal stykker) at man med bedst mulig tilnærmelse når frem til en såkaldt *typisk* eller *normal fordeling*. Årsagen til bestræbelser i denne retning er – udtrykt meget populært – at det aritmetiske gennemsnit er uegnet som bekrivende element, når det drejer sig om skæve fordelinger. Kan man derimod få den skæve fordeling omdannet til en exponentiel fordeling, er det velkendt, at den matematiske form for denne fordeling kun



indeholder 2 „faste" parametre, som netop er fordelingens aritmetiske middeltal og spredning. Ad denne vej kan man da opnå en beskrivelse af den oprindelige skæve fordeling ved hjælp af den valgte transformation i forbindelse med aritmetisk middeltal og spredning i den transformerede fordeling.

I exemplet ovenfor foreligger en meget „nærliggende" transformationsmulighed. I stedet for at måle ydedygtigheden ved antal stykker pr. 30 timer, kunne man se på, hvor mange timer, som det tager at fremstille eet stykke. Når 2 arbejdere udførte hver 10 stk. på 30 timer svarer dette til, at hver af disse 2 arbejdere var 3 timer om at fremstille eet stykke. De 7 arbejdere, der hver lavede 12 stykker på 30 timer, brugte hver 2½ time om hvert stykke o. s. v.

Vi får da den nye statistiske variable y, der fremkommer af den oprindelige x gennem denne enkle *transformation*

$$y = \frac{30}{x}$$

og den transformerede fordeling bliver nu:

Antal arbejdere (a)	Timer om at lave 1 stk. (y)	$a \cdot y$
2	3	6
7	2,5	17,5
10	2	20
7	1,5	10,5
2	1	2
28		56

Det ses let uden tegning, at der nu er fremkommet en symmetrisk fordeling. Det aritmetiske middeltal bliver

$$\bar{y} = \frac{56}{28} = 2$$

d. v. s. hver arbejder var gennemsnitligt 2 timer om at fremstille 1 stk. Regner man nu tilbage til den oprindelige statistiske måleenhed, findes, at den gennemsnitlige produktion pr. arbejder *ikke* bliver 16¾ stk., men 15 stk., nemlig 30:2. Man kommer altså ikke til det samme. Det kan også udtrykket på den måde, at man *ikke* kommer til det aritmetiske middeltal i den transformerede fordeling (der blev 2) ved en transformation af det aritmetiske middeltal i den oprindelige fordeling, d. v. s. 30 divideret med 16½. Man vil kun komme til samme resultat i det specielle tilfælde, hvor transformationen er lineær[5]). Gennemsnittet på 15 stk., der fandtes af den transformerede fordeling (ved „re-transformation" af gennemsnittet på 2 timer), kunne imidlertid godt være fundet direkte ud fra den oprindelige skæve fordeling, men blot ikke som et *aritmetisk* gennemsnit. Tallet 15 findes som et vejet *harmonisk* gennemsnit anvendt på den oprindelige fordeling, nemlig

$$\frac{28}{\frac{2}{30} + \frac{7}{12} + \frac{10}{15} + \frac{7}{20} + \frac{2}{30}} = 15$$

Det kan næppe overraske, at man i exemplet her med den „reciproke" transformation finder, at det harmoniske gennemsnit er, hvad man kunne kalde *korrespondensmiddeltallet* til det aritmetiske middel i den transformerede fordeling.

[5]) Beviset herfor findes hos A. Hald, Statistical Theory with Engineering Applications, 1952, pag. 106.

Ikke helt sjældent kan en skæv fordeling omdannes til en exponentiel fordeling ved den simple logaritmiske transformation

$$y = \log x$$

Korrespondensmiddeltallet bliver da det geometriske gennemsnit.

Hvis transformationen er

$$y = x^2$$

bliver korrespondensmiddeltallet det kvadratiske gennemsnit o. s. v.

VIII.

Man kan under „samme hat" sammenfatte en lang række typer af gennemsnit på følgende måde:

$$\text{Gennemsnittet}^{6)} \text{ af } x \text{ og } y = \left(\frac{x^{\alpha+\beta} + y^{\alpha+\beta}}{x^{\alpha} + y^{\alpha}}\right)^{\frac{1}{\beta}}$$

Der kan „hægtes" så mange tal på, det skal være, udover de 2, som er vist i formlen. Og skal gennemsnittet vejes, anbringes blot de respektive vægte som faktorer ved leddene i både tæller og nævner. Ovenfor omtalte typer „trækkes ud af hatten" ved indsættelse af følgende talværdier for α og β:

α	β	
0	1	Aritmetiske
0	$\to 0$	Geometriske
0	$\div 1$	Harmoniske
0	2	Kvadratiske
0	½	Kvadratrodsgennemsnittet
0	$\div 2$	Harmoniske af anden orden
1	1	Kontraharmoniske

Det kontraharmoniske, der dog ikke har været nævnt før, ser således ud:

$$G = \frac{x^2 + y^2}{x + y}$$

α og β kan være hele eller brudne tal, positive, nul eller negative, og formlen indeholder derfor en rig mangfoldighed af gennemsnitstyper.

6) Om denne generalisering og beviset for, at det geometriske genemsnit er „indeholdt" i formlen, se nærværende forfatter Om forbrugsteori og teorien om indextal, NTTØ 1942.

Alligevel existerer der endnu talrige andre typer gennemsnit udenfor denne mangfoldighed. Som et enkelt exempel kan nævnes

$$G = \frac{x - y}{\ln x - \ln y}$$

altså tallenes *differens* divideret med differensen imellem den naturlige logaritme til tallene. Formlen finder anvendelse i visse teknisk-fysiske relationer.

For fagstatistikeren og mange andre, der mere „fagmæssigt" kommer i kontakt med gennemsnitsberegninger, indeholder disse linier ingen nyheder. Men vi benytter os jo alle i ny og næ af gennemsnitsbetragtninger i forbindelse med dagliglivets foreteelser, og hensigten med exemplerne ovenfor har blot været at give nogle få illustrationer, der måske kunne vække til eftertanke og vise, at den „dagligdags" gennemsnitsopfattelse ikke altid slår til. Gennemsnittet af 2 og 8 kan efter omstændighederne være meget andet end akkurat 5!

Exemplernes tal kunne øges uden ende; helt uomtalt har været de komplicerede gennemsnitsproblemer, der rejser sig i forbindelse med beregning af pristal og andre indextal.

ERHVERVSØKONOMISK TIDSSKRIFT. Pris pr. hefte kr. 5,00, pr. årgang (4 hefter) kr. 18,00. HOVEDREDAKTION: Professor, dr. polit. Bjarke Fog, Handelshøjskolen, København V., telf. Nora 9260, privat Søborg 1440. AARHUS REDAKTION (nr. 3): Professor Svend Fredens, Aarhus Universitet, Aarhus, telf. Aarhus 34311, privat Aarhus 77916. Lektor, statsautoriseret revisor G. Graversen, Handelshøjskolen, Aarhus, telf. Aarhus 35011, privat Aarhus 40365. KØBENHAVNS REDAKTION (nr. 1, 2 og 4): Lektor, cand. merc. O. Loff, Handelshøjskolen, København V., telf. Nora 9260, privat Hellerup 2160. Lektor, kontorchef P. P. Sveistrup, Københavns Universitet, København K., telf. privat 87 09 14. REDAKTIONSSEKRETÆR Vidensk. ass., cand. oecon. Erik Johnsen, Handelshøjskolen, København V., telf. Nora 9260, privat Birkerød 2141. EKSPEDITION OG ADMINISTRATION: Foreningen af Danske Erhvervsøkonomer, landsretssagfører Axel Jacobsen, Lyngby Hovedgade 57 B, Kongens Lyngby, telf. 87 75 25. Trykt hos

Hansen & Andreasen, Godthåbsvej 22, København F.