THORKILD STEENBERG
BIOLOGI
KOMPENDIUM

B I O L O G I

Kugletusindben på blad

BIND II
TALBEHANDLING
&
REGNEARK

Indhold

	Indledning		*side 1*
1	**Resultatbehandling**		
	I	*Skemaopstilling*	*side 3*
	II	*Variation og spredning*	*side 4*
	III	*Grafisk vurdering*	*side 6*
2	**Statistisk analyse**		
	I	*Test af normalfordeling*	*side 7*
	II	*F-test og t-test*	*side 7*
	III	*Variansanalyse*	*side 10*
	IV	*Lineær tilpasning*	*side 12*
3	**Regneark**		
	I	*Excel*	*side 17*
	II	*QuattroPro*	*side 29*
	Register		*side 43*

□ *Thorkild Steenberg* □ *Århus Akademi* □

3. udg. 2003

Indledning

Resultater af biologiske forsøg eller undersøgelser vil ofte have form af en række måledata eller beregninger på basis af målinger.
Hvordan overskues og behandles sådanne forsøgsresultater?
Ofte ønsker man tillige svar på spørgsmålet: "Er der forskel på resultaterne af denne og hin måling?"

Hæftet beskriver opstilling af resultater i tabeller, import af data, beregning af gennemsnit og spredning, tegning af fordelingsdiagram, statistisk test af om to gennemsnit er ens, statistisk test af om flere gennemsnit er ens og tilpasning af data til en ret linie.

Desuden er der vejledning til brug af regneark (Excel og QuattroPro) til beregninger og statistisk analyse.

□ □ □

Resultatbehandling

I Oversigtsskema og gennemsnitsberegning

To stikprøver på hver 50 stk grønne bønner og gule bønner vejes for at kunne besvare følgende spørgsmål:

1 *Hvad er gennemsnitsvægten af grønne bønner og gule bønner?*
2 *Er der forskel på bønnerne med hensyn til vægt?*

Måledata indtastes i et regnearkskema og gennemsnitsværdier beregnes (figur 1). Se i afsnittet om regneark side 15ff, hvordan et regneark generelt bruges og hvor-dan gennemsnit beregnes.

Er der i stedet tale om data i fil-format (fx på diskette) - typisk udtræk fra databaser eller lignende - indlæses disse data direkte i regnearket (se side 15ff for en beskrivelse af hvordan det gøres i de forskellige regneark).

	Grønne bønner		Gule bønner	
	vægt i mg			
	161	127	230	160
	248	92	180	170
	169	143	240	240
	108	157	200	220
	221	124	200	220
	102	168	230	110
	251	234	190	180
	182	159	230	140
	148	109	220	150
	180	182	220	180
	128	182	310	210
	161	152	190	180
	205	159	270	250
	178	129	210	150
	178	153	230	240
	164	153	250	170
	136	138	310	210
	153	136	150	250
	146	155	150	270
	215	164	190	240
	116	307	190	280
	161	183	150	190
	119	141	190	220
	108	180	200	180
	185	133	180	250
Gns	**162**		**207**	

Figur 1
Vægte for grønne og gule bønner; nederst i tabellen er gennemsnitsvægtene beregnet.

Svaret på spørgsmål 1 er gennemsnitsvægtene i figur 1:
162 mg for grønne bønner og 207 mg for gule bønner.

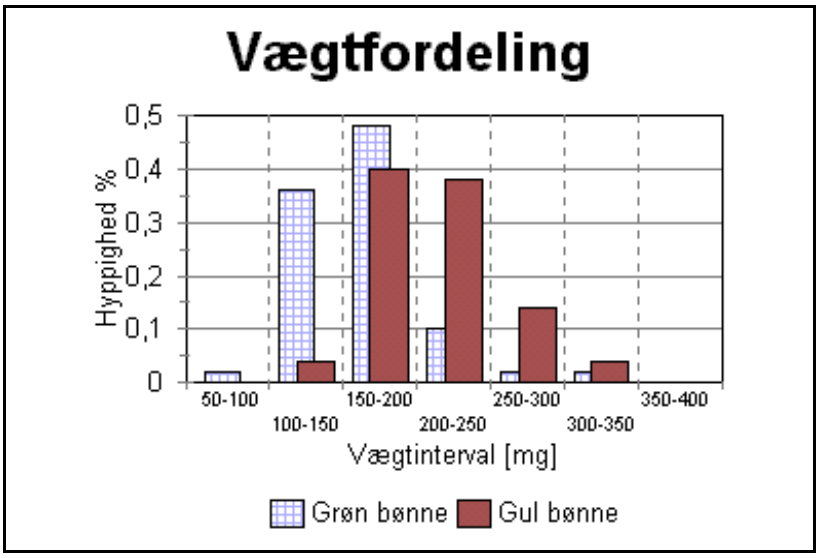
II Resultaternes variation

Spørgsmål 2 kan derimod ikke umiddelbart besvares. Om der er reel forskel på gennemsnitsværdierne afhænger af hvor "godt" de er bestemt, dvs. hvor stor variation der er imellem de enkelte måleresultater.

Det kan give et oerblik at tegne søjlediagrammer (histogrammer) over resultaternes fordeling.
Vælg et passende antal, lige store intervaller og optæl hvor mange værdier, der falder i hvert interval. Værdier der er lig med et intervals øvre grænse medtages i dette interval, medens værdier der er lig med nedre intervalgrænse medtages i forrige interval.
Figur 2 og 3 viser en fordelingstabel og det tilsvarende søjlediagram (se også side 15ff).

	Grøn bønne		Gul bønne	
Interval	Antal	Hyppighed	Antal	Hyppighed
50-100	1	0,02	0	0
100-150	18	0,36	2	0,04
150-200	24	0,48	20	0,4
200-250	5	0,1	19	0,38
250-300	1	0,02	7	0,14
300-350	1	0,02	2	0,04
350-400	0	0	0	0
Sum	50	1	50	1

Figur 2.
Eksempel på fordelingstabel. Intervalhyppighederne er angivet som antal og procent.



Figur 3
Diagram med vægtintervaller afsat på x-aksen og intervalhyppighed i % afsat på y-aksen.

Resultaterne fordeler sig mere eller mindre jævnt på begge sider af en gennemsnitsværdi. Hvis man forestiller sig et ideelt, meget stort forsøgsmateriale ville resultaterne fordele sig symmetrisk omkring gennemsnitsværdien og udgøre det man kalder en normalfordeling.

Det faktiske måleresultat kan så betragtes som en stikprøve af dette ideelle, normalfordelte forsøgsmateriale, og det man behøver er et mål for stikprøvens pålidelighed som udtryk for denne fordeling.

Beregner man den gennemsnitlige afvigelse mellem måleværdierne og gennemsnittet fås resultaternes **spredning (σ)**:

$$\sigma = \sqrt{\frac{\sum (\bar{x} - x_i)^2}{n-1}} \quad ; \quad \left[\begin{array}{l} \Sigma = \textit{summationstegn} \\ \bar{x} = \textit{gennemsnit} \\ x_i = \textit{måleværdi} \\ n = \textit{antal måleværdier} \end{array} \right]$$

Spredningen udtrykker resultaternes "samling" om gennemsnittet. Jo mindre σ er, des bedre er gennemsnitsværdien bestemt og man kan så tillade sig at betragte to gennemsnit som forskellige, hvis deres respektive spredningsintervaller (gns -σ til gns + σ) ikke overlapper eller kun overlapper lidt.
I figur 4 er resultaterne vist med udregnet spredning.

	Grønne bønner	Gule bønner
Gns	162	207
Spr	41	43

Figur 4. Beregnet gennemsnit og beregnet spredning

Man har valgt $\bar{x} \pm \sigma$ som en standard resultatangivelse.

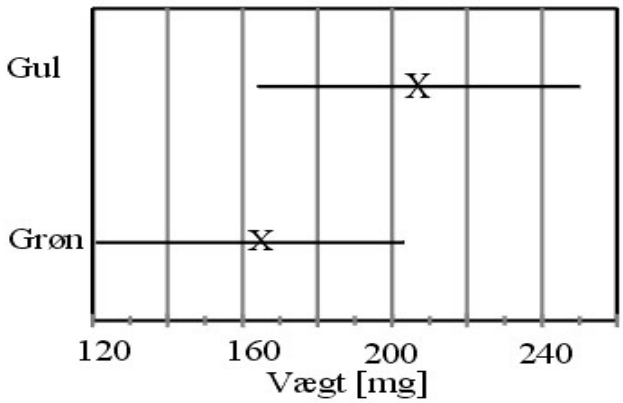
Svaret på spørgsmålet bør derfor korrekt skrives:

Grønne bønner: 162 ± 41 mg Gule bønner: 207 ± 43 mg

□ □ □

III Grafisk vurdering af resultater

Svaret på spørgsmål 2 kan man nu få ved at sammenligne spredningsintervallerne (gennemsnit - σ til gennemsnit + σ; figur 6) for forsøgsresultaterne.



Figur 5.
Diagram der viser spredningsintervaller for måleresultarer. (X markerer gennemsnit)

Spredningsintervallerne overlapper noget; men ikke mere end at det må kunne påståes, at gennemsnitsvægtene er forskellige.

□ □ □

Statistisk analyse

Længere end ovenstående kan man ikke komme uden en egentlig statistisk analyse.

I Test af normalfordeling

Den grundlæggende forudsætning bør testes: nemlig at der er tale om normalfordelte værdier i en stikprøve.

Ud fra fordelingstabellerne (side 4) beregnes den kumulerede hyppighed ved at fordelingshyppighederne (%) lægges sammen løbende (1. felt alene, 2. felt lægges sammen med 1. felt, 3. felt lægges sammen med 1. og 2. felt, etc). Disse kumulerede hyppigheder afsættes derefter i et digram på statistikpapir (normalfordelingspapir). Ligger de afsatte punkter jævnt om - eller på - en ret linie er der tale om en normalfordeling.

II Afgørelse af om der er forskel på gennemsnitsværdier

Her må man skelne mellem 1) tilfælde hvor kun to gennemsnitsværdier skal sammenlignes og 2) tilfælde hvor flere end to gennemsnitsværdier skal sammenlignes.

II A Sammenligning af to middelværdier (*t*-test)

Den statistiske test der anvendes til sammenligning af to middelværdier (***t*-test**) har som forudsætning, at varianserne (**var** = σ^2) for de to stikprøver, der skal sammenlignes, er ens. Derfor må det først afgøres, om de foreliggende varianser er tilstrækkeligt ens.

II A-1 Denne test kaldes en *F*-test

$$F = \frac{\text{største varians}}{\text{mindste varians}} \; ; \begin{bmatrix} m = \text{antal målinger i tælleren} - 1 \\ n = \text{antal målinger i nævneren} - 1 \end{bmatrix}$$

Hvis $\sigma_1^2 = \sigma_2^2$ er $\boldsymbol{F} = 1$.

Det vil sige, hvis *F* værdien er tilstrækkeligt tæt på 1, er der ikke noget, der strider mod at acceptere, at de to stikprøvevarianser er ens.

F værdien kan slåes op i en tabel (m og n kaldes frihedsgrader), og sandsynligheden for at varianserne er ens kan aflæses. Alternativt kan man med de indbyggede statistiske funktioner i regnearket beregne ***F***-værdien og sandsynligheden på én gang (eksempel figur 6, se fremgangsmåde i regnearksafsnittet side 15ff).

F-test: Varianser for to stikprøver		
	Grøn	***Gul***
Middelværdi	161,66	207,4
Varians	1705,3718	1848,20408
Observationer	50	50
df	49	49
F	1,0837543	
P(F<=f) enhalet	0,4227166	
F-kritisk enhalet	1,4463807	

Figur 6 Eksempel på F-test v.h.a. regneark

Hvis ***F*** er mindre end den kritiske værdi - det samme som at sandsynligheden $P(\boldsymbol{F} \leq f) \geq 0.05$ - accepteres det, at varianserne er ens.
I figur 6 sammenlignes varianserne for de to bønner ved en ***F***-test:
***F* er mindre end den kritiske værdi, dvs det accepteres at varianserne er ens.**

II A-2 Derefter fortsættes med den egentlige sammenligning af middelværdierne: en *t*-test

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\left(\frac{1}{n_x} + \frac{1}{n_y}\right) \frac{\sum x_i^2 - n_x\bar{x}^2 + \sum y_j^2 - n_y\bar{y}^2}{n_x + n_y - 2}}} \; ; \quad \left| \begin{array}{l} \bar{x} \text{ og } \bar{y} = \text{gns. 1 og 2} \\ n_x \text{ og } n_y = \text{antal data} \\ x_i \text{ og } y_j = \text{måleværdier} \end{array} \right|$$

Hvis $\boldsymbol{gns_1} = \boldsymbol{gns_2}$ er $\boldsymbol{t} = 0$.

Det vil sige, hvis $|t|$-værdien er tilstrækkeligt tæt på 0, er der ikke noget, der strider mod at acceptere, at de to gennemsnit er ens.

t-værdien kan slåes op i en tabel ($n_x + n_y - 2$ = antal frihedsgrader) og sandsynligheden for at gennemsnittene er ens kan aflæses; men her er det absolut en fordel at bruge regnearkets statistiske funktioner. (se fremgangsmåde i afsnittet om regneark side 15ff).

t-Test: to stikprøver med ens varianser		
	Grøn	***Gul***
Middelværdi	161,66	207,4
Varians	1705,3718	1848,20408
Observationer	50	50
Samlet varians	1739,205	
Hypotetisk middelforskel	0	
df	48	
t	-4,025227	
P(T<=t) enhalet	0,0001006	
t-kritisk enhalet	1,6772242	
P(T<=t) tohalet	0,0002013	
t-kritisk tohalet	2,0106348	

Figur 7 *Eksempel på t-test vha. regneark*

Hvis $|t|$ er mindre end den kritiske værdi - det vil sige det samme som at sandsynligheden $P(T \leq t) \geq 0{,}05$ - accepteres, at gennemsnittene er ens.

I figur 7 sammenlignes middelværdierne for grønne bønner og gule bønner ved en ***t***-test:

$|t|$ er større end den kritiske værdi, dvs de to middelværdier er ikke ens.

□ □ □

II B Sammenligning af flere middelværdier (variansanalyse)

Man kan bruge en *variansanalyse* til at teste hypotesen at et antal stikprøver har det samme gennemsnit. Variansen mellem grupper sammenlignes med variansen inden for grupperne; der anvendes også her en ***F***-test:

$$F = \frac{\dfrac{\sum_1^G (\overline{x_g} - \overline{x})^2}{(G-1)}}{\dfrac{\sum_1^G \sum_1^{n_g} (x - \overline{x_g})^2}{(N-G)}} \; ; \qquad \begin{array}{l} x = \textit{enkelte data} \\ \overline{x_g} = \textit{gruppegns.} \\ G = \textit{antal grupper} \\ N = \textit{total antal} \\ \overline{x} = \textit{total gns} \\ n_g = \textit{antal data i gruppen} \end{array}$$

Tælleren udtrykker gruppegennemsnittenes afvigelse fra totalgennemsnit og nævneren udtrykker de enkelte datas afvigelse fra deres gruppegennemsnit. Hvis de to variansudtryk er ens er ***F*** = 1.

Det vil sige, at hvis *F*-værdien er tilstrækkeligt tæt på 1, er der ikke noget der strider mod at acceptere, at stikprøvegennemsnittene er ens.

Hvis ***F*** er mindre end den kritiske værdi - det samme som at sandsynligheden P(***F***≤ f) ≥ 0,05 - accepteres det at gennemsnittene er ens.

Afstand mellem træer i seks prøveflader: 1-4: 25 kvm; 4-6 100 kvm						
	1	2	3	4	5	6
	2,300	0,450	2,630	1,200	2,300	1,750
	2,800	2,450	2,630	1,700	2,300	1,000
		1,000	1,220	0,710	1,900	1,000
		0,450	1,220	3,740	2,270	2,300
		0,450		0,870	2,000	2,500
		0,550		2,500	1,900	1,800
				1,900	3,470	1,800
				1,400	1,550	2,280
					0,800	2,870
					0,800	1,200
						0,830
						0,830
gns	2,550	0,892	1,925	1,753	1,929	1,680
spr	0,354	0,793	0,814	0,988	0,779	0,705

Figur 8. Afstandsmålinger fra seks prøveflade skal sammenlignes

I figur 8 er vist målinger af afstanden mellem træer i seks prøveflader. Er de seks prøveflader sammenlignelige, således at de kan betragtes som en stor prøveflade? Her skal bruges en variansanalyse. Resultatet af analysen ses i figur 9.

Variansanalyse - enkelt faktor

RESUME

Grupper	*Antal*	*Sum*	*Gennemsnit*	*Varians*
Kolonne 1	2	5,1	2,55	0,125
Kolonne 2	6	5,35	0,8916667	0,628417
Kolonne 3	4	7,7	1,925	0,6627
Kolonne 4	8	14,02	1,7525	0,975507
Kolonne 5	10	19,29	1,929	0,607321
Kolonne 6	12	20,16	1,68	0,496982

Beregning

Variationskilde	*SK*	*fg*	*MK*	*F*	*P-værdi*	*F krit*
Mellem grupper	6,118024	5	1,2236049	1,913841	0,116175	2,477165
Inden for gruppe	23,01642	36	0,6393451			
I alt	29,13445	41				

Figur 9. Variansanalyse udført på resultaterne i figur 8

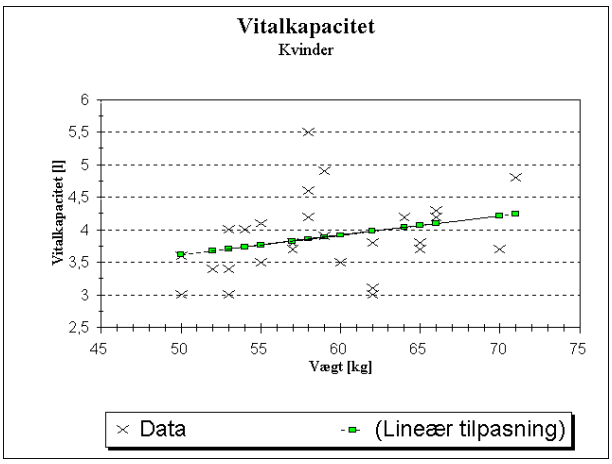
F er mindre end den kritiske værdi; det vil sige at gennemsnitsværdierne i eksemplet er sammenlignelige.

□ □ □

III Lineær tilpasning (regression)

Ofte har man i biologiske forsøg samtidige målinger af to eller flere data, som man vil undersøge den indbyrdes sammenhæng imellem.

Figur 10 viser som eksempel målinger af vægt og vitalkapacitet (dvs forskel mellem maksimal indånding og maksimal udånding) hos kvinder.
Målingerne er afsat i et x-y diagram (se regnearksafsnit side 15ff) og opgaven består i at bestemme et udtryk for den rette line, der bedst muligt tilpasses punktmængden.



Figur 10 Vitalkapacitet som funktion af vægt hos kvinder

Man vælger linien således, at summmen af kvadratet på afstanden mellem de beregnede y-koordinater og de faktiske y-koordinater bliver mindst mulig:

$$\sum (y_i - y_{bi})^2 \; ; \quad \begin{bmatrix} y_i = \textit{målt y-koordinat} \\ y_{bi} = \textit{beregnet y-koordinat} \end{bmatrix}$$

Metoden kaldes *lineær tilpasning* efter mindste kvadraters metode.
Ligningen for linien er $y = \boldsymbol{\alpha} x + \boldsymbol{b}$ og hældningskoeeficienten $\boldsymbol{\alpha}$ og skæringspunktet med y-aksen $\boldsymbol{b}$ kan beregnes efter følgende udtryk:

$$\alpha = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \; , \quad b = \bar{y} - \alpha \bar{x} \; ; \qquad \begin{bmatrix} \bar{x} = \textit{gns. af x-data} \\ \bar{y} = \textit{gns. af y-data} \\ x_i, y_i = \textit{datasæt} \end{bmatrix}$$

Regnearket kan beregne hældningskoeeficent og skæringspunkt med y-aksen (se fig 10 og fremgangsmåde i regnearksafsnittet side 15ff).

Desuden får man en beregning af korrelationskoefficienten ***r***, som angiver graden af indbyrdes afhængighed mellem de to variable:

r = 0 ingen afhængighed til ***r*** = 1 total afhængighed.

Det er vigtigt at understrege, at der ikke nødvendigvis er tale om en årsagssammenhæng mellem de to variable selv om korrelationskoefficienten er høj.

Sandsynligheden for at en ***r***-værdi er forskellig fra 0 kan findes med en en ***t***-test:

$$t = \sqrt{\nu\, r^2 / (1 - r^2)}\ ; \quad \nu = n - 2$$

Hvis $t > t_k$ (alternativt $0{,}025 \le P(T \le t) \le 0{,}0975$) kan det ikke accepteres at ***r*** er lig med 0 (dvs. ***r*** signifikant forskellig fra 0).
Se tabel over udvalgte værdier af t_k og n side 39.

Regressionsdata for eksempel i figur 10			
Konstant			2,129942
Standardafv. for Y-estimat			0,577577
Kvadratet på R			0,085179
Antal observationer			28
Frihedsgrader			26
X-koefficient(er)		0,029788	
Standardafv. for koefficient		0,019145	

Figur 11 Eksempel på lineær tilpasning vha regneark

Ligningen for eksemplet i figur 10 bliver ved indsætning af værdierne i figur 11:

$$\text{vitalkap}_{\text{kvinder}} = 0{,}0298 * \text{vægt}_{\text{kvinder}} + 2{,}1299; \quad r^2 = 0{,}085179;$$

Udføres en ***t***-test på korrelationskoefficienten som beskrevet ovenfor fås at ***r*** ikke adskiller sig fra 0; det vil sige at datamaterialet ikke tilstrækkeligt til at vise en sammenhæng mellem vitalkapacitet og vægt.

□ □ □

Brug af regneark til beregninger, statistik og grafisk afbildning

Afsnittet indeholder en beskrivelse af de vigtigste regnearksdefinitioner og en vejledning til anvendelse af regnearksfunktioner til talbehandling og simpel statistisk analyse.

Desuden vejledning til import af datasæt fra databaser eller datafangstmoduler og vejledning til at oprette diagrammer over datamaterialet.

Vejledningen findes til følgende udgaver:

- Excel 2000 side 17
- Excel 97 side 23
- QuattroPro 8 side 29
- QuattroPro 5 side 34

Vejledningen til QuattroPro 5 indeholder lidt flere detaljer end de tre øvrige regnearksudgaver, så dette afsnit kan supplere de øvrige.

□ □ □

I Excel 2000

- Definitioner side 18
- Import af datafiler side 19
- Beregning af gennemsnit og spredning side 20
- Statistik side 20

Se også QuattroPro 5 afsnittet (side 34) for flere detaljer.

□ □ □

I Definitioner

BLOK

En eller flere celler eller rækker eller kolonner

MARKER BLOK

Peg på øverste celle, hold venstre museknap nede og træk til nederste celle.
Slip museknappen (blokken er markeret med en afvigende farve).
Eller peg på øverste celle, hold ***<skift>*** nede og tryk dernæst ***<pil op/ned>*** eller ***<pil højre/venstre>*** sålænge det er nødvendigt.
Skal ikke sammenhængende blokke markeres holdes <***crtl***>- tasten nede samtidig med at der markeres med mus eller <***skift***> og piletaster.

Er det en hel kolonne, der skal markeres, kan det gøres på en gang ved at trykke samtidig på ***<crtl><skift><pil ned eller op>***.

NAVNGIV BLOK

Når en blok er markeret som beskrevet ovenfor kan den navngives, således at man blot angiver navnet, hver gang den samme blok skal bruges i formlerog lign.
Det er især en fordel ved lange uoverskuelige kolonner.
Vælg ***<Indsæt/Navn/Definer>*** i menuen. Skriv et passende navn og tryk ***<retur>***.

ÆNDRE BLOKKENS EGENSKABER

Klik med højre museknap på en celle eller blok, vælg ***<formater celler>***.
Her kan vælges skrifttype, skriftstørrelse, fed, kursiv, farve, format, linietegning, m.m.

KOLONNEBREDDE

Kolonnebredden kan automatisk afpasses efter det bredeste celleindhold med menupunktet: ***<Formater/Kolonne/Autotilpas>***.

KOPIÉR

Markér celle eller celler der skal kopieres, tryk på ***<crtl+c>*** eller højreklik og vælg ***<kopier>*** .
Marker øverste celle, der skal kopieres til og tryk på ***<crtl+v >*** eller højreklik og vælg ***<indsæt>.***

FLYT

Peg på cellen eller blokken med venstre museknap. Hold knappen indtrykket og flyt markøren ud til kanten indtil en lille pil viser sig.
Flyt cellen eller - med museknappen indtrykket - til den nye placering.
Slip museknappen.

SUMMERING AF RÆKKER ELLER KOLONNER: <AUTOSUM> ∑ -IKONEN

Afmærk kolonnen eller rækken + én ekstra celle.
Tryk på ∑-ikonen.

REFERENCE OG ABSOLUT REFERENCE

Alle beregninger foretages ved at refere til andre celler: fx indhold i B2 "=
A2*2" (se eksempel nedenfor).
Denne formel kan kopieres nedad i B-kolonnen så alle celler får den samme beregning.
Tilsvarende er celle B6 "=B2+B3+B4+B5 ".

Hvis der beregnes % i en kolonne på basis af værdier i kolonnen før og en sumværdi i cellen under denne, skal der anvendes absolut reference til sum-cellen; dvs cellen markeres med $ foran tallet, fx B$6 som i eksemplet nedenfor.
Når en formel med absolut reference kopieres nedad til andre celler i samme kolonne divideres med tallet i sumfeltet i alle celler.

	A	B	C
1		Antal	Antal i %
2	17	34 (formel: “=A2*2")	“=B2/B$6"
3	16	32	“=B3/B$6"
4	17,5	35	“=B4/B$6"
5	15	30	“=B5/B$6"
6	Total	131	

II Import af datafiler

Filerne er såkaldte kommaseparerede filer - et format som databasedata eksporteres til andre systemer i.
Hver information (databasefelt) er adskilt fra andre med ‘ ,’(komma - deraf navnet).
Talværdier er angivet umiddelbart, medens tekst er omkranset af “ ”.
Tal lagres i internationalt format, dvs med decimalpunktum.

- I regnearket vælges ***<data/hent eksterne data/importer tekstfil>*** i menuen.
  Søg fil på “drev A:” og sæt filtype til “alle filer”, dvs *.*
  Find filen ***<a:\data.kom>*** eller ***<a:\standard.kom>***.
  Åben filen..

- 1. trin. Accepter 1. trin i tekstimportguiden. Tryk på ***<næste>***.
  2. trin. Vælg “komma” som afgrænser og fjern afkrydsning i “tabulator”. Tryk på ***<næste>***.
  3.trin. Tryk ***<avanceret>***: Vælg “punktum” som decimalseparator og “komma” som tusindadskiller. Tryk <***OK***>
  Accepter resten af 3. trin. Tryk på ***<udfør>***.

- Accepter placering i eksisterende regneark og øverste venstre hjørne.
  Tryk ***<OK>***
  Data indlæses i regnearket klar til videre behandling.

Overskrifter på kolonner og tilretning.

Den store udskrift fra databasen (standard.kom) kan forsynes med kolonneoverskrifter.
Slet kolonner G til og med N.
Importer på samme måde som 2 og 3 ovenfor overskrifttekster fra filen st-tekst.kom.
Marker og kopier tekstlinien. Skift til regnearksfanen med data. Indsæt en linie øverst på siden og indsæt tekstlinen her.
Slet evt. kolonner der ikke skal bruges så regnearket bliver lidt mere overskueligt.
Desuden kan det være en fordel at flytte alle data for mænd til et regnearksblad for sig selv:
Marker blokken med alle data for mænd som beskrevet ovenfor.

Højreklik i blokken og tryk på ***<klip>***.
Marker øverste venstre celle i en nyt ark med musen, højreklik og vælg ***<indsæt>***
Begge ark kan dernæst navngives henholdsvis kvinder og mænd

III Beregninger af gennemsnit og spredning

Beregning af gennemsnit for en blok: =MIDDEL(blok)

Vælg en fri celle under blokken.. Skriv =MIDDEL("første celle":"sidste celle" eller bloknavn) og tryk <***retur***>.
Blokken kan også markeres med musen eller med <***skift***> og piletasterne.

Ret dernæst antal decimaler til som ønskeligt og sæt en forklarende tekst ind først i samme linie.

Beregning af spredning for en blok: =STDAFV(blok)

Vælg en fri celle under blokken. Skriv =STDAFV("første celle":"sidste celle", eller bloknavn) og tryk <***retur***>.
Blokken kan også markeres med musen eller med <***skift***> og piletasterne.

Ret dernæst antal decimaler til som ønskeligt og sæt en forklarende tekst ind først i samme linie.

IV Statistik

Oprette fordelingstabeller til stolpediagrammer og tegning af stolpediagram.

1 Oprette fordelingstabeller:

Vælg et passende antal, lige store intervaller og skriv dem i stigende rækkefølge i en tom kolonne under eller til højre for de øvrige data.
Værdier der er lig med et intervals øvre grænse bliver medtaget i dette interval, medens værdier der er lig med nedre intervalgrænse bliver medtaget i forrige interval.

Opret en kolonne med øvre grænseværdier til højre for intervalkolonnen.

Vælg <***Funktioner/Dataanalyse/Histogram***> i menuen.
Der åbnes en dialogboks.

- Skriv datablokken i feltet "Inputområde" - dvs navn eller "første celle": "sidste celle" (blokken kan også udpeges ved at klikke på det lille ikon til højre i skrivefeltet).
- Skriv den netop oprettede grænseværdikolonne i feltet "Intervalområde".
- Marker "Outputområde" og skriv en cellekoordinat til højre for grænseværdikolonnen.
- Afkryds evt feltet "Kumulativ frekvens". Tryk <***OK***>

Regnearket beregner og opretter en tabel med intervalhyppighed.

2 Tilpasning:

- Kopier kolonnen med intervaller ind i den oprettede tabel i stedet for kolonnen med øvre grænseværdier.
- Summér hyppighederne i cellen under hyppighedskolonnen.
- Flyt kolonnen "Kumulativ %" en celle til højre.
- Brug den tomme kolonne til højre for hyppighed til at beregne hyppigheden i procent.
- Juster antal decimaler i kolonnerne.

3 Stolpediagram:

Marker % hyppighed-kolonnen og intervalkolonnen.
Vælg graf-ikonet i menuen. Der åbnes en række dialogbokse.

- Vælg "søjle" som diagramtype og den simple udgave øverst til venstre. Tryk ***<næste>***.
- Accepter indstillinger på næste side; tryk ***<næste>***.
- Titelfane: Skriv en passende overskrift og en forklaring til x-akse: fx højdeintervaller]x', x''] cm, og y-akse: hyppighed i % ;
- Forklaring fane: Fjern afkrydsning i "vis forklaring". Tryk ***<næste>.***
- Skal diagrammet anbriges på regnearkbladet eller i et vindue for sig selv. Tryk ***<udfør>***.

Stolpediagram over fordelingstabellen tegnes. Skal der yderligere redigering til gøres det ved at højreklikke forskellige steder i diagrammet (x-akse, y-akse, serier, etc) og ændre indstillingerne så de passer.

F-test. Test for at varianser i to stikprøver er ens

Vælg ***<Funktioner/Dataanalyse/F-test: dobbelt-stikprøve for varians>*** i menuen.
Der åbnes en dialogboks.

- Udfyld "område for variabel 1" og "område for variabel 2" - *dvs navn eller "Arknavn!første celle:sidste celle" (arknavn skal kun bruges hvis talværdierne er fordelt på flere faner).*
- Marker "outputområde" og skriv en cellekoordinat til højre for datablokken i feltet.
- Tryk ***<OK>***. F-test resultatet indsættes som en lille tabel i regnearket; juster kolonnebredde således at statistiktabellen står pænt.
- Aflæs resultat: hvis sandsynligheden $P(F \leq f) \geq 0{,}05$ er der ikke noget der strider mod at acceptere, at varianserne er ens.

t-test. Test for at to gennemsnitsværdier er ens

Vælg ***<Funktioner/Dataanalyse/t-test: to stikprøver med ens varians>*** i menuen.
Der åbnes en dialogboks.

- Udfyld "område for variabel 1" og "område for variabel 2" - dvs *navn eller "Arknavn!første celle:sidste celle" (arknavn skal kun bruges hvis talværdierne er fordelt på flere faner).*
- Hypotese for forskel i middelværdi: skriv 0 (dvs der testes for at gennemsnitsværdierne er ens).
- Marker "outputområde" og skriv en koordinat for en celle til højre for datablokken i feltet.
- Tryk ***<OK>***. t-test resultatet indsættes som en lille tabel i regnearket; juster kolonnebredde således at statistiktabellen står pænt.
- Aflæs resultat: hvis sandsynligheden $P(T \leq t) \geq 0{,}05$ (tohalet) er der ikke noget der strider mod at acceptere, at gennemsnittene er ens.

Variansanalyse

Vælg ***<Funktioner/Dataanalyse/ANAVA: enkeltfaktor>*** i menuen. Der åbnes en dialogboks.

- Skriv datablokken i feltet "Inputområde" - dvs navn eller "første celle": "sidste celle" (blokken kan også udpeges ved at klikke på det lille ikon til højre i skrivefeltet).

- Marker "Outputområde" og skriv en cellekoordinat til højre for eller under datablokken.
- Tryk ***<OK>***. Variansanalysen indsættes som to tabeller i regnearket.
- Aflæs sandsynligheden (P-værdi) eller F-værdierne. Hvis sandsynligheden er større end 0,05 (eller F er mindre end $F_{kritisk}$) er der ikke noget der strider mod at acceptere, at gennemsnittene er ens.

Lineær tilpasning (regression)

Vælg ***<Funktioner/Dataanalyse/regression>*** i menuen. Der åbnes en dialogboks.

- Udfyld "input for Y-område" (dvs den afhængigt variable: fx vægt) med: *navn eller "første celle:sidste celle" eller de*

andre måder beskrevet ovenfor.

- Udfyld "input for X-område" (dvs den uafhængigt variable: fx højde) på tilsvarende måde
- Marker "outputområde" og skriv en koordinat for en celle til højre for datablokken i feltet.
- Tryk ***<OK>***. Regressionsresultatet indsættes som en (noget indviklet!) tabel i regnearket.
- Aflæs resultat: Det der skal bruges er koefficienter til x-variabel (dvs α) og skæring (dvs b) i forskriften for den rette linie:

$f(x) = \alpha x + b$ i eksemplet: vægt = α højde + b

□ □ □

II Excel 97

- Definitioner side 24
- Import af datafiler side 25
- Beregning af gennemsnit og spredning side 26
- Statistik side 26

Se også QuattroPro 5 afsnittet (side 34) for flere detaljer.

□ □ □

I Definitioner

BLOK

En eller flere celler eller rækker eller kolonner

MARKER BLOK

Peg på øverste celle, hold venstre museknap nede og træk til nederste celle.
Slip museknappen (blokken er markeret med en afvigende farve).
Eller peg på øverste celle, hold ***<skift>*** nede og tryk dernæst ***<pil op/ned>*** eller ***<pil højre/venstre>*** sålænge det er nødvendigt.
Skal ikke sammenhængende blokke markeres holdes <***crtl***>- tasten nede samtidig med at der markeres med mus eller <***skift***> og piletaster.

Er det en hel kolonne, der skal markeres, kan det gøres på en gang ved at trykke samtidig på ***<crtl><skift><pil ned eller op>***.

NAVNGIV BLOK

Når en blok er markeret som beskrevet ovenfor kan den navngives, således at man blot angiver navnet, hver gang den samme blok skal bruges i formlerog lign.
Det er især en fordel ved lange uoverskuelige kolonner.
Vælg ***<Indsæt/Navn/Definer>*** i menuen. Skriv et passende navn og tryk ***<retur>***.

ÆNDRE BLOKKENS EGENSKABER

Klik med højre museknap på en celle eller blok, vælg ***<formater celler>***.
Her kan vælges skrifttype, skriftstørrelse, fed, kursiv, farve, format, linietegning, m.m.

KOLONNEBREDDE

Kolonnebredden kan automatisk afpasses efter det bredeste celleindhold med menupunktet: ***<Formater/Kolonne/Autotilpas>***.

KOPIÉR

Markér celle eller celler der skal kopieres, tryk på ***<crtl+c>*** eller højreklik og vælg ***<kopier>*** .
Marker øverste celle, der skal kopieres til og tryk på ***<crtl+v >*** eller højreklik og vælg ***<indsæt>.***

FLYT

Peg på cellen eller blokken med venstre museknap. Hold knappen indtrykket og flyt markøren ud til kanten indtil en lille pil viser sig.
Flyt cellen eller - med museknappen indtrykket - til den nye placering.
Slip museknappen.

SUMMERING AF RÆKKER ELLER KOLONNER: <AUTOSUM> ∑ -IKONEN

Afmærk kolonnen eller rækken + én ekstra celle. Tryk på ∑-ikonen.

REFERENCE OG ABSOLUT REFERENCE

Alle beregninger foretages ved at refere til andre celler: fx indhold i B2 "=
A2*2" (se eksempel nedenfor).
Denne formel kan kopieres nedad i B-kolonnen så alle celler får den samme beregning.
Tilsvarende er celle B6 "=B2+B3+B4+B5 ".

Hvis der beregnes % i en kolonne på basis af værdier i kolonnen før og en sumværdi i cellen under denne, skal der anvendes absolut reference til sum-cellen; dvs cellen markeres med $ foran tallet, fx B$6 som i eksemplet nedenfor.
Når en formel med absolut reference kopieres nedad til andre celler i samme kolonne divideres med tallet i sumfeltet i alle celler.

	A	B	C
1		Antal	Antal i %
2	17	34 (formel: “=A2*2")	“=B2/B$6"
3	16	32	“=B3/B$6"
4	17,5	35	“=B4/B$6"
5	15	30	“=B5/B$6"
6	Total	131	

II Import af datafiler

Filerne er såkaldte kommaseparerede filer - et format som databasedata eksporteres til andre systemer i.
Hver information (databasefelt) er adskilt fra andre med ‘ ,’(komma - deraf navnet).
Talværdier er angivet umiddelbart, medens tekst er omkranset af “ ”.
Tal lagres i internationalt format, dvs med decimalpunktum.

I regnearket (excel 97) kan der ikke ændres talformatering; det skal gøres i Windows/ Kontrolpanel før og efter import.

- Vælg ***<start/ indstillinger/ kontrolpanel/ international/ tal>*** på skrivebordet.
  Sæt decimaltegn til punktum og tusindseparator til komma.
  Dernæst *<ok>*

- I regnearket vælges ***<filer/ åben fil>*** i menuen.
  Søg fil på “drev A:” og sæt filtype til “alle filer”, dvs *.*
  Find filen ***<a:\data.kom>*** eller ***<a:\standard.kom>***. Åben filen..

- 1. trin: Accepter 1. trin i tekstimportguiden. Tryk på ***<næste>***.
  2. trin: Vælg komma som afgrænser og fjern afkrydsning i tabulator. Tryk på ***<næste>***.
  3.trin: Accepter 3. trin. Tryk på ***<udfør>***.
  Data indlæses i regnearket.

- Skift til skrivebordet.
  Vælg ***<start/ indstillinger/ kontrolpanel/ international/ tal>*** på skrivebordet.
  Sæt decimaltegn til komma og tusindseparator til punktum. Dernæst ***<OK>***
  Skift tilbage til regnearket.

Herefter er data klar til videre behandling.

Overskrifter på kolonner og tilretning.

Den store udskrift fra databasen (standard.kom) kan forsynes med kolonneoverskrifter.
Slet kolonner G til og med N.
Importer på samme måde som 2 og 3 ovenfor overskrifttekster fra filen st-tekst.kom. Marker og kopier tekstlinien. Skift til regnearkvinduet med data. Indsæt en linie øverst på siden og indsæt tekstlinen her. Slet evt. kolonner der ikke skal bruges så regnearket bliver lidt mere overskueligt. Desuden kan det være en fordel at flytte alle data for mænd til et regnearksblad for sig selv: Marker blokken med alle data for mænd som beskrevet ovenfor. Højreklik i blokken og tryk på ***<klip>***. Vælg ***<Indsæt/Regneark>*** i menuen; marker øverste venstre celle i det nye ark med musen, højreklik og vælg ***<indsæt>***. Begge ark kan dernæst navngives henholdsvis kvinder og mænd.

III Beregninger af gennemsnit og spredning

Beregning af gennemsnit for en blok: =MIDDEL(blok)

Vælg en fri celle under blokken.. Skriv =MIDDEL("første celle":"sidste celle" eller bloknavn) og tryk <***retur***>.
Blokken kan også markeres med musen eller med <***skift***> og piletasterne.

Ret dernæst antal decimaler til som ønskeligt og sæt en forklarende tekst ind først i samme linie.

Beregning af spredning for en blok: =STDAFV(blok)

Vælg en fri celle under blokken. Skriv =STDAFV("første celle":"sidste celle", eller bloknavn) og tryk <***retur***>.
Blokken kan også markeres med musen eller med <***skift***> og piletasterne.

Ret dernæst antal decimaler til som ønskeligt og sæt en forklarende tekst ind først i samme linie.

IV Statistik

Oprette fordelingstabeller til stolpediagrammer og tegning af stolpediagram.

1 Oprette fordelingstabeller:

Vælg et passende antal, lige store intervaller og skriv dem i stigende rækkefølge i en tom kolonne under eller til højre for de øvrige data.
Værdier der er lig med et intervals øvre grænse bliver medtaget i dette interval, medens værdier der er lig med nedre intervalgrænse bliver medtaget i forrige interval.

Opret en kolonne med øvre grænseværdier til højre for intervalkolonnen.

Vælg <***Funktioner/Dataanalyse/Histogram***> i menuen. Der åbnes en dialogboks.

- Skriv datablokken i feltet "Inputområde" - dvs navn eller "første celle": "sidste celle" (blokken kan også udpeges ved at klikke på det lille ikon til højre i skrivefeltet).
- Skriv den netop oprettede grænseværdikolonne i feltet "Intervalområde".
- Marker "Outputområde" og skriv en cellekoordinat til højre for grænseværdikolonnen.
- Afkryds evt feltet "Kumulativ frekvens". Tryk <***OK***>

Regnearket beregner og opretter en tabel med intervalhyppighed.

2 Tilpasning:

- Kopier kolonnen med intervaller ind i den oprettede tabel i stedet for kolonnen med øvre grænseværdier.
- Summér hyppighederne i cellen under hyppighedskolonnen.
- Flyt kolonnen "Kumulativ %" en celle til højre.
- Brug den tomme kolonne til højre for hyppighed til at beregne hyppigheden i procent.
- Juster antal decimaler i kolonnerne.

3 Stolpediagram:

- Marker % hyppighed-kolonnen og intervalkolonnen.
- Vælg graf-ikonet i menuen. Der åbnes en række dialogbokse.
- Vælg "søjle" som diagramtype og den simple udgave øverst til venstre. Tryk <***næste***>.
- Accepter indstillinger på næste side; tryk <***næste***>.
- Titelfane: Skriv en passende overskrift og

en forklaring til x-akse: fx højdeintervaller]x', x''] cm, og y-akse: hyppighed i % ;

- Forklaring fane: Fjern afkrydsning i "vis forklaring". Tryk ***<næste>.***
- Skal diagrammet anbriges på regnearkbladet eller i et vindue for sig selv. Tryk ***<udfør>***.

Stolpediagram over fordelingstabellen tegnes. Skal der yderligere redigering til gøres det ved at højreklikke forskellige steder i diagrammet (x-akse, y-akse, serier, etc) og ændre indstillingerne så de passer.

F-test. Test for at varianser i to stikprøver er ens

Vælg ***<Funktioner/Dataanalyse/F-test: dobbelt-stikprøve for varians>*** i menuen.
Der åbnes en dialogboks.

- Udfyld "område for variabel 1" og "område for variabel 2" - *dvs navn eller "Arknavn!første celle:sidste celle" (arknavn skal kun bruges hvis talværdierne er fordelt på flere faner).*
- Marker "outputområde" og skriv en cellekoordinat til højre for datablokken i feltet.
- Tryk ***<OK>***. F-test resultatet indsættes som en lille tabel i regnearket; juster kolonnebredde således at statistiktabellen står pænt.
- Aflæs resultat: hvis sandsynligheden $P(F \leq f) \geq 0{,}05$ er der ikke noget der strider mod at acceptere, at varianserne er ens.

t-test. Test for at to gennemsnitsværdier er ens

Vælg ***<Funktioner/Dataanalyse/t-test: to stikprøver med ens varians>*** i menuen.
Der åbnes en dialogboks.

- Udfyld "område for variabel 1" og "område for variabel 2" - dvs *navn eller "Arknavn!første celle:sidste celle" (arknavn skal kun bruges hvis talværdierne er fordelt på flere faner).*
- Hypotese for forskel i middelværdi: skriv 0 (dvs der testes for at gennemsnitsværdierne er ens).
- Marker "outputområde" og skriv en koordinat for en celle til højre for datablokken i feltet.
- Tryk ***<OK>***. t-test resultatet indsættes som en lille tabel i regnearket; juster kolonnebredde således at statistiktabellen står pænt.
- Aflæs resultat: hvis sandsynligheden $P(T \leq t) \geq 0{,}05$ (tohalet) er der ikke noget der strider mod at acceptere, at gennemsnittene er ens.

Variansanalyse

Vælg ***<Funktioner/Dataanalyse/ANAVA: enkeltfaktor>*** i menuen. Der åbnes en dialogboks.

- Skriv datablokken i feltet "Inputområde" - dvs navn eller "første celle": "sidste celle" (blokken kan også udpeges ved at klikke på det lille ikon til højre i skrivefeltet).
- Marker "Outputområde" og skriv en cellekoordinat til højre for eller under datablokken.
- Tryk ***<OK>***. Variansanalysen indsættes som to tabeller i regnearket.
- Aflæs sandsynligheden (P-værdi) eller F-værdierne. Hvis sandsynligheden er større end 0,05 (eller F er mindre end $F_{kritisk}$) er der ikke noget der strider mod at acceptere, at gennemsnittene er ens.

Lineær tilpasning (regression)

Vælg ***<Funktioner/Dataanalyse/regression>*** i menuen.. Der åbnes en dialogboks.

- Udfyld "input for Y-område" (dvs den afhængigt variable: fx vægt) med: *navn eller "første celle:sidste celle" eller de andre måder beskrevet ovenfor.*
- Udfyld "input for X-område" (dvs den uafhængigt variable: fx højde) på tilsvarende måde

- Marker "outputområde" og skriv en koordinat for en celle til højre for datablokken i feltet.
- Tryk **<*OK*>**. Regressionsresultatet indsættes som en (noget indviklet!) tabel i regnearket.
- Aflæs resultat: Det der skal bruges er koefficienter til x-variabel (dvs α) og skæring (dvs b) i forskriften for den rette linie:

$f(x) = \alpha x + b$ i eksemplet: vægt = α højde + b

□ □ □

III QuattroPro 8

- Definitioner side 30
- Import af datafiler side 31
- Beregning af gennemsnit og spredning side 31
- Statistik side 31

□ □ □

I DEFINITIONER

BLOK

En eller flere celler eller rækker eller kolonner

MARKER BLOK

Peg på øverste celle, hold venstre museknap nede og træk til nederste celle.
Slip museknappen (blokken er markeret med en afvigende farve).
Eller peg på øverste celle, tryk <***skift F7***> og dernæst <***pil ned***> /<***pil højre***> sålænge det er nødvendigt.

ÆNDRE BLOKKENS EGENSKABER

Klik med højre museknap på en celle eller blok, vælg <***egenskaber***>.
Her kan vælges skrifttype, skriftstørrelse, fed, kursiv, farve, format, linietegning, m.m.

KOLONNEBREDDE

Kolonnebredden kan automatisk afpasses efter det bredeste celleindhold med ikonen: **< ↔ >.**

KOPIÉR

Markér celle eller celler der skal kopieres, tryk på <***kopiér***> ikonen. Marker øverste celle, der skal kopieres til og tryk på <***indsæt***> ikonen.

FLYT

Peg på cellen eller blokken med venstre museknap. Hold knappen indtrykket og vent et lille øjeblik, indtil en lille hånd viser sig. Flyt cellen eller blokken - med museknappen indtrykket - til den nye placering.
Slip museknappen.

SUMMERING AF RÆKKER ELLER KOLONNER: <AUTOSUM> ∑ -IKONEN

Afmærk kolonnen eller rækken + én ekstra celle. Tryk på ∑-ikonen.

REFERENCE OG ABSOLUT REFERENCE

Alle beregninger foretages ved at refere til andre celler: fx indhold i B2 "+ A2*2" (se eksempel nedenfor).
Denne formel kan kopieres nedad i B-kolonnen så alle celler får den samme beregning.
Tilsvarende er celle B6 "+B2+B3+B4+B5 ".
Hvis der beregnes % i en kolonne på basis af værdier i kolonnen før og en sumværdi i cellen under denne, skal der anvendes absolut reference til sum-cellen; dvs cellen markeres med $foran tallet, fx B$6 som i eksemplet nedenfor.
Når en formel med absolut reference kopieres nedad til andre celler i samme kolonne divideres med tallet i sumfeltet i alle celler.

	A	B	C
1		Antal	Antal i %
2	17	34 (formel: "+A2*2")	"+B2/B$6"
3	16	32	"+B3/B$6"
4	17,5	35	"+B4/B$6"
5	15	30	"+B5/B$6"
6	Total	131	

II Import af datafiler

Talformat kan ændres i regnearket og sættes til internationalt før import og sættes tilbage til dansk standard efter import.

- Vælg <*funktioner/ indstillinger* > i menuen.
  Vælg <*international/ skilletegn*> og afkryds 4. nederste mulighed : 1 234.56 (a1,a2).
  Tryk på <*ok*>.

- Vælg <*indsæt/ filer*> i menuen
  Søg fil på "drev A:" og sæt filtype til "alle filer", dvs *.*
  Find filen <*a:\data.kom*> eller <*a:\standard.kom*>.
  Åben filen. Tryk på <*ok*>.
  Opdelingsguide starter: kontroller at der i indstillinger er afkrydset "komma og anførselstegn" og accepter resten.
  Tryk på <opdel>.

Data indlæses i regnearket.

- Vælg <*funktioner/ indstillinger*> i menuen.
  Vælg <*international/ skilletegn*> og afkryds 3. nederste mulighed : 1 234,56 (a1.a2).
  Tryk på <*ok*>.

Herefter er data klar til videre behandling.

III Beregning af gennemsnit og spredning

Beregning af gennemsnit for en blok: @AVG(blok).

Vælg en fri celle under eller til højre for blokken.
Skriv @AVG("første celle"."sidste celle"), og tryk <***retur***>.
Blokken kan også udpeges med musen eller med <***skift F7***> og piletasterne.

Beregning af spredning for en blok: @STDS(blok).

Vælg en fri celle under eller til højre for blokken.
Skriv @STDS("første celle"."sidste celle"), og tryk <***retur***>.
Blokken kan også udpeges med musen eller med <***skift F7***> og piletasterne.

IV Statistik

Oprette fordelingstabeller til stolpediagrammer og tegning af stolpediagram.

1 Oprette fordelingstabeller:

Vælg et passende antal, lige store intervaller og skriv dem i stigende rækkefølge i en tom kolonne under eller til højre for de øvrige data.
Værdier der er lig med et intervals øvre grænse bliver medtaget i dette interval, medens værdier der er lig med nedre intervalgrænse bliver medtaget i forrige interval.

Opret en kolonne med øvre grænseværdier til højre for intervalkolonnen.

Vælg ***<Funktioner/Numeriske værktøjer/Analyse/>*** i menuen.
Der åbnes en oversigt. Vælg ***Histogram*** og tryk ***<Næste>***.

- Skriv datablokken i feltet "Inputceller" - dvs navn eller "første celle". "sidste cel-

le" (blokken kan også udpeges ved at klikke på det lille ikon til højre i skrivefeltet).
- Skriv den netop oprettede grænseværdikolonne i feltet "Intervalceller".
- Marker "Outputceller" og skriv en cellekoordinat til højre for grænseværdikolonnen.
- Afkryds evt feltet "Kumulativ frekvens". Tryk ***<Udfør>***

Regnearket beregner og opretter en tabel med intervalhyppighed.

2 Tilpasning:

- Kopier kolonnen med intervaller ind i den oprettede tabel i stedet for kolonnen med øvre grænseværdier.Summér hyppighederne i cellen under hyppighedskolonnen.
- Flyt kolonnen "Kumulativ %" en celle til højre.
- Brug den tomme kolonne til højre for hyppighed til at beregne hyppigheden i procent.
- Juster antal decimaler i kolonnerne.

3 Stolpediagram:

Vælg graf-ikonet i menuen. Der åbnes en række dialogbokse.
- Marker % hyppighed-kolonnen. Tryk ***<næste>***
- Vælg "søjle" som diagramtype; tryk ***<næste>***. Væl den simple udgave øverst til venstre. Tryk ***<næste>***.
- Accepter indstillinger på næste side; tryk ***<næste>***.
- Titelfane: Skriv en passende overskrift og en forklaring til x-akse: fx højdeintervaller]x', x''] cm, og y-akse: hyppighed i % ; vælg om grafen skal indsættes i regnearket eller være i digramvindue; tryk ***<udfør>***.

Stolpediagram over fordelingstabellen tegnes (klik for at indsætte det i regnearket).

- Klik på x-aksen, vælg *<x-akse egenskaber>*. Skriv intervalkolonnen i vinduet eller udpeg på sædvanlig måde. Lav evt andre justeringer af x-aksen. Tryk ***<ok>***.

Skal der yderligere redigering til gøres det ved at højreklikke forskellige steder i diagrammet (x-akse, y-akse, serier, etc) og ændre indstillingerne så de passer.

F-test. Test for at varianser i to stikprøver er ens

Vælg ***<Funktioner/Numeriske værktøjer/Analyse> i menuen***

Der åbnes en oversigt. Vælg ***F-test.*** Tryk ***<næste>***.

- Udfyld "variabel 1" og "variabel 2" - *dvs navn eller "Arknavn!første celle . sidste celle" (arknavn skal kun bruges hvis talværdierne er fordelt på flere faner).*
- Marker "outputceller" og skriv en cellekoordinat til højre for datablokken i feltet.
- Tryk ***<udfør>***.
- F-test resultatet indsættes som en lille tabel i regnearket; juster kolonnebredde således at statistiktabellen står pænt.
- Aflæs resultat: hvis sandsynligheden $P(F \leq f) \geq 0{,}05$ er der ikke noget der strider mod at acceptere, at varianserne er ens.

t-test. Test for at to gennemsnitsværdier er ens

Vælg ***<Funktioner/Numeriske værktøjer/Analyse>*** i menuen.

Der åbnes en oversigt. Vælg ***t-test.***

- Udfyld "variabel 1" og "variabel 2" - dvs *navn eller "Arknavn!første celle . sidste celle" (arknavn skal kun bruges hvis talværdierne er fordelt på flere faner).*
- Marker "outputområde" og skriv en koordinat for en celle til højre for datablokken i feltet. Tryk ***<næste>***
- Hypotese for forskel i middelværdi: skriv 0 (dvs der testes for at gennemsnitsværdierne er ens). Accepter testtype
- Tryk ***<udfør>***. t-test resultatet indsættes som en lille tabel i regnearket; juster kolonnebredde således at statistiktabellen står pænt.
- Aflæs resultat: hvis sandsynligheden $P(T \leq t) \geq 0{,}05$ (tohalet) er der ikke noget der strider mod at acceptere, at gennemsnittene er ens.

Variansanalyse

Vælg ***<Funktioner/Numeriske værktøjer/Analyse>*** i menuen.
Der åbnes en oversigt. Vælg ***ANOVA: enkeltfaktor.*** Tryk ***<næste>***.

- Skriv datablokken i feltet "Inputceller" - dvs navn eller "første celle". "sidste celle" (blokken kan også udpeges ved at klikke på det lille ikon til højre i skrivefeltet).
- Marker "Outputceller" og skriv en cellekoordinat til højre for eller under datablokken.
- Tryk ***<udfør>***.
- Variansanalysen indsættes som to tabeller i regnearket.
- Aflæs sandsynligheden (P-værdi) eller F-værdierne. Hvis sandsynligheden er større end 0,05 (eller F er mindre end $F_{kritisk}$) er der ikke noget der strider mod at acceptere, at gennemsnittene er ens.

Lineær tilpasning (regression)

Vælg ***<Funktioner/Numeriske værktøjer/Regression>*** i menuen..
Der åbnes en dialogboks.

- Udfyld "uafhængig" (dvs den uafhængigt variable (x): fx højde) med: *navn eller "første celle . sidste celle" eller de andre måder beskrevet ovenfor.*
- Udfyld "afhængig" (dvs den afhængigt variable (y): fx vægt)på tilsvarende måde
- Marker "outputceller" og skriv en koordinat for en celle til højre for datablokken i feltet.
- Tryk ***<udfør>***. Regressionsresultatet indsættes som en (noget indviklet!) tabel i regnearket.
- Aflæs resultat: Det der skal bruges er koefficienter til x-variabel (dvs α) og skæring (dvs b) i forskriften for den rette linie:

$f(x) = \alpha x + b$ i eksemplet: vægt = α højde + b

□ □ □

IV QuattroPro 5

- Definitioner side 35
- Import af datafiler side 36
- Beregning af gennemsnit og spredning side 36
- Grafer side 37
- Statistik side 38

□ □ □

I DEFINITIONER

BLOK

En eller flere celler eller rækker eller kolonner

MARKER BLOK

Peg på øverste celle, hold venstre museknap nede og træk til nederste celle.
Slip museknappen (blokken er markeret med en afvigende farve).
Eller peg på øverste celle, tryk <***skift F7***> og dernæst <***pil ned***> /<***pil højre***> sålænge det er nødvendigt.

ÆNDRE BLOKKENS EGENSKABER

Klik med højre museknap på en celle eller blok, vælg <***egenskaber***>.
Her kan vælges skrifttype, skriftstørrelse, fed, kursiv, farve, format, linietegning, m.m.

KOLONNEBREDDE

Kolonnebredden kan automatisk afpasses efter det bredeste celleindhold med ikonen: **< ↔ >.**

KOPIÉR

Markér celle eller celler der skal kopieres, tryk på <***kopiér***> ikonen. Marker øverste celle, der skal kopieres til og tryk på <***indsæt***> ikonen.

FLYT

Peg på cellen eller blokken med venstre museknap. Hold knappen indtrykket og vent et lille øjeblik, indtil en lille hånd viser sig. Flyt cellen eller blokken - med museknappen indtrykket - til den nye placering.
Slip museknappen.

SUMMERING AF RÆKKER ELLER KOLONNER: <AUTOSUM> ∑ -IKONEN

Afmærk kolonnen eller rækken + én ekstra celle. Tryk på ∑-ikonen.

REFERENCE OG ABSOLUT REFERENCE

Alle beregninger foretages ved at refere til andre celler: fx indhold i B2 "+ A2*2" (se eksempel nedenfor).
Denne formel kan kopieres nedad i B-kolonnen så alle celler får den samme beregning.
Tilsvarende er celle B6 "+B2+B3+B4+B5 ".
Hvis der beregnes % i en kolonne på basis af værdier i kolonnen før og en sumværdi i cellen under denne, skal der anvendes absolut reference til sum-cellen; dvs cellen markeres med $foran tallet, fx B$6 som i eksemplet nedenfor.
Når en formel med absolut reference kopieres nedad til andre celler i samme kolonne divideres med tallet i sumfeltet i alle celler.

	A	B	C
1		Antal	Antal i %
2	17	34 formel: "+A2*2"	"+B2/B$6"
3	16	32	"+B3/B$6"
4	17,5	35	"+B4/B$6"
5	15	30	"+B5/B$6"
6	Total	131	

II Import af datafiler

Import af kommaseparerede filer fra databaser eller datafangstmoduler.
Importen foregår i tre trin.

1) midlertidig ændring af regnearkets decimaltegn og adskillelsestegn, så det passer til databasens,
2) selve importen,
3) tilbageføring af regnearkets decimaltegn og adskillelsestegn til sædvanlig standard.

- Vælg <***Egenskaber> /applikation /international.***
  Der fremkommer en liste med forskellige muligheder. Vælg kombinationen med decimalpunktum og komma-separering: 1 234.56 (a1,a2). Afslut med <***ok***>.

- Vælg <***Værktøj> /importér.***

  a Erstat *.prn i filruden med *.kom
  b Vælg drev og bibliotek (fx a:\).
  Udpeg den ønskede fil.
  c Afkryds "komma og anførselstegn".
  d Afslut med <***ok***>.

- Vælg <***Egenskaber> /applikation /international.***
  Vælg kombinationen med decimalkomma og punktumseparering: 1 234,56 (a1.a2).
  Afslut med <***ok***>.

Efter dataimporten kan regnearket tilrettes med overskrifter, flytning af importerede overskrifter, rettelse af fejlimporterede tegn (ø, ≥, ≤ og æ), m.m.

III Beregning af gennemsnit og spredning

Beregning af gennemsnit for en blok: @AVG(blok).

Vælg en fri celle under eller til højre for blokken. Skriv @AVG("første celle"."sidste celle"), og tryk <***retur***>.
Blokken kan også udpeges med musen eller med <***skift F7***> og piletasterne.

Beregning af spredning for en blok: @STDS(blok).

Vælg en fri celle under eller til højre for blokken. Skriv @STDS("første celle"."sidste celle"), og tryk <***retur***>.
Blokken kan også udpeges med musen eller med <***skift F7***> og piletasterne.

Beregning af gennemsnit ud fra fordelingstabeller (grupperede data):

$$\bar{x} \approx \frac{\sum (n_x\, x)}{\sum n_x}\ ; \quad \left[\begin{array}{rcl} \bar{x} & = & \text{gennemsnit} \\ \sum & = & \text{summationstegn} \\ n_x & = & \text{antal i klasse} \\ x & = & \text{klassemidte} \end{array}\right]$$

Lav et antal kolonner til udregning af formelen.

Beregning af spredning ud fra fordelingstabeller (grupperede data):

$$\sigma \approx \sqrt{\frac{\sum (n_x\, x^2) - \frac{(\sum (n_x\, x))^2}{\sum n_x}}{(\sum n_x) - 1}} \; ; \quad \left[\begin{array}{rcl} \sigma & = & \text{spredning} \\ \sum & = & \text{summationstegn} \\ n_x & = & \text{antal i klasse} \\ x & = & \text{klassemidte} \end{array} \right]$$

Lav et antal kolonner til udregning af formelen

IV GRAFER

1 Stolpediagrammer

Vælg <***Graf> /ny.***

Peg på knappen <***x-akse***>. Pilen skifter til et miniature-regneark.
Klik på knappen og x-akse tallene (dvs intervallerne i fordelingstabellen, fx. figur 3, side 4) kan markeres med musen eller <***skift F7***> og piletaster, jvf ovenfor.
Returner ved klik på pilen i grafbjælken ovenover regnearket.

Samme fremgangsmåde for 1. serie (hyppighed i %), 2. serie, ... og evt. ledetekst (celler med forklaring af, hvad serierne viser).

Tryk <***ok***> og grafen vises.

Højreklik på grafen, vælg titler og skriv overskrift og tekst på akserne.
Afslut med <***ok***>.

2 x-y diagrammer

Vælg <***Graf>/ny***.

Peg på knappen <***x-akse***>. Pilen skifter til et miniatureregneark.
Klik på knappen og marker x-akse tallene (fx højde i et højde-vægt diagram eller systolisk blodtryk i systolisk-diastolisk blodtryksdiagram) med musen eller <***skift F7***> og piletaster, jvf ovenfor.
Returner ved klik på pilen i grafbjælken ovenover regnearket.

Samme fremgangsmåde for 1. y-serie og evt. 2. y- serie og evt. ledetekst (celler med forklaring af, hvad serierne viser).

Hvis punktmængder skal vises separat for kvinder og mænd som i et højde-vægt diagram, skal vægtkolonnen i regnearket deles i to ved at fx mændenes vægt flyttes til en nyoprettet kolonne ved siden af den oprindelige vægtkolonne.

Alle højder (kvinder+mænd) afmærkes som én sammenhængende x-serie.
Den oprindelige vægtkolonne (nu kun med kvinde-vægte) markeres som 1. y-serie, men lige så langt som der er x-værdier, selv om de nederste celler er tomme.
Den nye vægtkolonne markeres som 2. y-serie; også lige så langt som der er x-vær-dier, selv om de øverste celler er tomme.

Tryk <***ok***> og grafen vises, men som et liniediagram.
Højreklik på grafen, vælg graftype og ret til x-y diagram.
Afslut med <***ok***>.

Højreklik på grafen, vælg linieserieegenskaber. Vælg liniestil, sæt den til ingen.
Ret evt mærkestil, størrelse og farve.
Afslut med <***ok***>.

Højreklik på grafen, væg titler og skriv overskrift og tekst på akserne. Afslut med <***ok***>.

V STATISTIK

1 Oprette fordelingstabeller til stolpediagrammer (jvf. IV, 1).

Vælg et passende antal, lige store intervaller og lad regnearket optælle hvor mange værdier, der falder i hvert interval.
Værdier der er lig med et intervals nedre grænse bliver medtaget i dette interval, medens værdier der er lig med øvre intervalgrænse bliver medtaget i næste interval.
(NB! dansk standard er den omvendte procedure: værdier der er lig ned et intervals øvre grænse medtages i dette interval, medens værdier der er lig med nedre intervalgrænse medtages i forrige interval).

A Opret en kolonne med nedre grænseværdier i de ønskede intervaller.

B Vælg **<værktøj> / analyseværktøj**
Der vises en ny knaplineal. Vælg histo-gram ikonen.

C 1 Markér datakolonnen som indlæsningsblok
2 Markér den netop oprettede grænseværdikolonne som intervalblok
3 Markér feltet lige over grænseværdikolonnen som udlæsningsblok
4 Afkryds evt kumulerede værdier
5 Tryk <**ok**>.

D Regnearket beregner og opretter en tabel med intervalhyppigheder (Bin = begyndelsesværdi i interval).

Før man går videre til at oprette et diagram, bør grænseværdierne erstattes af en intervalangivelse (fx nedre grænseværdi: 10 ▸ interval: 10 - 20).

E Tegn stolpediagrammet som beskrevet i **IV , 1**.

2 *F*-test: test for at varianser i to stikprøver er ens

Vælg <***værktøj>/analyseværktøj.***

Der vises en ny knaplineal.
Vælg <***F***>. Udpeg 1. variabelblok (dvs 1. kolonne med data) med musen eller <***skift F7***> og piletasterne.
Dernæst 2. variabelblok. Hvis øverste linie indeholder kolonneoverskrifter afkrydses feltet: Labels.

Vælg et frit område nedenfor eller til højre for datablokkene og markér udlæsningsblok her.
Afslut med <***ok***> og regnearket opstiller en tabel med resultater.

Hvis $P(F \leq f) \geq 0{,}05$ accepteres det, at varianserne er ens.

2 *t*-test: test for at to gennemsnitsværdier er ens

Vælg <***værktøj>/analyseværktøj.***

Der vises en ny knaplineal.
Vælg <***t***>.
Afkryds feltet: "ens varianser".

Udpeg 1. variabelblok (dvs 1. kolonne med data) med musen eller <***skift F7***> og piletasterne.
Dernæst 2. variabelblok. Hvis øverste linie indeholder kolonneoverskrifter afkrydses feltet: Labels.

Vælg et frit område nedenfor eller til højre for datablokkene og markér udlæsningsblok her.
Afslut med <***ok***> og regnearket opstiller en tabel med resultater.

Hvis P(T≤t) ≥ 0,05 accepteres det, at gennemsnittene er ens.

4 Variansanalyse: test for forskelle mellem flere stikprøver.

Vælg <***værktøj>/analyseværktøj.***

Der vises en ny knaplineal. Vælg <***σ/1***> *for én-faktoranalyse eller* <***σ/3***> for to-faktoranalyse.
Udpeg datablokken med musen eller <***skift F7***> og piletasterne. **NB! der skal være lige store antal observationer i de to eller flere kolonner med data, der indgår i analysen.**

Vælg et frit område nedenfor eller til højre for datablokkene og markér udlæsningsblok her. Afslut med <***ok***> og regnearket opstiller en tabel med resultater.

Hvis P(F≤f) ≥ 0,05 accepteres det, at stikprøverne og dermed gennemsnittene er ens.

5 Lineær regression: tilpasning af en punktmængde til bedste rette linie.

Vælg <***værktøj>/avanceret matematik/regression.***

Der vises et nyt vindue. Udpeg den uafhængigt-variable blok (dvs x-aksen) med musen eller <***skift F7***> og piletasterne.
Dernæst den afhængigt-variable blok (dvs y-aksen).

Vælg et frit område nedenfor eller til højre for datablokkene og markér udlæsningsblok her.
Afslut med <***ok***> og regnearket opstiller en tabel med resultater.
Y-akse skæringspunkt er som standard sat til "beregnes"; hvis man vil have grafen til at gå gennem [0,0] skal afkrydsningen ændres.

Korrelationskoefficienten (r) angiver graden af indbyrdes afhængighed mellem de to variable: r=0 ingen afhænghed til r=1 total afhængighed.
Sandsynligheden for at en r-værdi er forskellig fra 0 kan findes med en en ***t***-test:

$$t = \sqrt{\nu\, r^2 / (1 - r^2)}\ ; \quad \nu = n - 2$$

Hvis $t > t_k$ (alternativt 0,025 ≤ P(T≤t) ≤ 0,0975) kan det ikke accepteres at ***r*** er lig med 0 (dvs. ***r*** signifikant forskellig fra 0).

Tabel over udvalgte værdier af t_k og n	
t_k	n
1,960	∞
1,980	120
2,048	30
2,101	20
2,306	10

Litteratur:

1 Bernhard Andersen & Erling B. Andersen:
Grundlæggende statistik
Gyldendal. 1978.

2 Bernhard Andersen & Erling B. Andersen:
Matematisk statistik i grundtræk
Gyldendal. 1973.

3 Arne Nielsen, Jørgen Hilden & Kirsten Fenger:
Statistik og sandsynlighed anvendt i medicin
FADL 2.udg. 1976.

4 Anon.
Applied Statistics with TI Programmable 58/59
Texas Instruments Inc. 1977.

□ □ □

Register

σ, spredning
- beregning 5
- definition 5
- resultatangivelse 5

Absolut reference 24, 30, 35

Blok
- definition 18, 24, 30, 35
- flyt 18, 24, 30, 35
- kopiér 18, 24, 30, 35
- markér 18, 24, 30, 35
- navngiv 18, 24
- reference 24, 30
- ændre egenskaber ved 30, 35

Excel
- 2000 17
- 97 23

F-test
- definition 7
- Excel 2000 . . . . . 21, 32
- Excel 97 27
- regnearksfunktion 38

Fordelingstabel 4
- beregning af gennemsnit 36
- beregning af spredning 37
- regneark 20, 31, 38

Gennemsnit
- beregning 20, 31, 36
- beregning (grupperede data) 36
- resultatangivelse 5
- t-test; sammenligning af to gennemsnit 7, 38
- variansanalyse; sammenligning af flere gns. 10, 39

Grafisk vurdering 6

Hyppigheder
- kumulerede 7

Import af dafiler
- Excel 2000 19
- Excel 97 25
- QuattroPro 5 36
- QuattroPro 8 31

Intervalhyppighed 4

Kommafiler
- data i fil-format 3
- import 19, 25, 31, 36

Lineær tilpasning
- brug af regneark 39
- definition 12
- regneark 21, 33

Måledata 1, 12

Normalfordeling 4
- test af 7

QuattroPro
- 5 34
- 8 29

Reference 24, 35

Regneark 15
- definitioner 18
- F-test 27, 38
- fordelingstabeller 38
- lineær tilpasning (regression) . . . 39
- sammenligning af flere gennemsnit 39
- sammenligning af to gennesmsnit 38
- t-test 38
- variansanalyse 39
- versioner 15

Resultatangivelse 5

Spredning, σ
- beregning 20, 31, 36
- beregning fra grupperede data 31, 36
- definition 5

Spredningsintervaller 5
- grafisk sammenligning af 6

Statistisk analyse 7

Stikprøve 5

Stolpediagrammer
- i regneark 37

Søjlediagram 4
- regneark 20, 32

t-test
- beregning 7
- definition 8
- Excel 2000 . . . . . 21, 33
- Excel 97 27
- regneark 38

t-test på korrelationskoefficient 39

Varians 7

Variansanalyse 10
- definition 10
- Excel 2000 21

QuattroPro 5 39
QuattroPro 8 33
Variation 4
x-y diagrammer
i regneark 37
lineær tilpasning 12

□ □ □