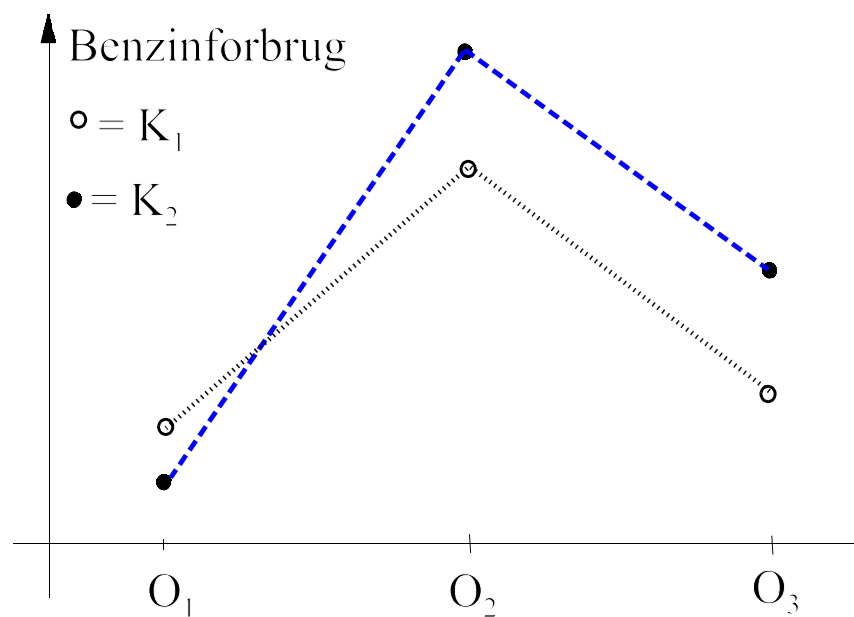


VIDEREGÅENDE STATISTIK

I

Sammenligning af to eller flere kvalitative variable

(TI 89 og Statgraphics)



FORORD

Dette notat kan læses på baggrund af en statistisk viden svarende til lærebogen “M. Oddershede Larsen : Statistiske grundbegreber”.

Notatet er bygget op således, at de væsentligste begreber søges forklaret anskueligt og ved hjælp af et stort antal eksempler.

Det forudsættes, at man har en lommeregner med de statistiske fordelinger indlagt. Der vil derfor i ringe omfang blive benyttet statistiske tabeller.

I “Statistiske grundbegreber” er der i appendix A en “brugsanvisning” på hvorledes dette kan gøres med bl.a lommeregnerne TI-83 og HP48G.

Sidst i dette notat (side 85) findes en tilsvarende “brugsanvisning” for TI-89.

Sidst i hvert kapitel findes en oversigt over de vigtigste formler samt nogle opgaver. En facitliste til opgaverne findes bagerst i notatet.

Fordelen ved direkte at bruge formlerne til løsning af eksemplerne er, at man derved opnår en større forståelse. De mere avancerede programmer bliver let en “sort” kasse, hvor der på mystisk vis dukker et facit op, som man ikke rigtig har noget forhold til.

Imidlertid er det naturligvis også vigtigt, at man kender de muligheder mere avancerede regnemidler tilbyder.

Lommeregneren TI-89 er en god “statistiklommeregner, så i slutningen af hvert kapitel bliver de samme eksempler som i hovedteksten regnet ved benyttelse af dens indbyggede programmer.

Ved behandling af store datamængder og ved mere regneteknisk komplicerede analyser er en lommeregner som TI - 89 ikke nok. Her er det nødvendigt at benytte en PC med et passende statistisk software. I dette notat er anvendt statistikprogrammet “Statgraphics”, således, at alle eksemplerne efter hvert kapitel også er regnet med dette program.

Der findes mange andre udmærkede statistikprogrammer. I et ganske tilsvarende notat (som kan findes på nedenstående adresse) er Statgraphics eksemplerne således udskiftet med SAS-JUMP. Udskifterne fra sådanne statistikprogrammer afviger ikke væsentligt fra hinanden, så skulle man i undervisningen benytte et tredje statistikprogram, kan de studerende uden vanskelighed på basis af disse udskrifter tolke egne udskrifter.

Data foreligger ofte som en fil i et regneark som eksempelvis Excel. Disse regneark har indbygget en del statistik bl.a. de almindeligste testfunktioner.

I notatet “Videregående statistik regnet med Excel” er en række af disse statistiske muligheder gennemgået.

Andre notater i samme “serie” er noterne

Videregående Statistik II: Regressionsanalyse

Videregående Statistik III: Ikke parametriske metoder. Statistisk kvalitetskontrol

Noterne (som både findes i en Statgraphics og en SAS-JUMP version) er søgt udarbejdet, så de kan læses uafhængigt af hinanden.

Alle de nævnte noter kan i pdf-format findes på adressen www.larsen-net.dk

INDHOLD

10 HYPOTSETESTNING (TO STATISTISKE VARIABLE)

10.1	Indledning	1
10.2	Planlægning af forsøg	1
10.3	Test og konfidensintervaller - 2 normalfordelte variable	4
10.3.1	De eksakte spredninger er ukendte, men antages at være ens	5
10.3.2	De eksakte spredninger er ukendte, men antages at være forskellige	7
10.3.3	De eksakte spredninger er kendte	8
10.3.4	Store stikrøvestørrelser (over 30)	9
10.4	Test og konfidensintervaller - 2 binomialfordelte variable	10
10.5	Test og konfidensintervaller - 2 Poissonfordelte variable	11

Oversigter.

10.1	Test af middelværdier μ_1 og μ_2 og konfidensinterval for differens for 2 normalfordelte variable ($\sigma_1 \approx \sigma_2$)	12
10.2.	Test af middelværdier μ_1 og μ_2 og konfidensinterval for differens for 2 normalfordelte variable ($\sigma_1 \neq \sigma_2$)	13
10.3.	Test af middelværdier μ_1 og μ_2 og konfidensinterval for differens for 2 normalfordelte variable (spredninger kendte eller stor stikprøvestørrelse)	14
10.4.	Test af varianser σ_1^2 og σ_2^2 og konfidensinterval for σ_1^2 / σ_2^2 for 2 variable	15
10.5.	Oversigt over test af parametre p_1 og p_2 for binomialfordelte variable	16
10.6.	Oversigt over test af parametre μ_1 og μ_2 for Poissonfordelte variable	17
10.7.	Mapleprogram til dimensionering: 1 og 2 normalfordelte variable	18

Appendix: Statistikprogrammer

10A	Eksempler regnet med TI - 89	19
11B	Eksempler regnet med Statgraphics	22
Opgaver		26

11 VARIANSANALYSE

11.1	Indledning	31
11.2	Ensidet variansanalyse	32
11.2.1	Forklaring af metoder og formler	32
11.2.2	Beregning af ensidet variansanalyse	35
11.3	Tosidet variansanalyse	36
11.3.1	Indledning	36

11.3.2	Planlægning af forsøg	36
11.3.2.1	En faktor ad gangen	37
11.3.2.2	Fuldstændig faktorstruktur	38
11.3.3	Formler og metode	39
11.3.4	Beregning af tosidet variansanalyse	40
11.3.4.1	Model med vekselvirkning	40
11.3.4.2	Additiv model, model uden vekselvirkning	42
11.4	Fuldstændigt randomiseret blokforsøg	44
	Oversigter	
11.1	Formler til beregning af ensidet variansanalyse	47
11.2	Formler til beregning af tosidet variansanalyse	50
	Appendix: Statistikprogrammer	
11A	Eksempler regnet med TI - 89	53
11B	Eksempler regnet med Statgraphics	59
	Opgaver	69
12	FLERE END TO BINOMIAL- OG POISSONFORDELTE VARIABLE	
12.1	Indledning	74
12.2	En faktor på mere end 2 niveauer	74
12.2.1	Binomialfordelt variabel	74
12.2.2	Poissonfordelt variabel	75
12.3	To faktorer i et fuldstændigt faktorforsøg	75
	Oversigter	
12.1.	Oversigt over test af parametre $p_1, p_2, \dots, p_k$ for binomialfordelte variable	78
12.2.	Oversigt over test af parametre $\mu_1, \mu_2, \dots, \mu_k$ for Poissonfordelte variable	78
12.3.	Oversigt over transformation af binomial - eller Poissonfordelte variable til tosidet variansanalyse	78
	Appendix: Statistikprogrammer	
12A	Eksempler regnet med TI - 89	79
12B	Eksempler regnet med Statgraphics	81
	Opgaver	83
	GRUNDLÆGGENDE OPERATIONER PÅ TI - 89	85
	STATISTISKE TABELLER	
TABEL 1.	Fraktiler i U - fordelingen	87
TABEL 2.	Fraktiler i t - fordelingen	88
TABEL 3.	Dimensioneringstabel for 2 variable	89
	FACITLISTE	90
	STIKORD	91

10 HYPOTHESETESTNING

(TO STATISTISKE VARIABLE)

10.1 Indledning

Afsnit 10.2 giver en kort beskrivelse af hvorledes forsøg bedst planlægges ud fra et statistisk synspunkt. Hertil anvendes bl.a. den i tabel 3 angivne dimensioneringsstabel. I tilfælde som ikke dækkes af tabellen kan anvendes det i oversigt 10.7 angivne program skrevet i matematikprogrammet "Maple".

I afsnit 10.3 sammenlignes to normalfordelte variable, mens de variable i afsnit 10.4 og 10.5 forudsættes at være henholdsvis binomial - og Poisson - fordelte.

De grundlæggende begreber vedrørende hypotesetest, konfidensintervaller og dimensionering af forsøg blev i "Statistiske Grundbegreber" grundigt beskrevet for én statistisk variabel.

Da det er de samme begreber og analoge udledninger der anvendes, når det drejer sig om at sammenligne 2 statistiske variable, vil de kun blive ganske kort blive gentaget i dette kapitel.

Som det nævnes i forordet vil vi forudsætte, at man har en lommeregner (som eksempelvis TI - 83, TI - 89 eller HP48G). I appendix 10A og 10B vil de samme beregninger ses udført med mere avancerede statistikprogrammer.

10.2 Planlægning af forsøg

Vi vil i dette afsnit som udgangspunkt tage følgende eksempel.

Eksempel 10.1. Hypotesetest.

En fabrik der producerer maling, har udviklet to nye additiver A_1 og A_2 , som bevirker en kortere tørretid. Additiv A_1 er det dyreste, men man forventer også, at det giver den korteste tørretid. På grund af prisforskellen, skal tørretiden dog være mindst 10 minutter kortere for A_1 , før man vil gå over til den. For at undersøge disse forhold produceres nogle liter maling, som derefter deles op i mindre portioner. Til nogle af portionerne tilsættes additiv A_1 og til andre additiv A_2 . Tørretiden måles derefter. ♦

Generelt gælder, at hvert delforsøg i et forsøg udføres under en række forsøgsbetingelser. De betingelser som med forsæt varieres som led i forsøgets formål, kaldes forsøgsbehandlingerne eller blot behandlingerne (engelsk: treatment). Alle andre delforsøgsbetingelser sammenfattes i et begreb, der kaldes forsøgsenheden. I eksempel 10.1 er additiverne = behandlingerne og forsøgsenhederne er den enkelte portion maling, anvendt apparatur og personale, tidspunkt for delforsøget og de forhold med hensyn til temperatur, luftfugtighed osv., som gælder på forsøgstidspunktet. Bemærk, at forsøgsenhederne ofte indeholder faktorer, som ikke kan gøres ensartet fra delforsøg til delforsøg. Dette bevirker, at resultatet af de enkelte delforsøg varierer. Dette giver forsøgsvariablen variation" eller kort forsøgets "støj".

Randomisering.

For at sikre et statistisk gyldigt forsøg foretager man en såkaldt fuldstændig randomisering. Dette betyder at man ved lodtrækning fordeler "forsøgsmønstrene" tilfældigt på behandlingerne. Dette sker, for at man ikke ubevidst kommer til at favorisere en af de to behandlinger. Hvis man eksempelvis helt systematisk i eksempel 10.1 først laver alle delforsøg med additiv A_1 , kunne dette bevirke en favorisering af A_1 nemlig hvis forsøgsomstændighederne (apparater, personale, luftfugtighed) er mest "gunstige" ved begyndelsen af forsøgsperioden.

For at anskueliggøre denne randomiseringsproces antager vi, at vi i eksempel 10.1 skal lave 4 delforsøg med hver additiv. Endvidere antages, at delforsøgene skal indgå i den almindelige produktionsgang, dvs. at man af tidsmæssige, personalemæssige og på grund af en begrænset mængde apparatur må lade forsøgene forløbe over flere dage. Man tror ikke, at dage, apparatur og laborant har nogen væsentlig betydning for forsøgsresultaterne. Der er sandsynligvis også andre forhold udenfor vor kontrol, og som tilsammen bevirker, at selv om man udfører gentagne delforsøg med samme behandling, så får vi afvigende resultater.

For en sikkerheds skyld vælger vi imidlertid at randomisere dage, apparatur og laboranter

Lad os antage at der gælder følgende:

Mandag er det kun muligt at lave 1 delforsøg, idet apparatur nr 1 og laborant A er de eneste der er ledige.

Tirsdag er der kapacitet ledig til 3 delforsøg:

Ét delforsøg hvor apparatur nr 2 og laborant A benyttes

Ét delforsøg hvor apparatur nr 1 og laborant B benyttes, og

Ét delforsøg hvor apparatur nr 3 og laborant C benyttes.

Onsdag kan der også laves 3 delforsøg osv. (se det følgende skema).

Forsøgsmønstre			Behandlinger (apparater)
Dag	Apparatur	Laborant	
mandag	1	A	
tirsdag	2	A	
tirsdag	1	B	
tirsdag	3	C	
onsdag	3	B	
onsdag	4	C	
onsdag	1	A	
torsdag	3	B	

Vi foretager nu randomiseringen, som kort sagt er en form for lodtrækning. Sædvanligvis vil man benytte et program, der kan generere tilfældige tal (mange lommeregner har et sådant program). For at anskueliggøre randomiseringen vil vi mere primitivt foretage lodtrækningen på følgende måde. På 4 sedler skrives A_1 , på andre 4 sedler skrives A_2 . Hver seddel krølles sammen til en kugle og placeres i en dåse. Sedlerne blandes ved at dåsen rystes (se figur). Hvis den første seddel der udtrækkes er A_2 så betyder det, at det delforsøg der mandag udføres med apparatur 1 og laborant A skal anvende additiv A_2 .

Hvis den næste seddel der udtrækkes er A_1 så betyder det, at det delforsøg der tirsdag udføres med apparatur 2 og råvareleverance 1 skal anvende additiv A_1 osv. Resultaterne kunne eksempelvis være som angivet på følgende skema:

A_1	A_2	
A_2	A_1	A_2
A_2	A_1	A_1

Forsøgsenheder			Behandlinger (apparater)
Dag	Apparatur	Laborant	
mandag	1	A	A_2
tirsdag	2	A	A_1
tirsdag	1	B	A_1
tirsdag	3	C	A_2
onsdag	3	B	A_2
onsdag	4	C	A_1
onsdag	1	A	A_2
torsdag	3	B	A_1

På denne måde sikrer man sig, at vi får et så vidt muligt "statistisk gyldigt" forsøg. Hvis vi derfor efter beregninger (som ses i de følgende kapitler) konkluderer, at der er forskel på additiverne, så er det "korrekt", idet det ville være helt tilfældigt, hvis én af additiverne har været begunstiget med særlig gode forsøgsenheder.

Herved har man også sikret sig, at de to stikprøver (variable) er statistisk uafhængige.

Forsøg bør udføres, så alle behandlinger får lige mange gentagelser.

Ved planlægningen af forsøget er det ganske klart, at hvis man eksempelvis har ressourcer til at lave 20 delforsøg, så ville det være en meget dårlig plan, hvis man lavede 18 delforsøg med A_1 og kun 2 delforsøg med A_2 . Der bør i naturligvis tilstræbes at lave 10 delforsøg med hver behandling. Delforsøg kan mislykkes, så målet i praksis ikke bliver opfyldt. I sådanne tilfælde kan de i de følgende kapitler anførte statistiske analyser dog stadig gennemføres. Testene bliver dog mindre "robuste" (dvs. mere afhængige af at forudsætningerne gælder), og beregningerne mere komplicerede.

Dimensionering

Analogt med forklaringen i "Statistiske Grundbegreber" side 57 kan man under visse forudsætninger beregne hvor mange gentagelser (portioner) der skal anvendes for hver behandling, hvis $P(\text{fejl af type I}) \leq \alpha$ og $P(\text{fejl af type II}) \leq \beta$. Man skal naturligvis angive en bagatelgrænse Δ , men desuden kræver beregningerne, at spredningerne ved de to behandlinger er (tilnærmelsesvis) ens, og at man kan give et nogenlunde realistisk skøn for denne fælles spredning σ .

Det er naturligvis en svaghed ved dimensioneringen, at man inden forsøget er udført skal give et sådant skøn. En vurdering heraf kunne baseres på erfaringer fra tilsvarende forsøg. Findes sådanne erfaringer må man først lave nogle få delforsøg og derfra få et rimeligt gæt på spredningen σ .

At spredningerne er nogenlunde ens vil i praksis ofte være tilfældet, da forsøgsenhederne jo er valgt ved randomisering. Når forsøget så er lavet, kan man (lidt sent) se, om man har skønnet rigtigt. Dimensioneringen skal jo bevirke, at man laver det "rigtige" antal gentagelser. Man kan fristes til at tro, at jo flere gentagelser jo bedre. Dette er imidlertid ikke tilfældet, idet man med "for mange" gentagelser kan risikere at opdage så små forskelle, at de ikke har praktisk betydning, og så er de mange delforsøg jo spild af arbejdskraft og penge.

Formler for dimensionering af 2 variable findes i oversigt 10.1 og 10.3, en tabel til dimensionering findes i tabel 3, ligesom der i oversigt 10.7 er angivet et Mapleprogram, som kan anvendes i de tilfælde, hvor man ikke kan finde resultatet i tabellen.

Eksempel 10.2. (fortsættelse af eksempel 10.1) Dimensionering.

- a) Hvor mange portioner skal anvendes ved forsøget, hvis man ønsker, at $P(\text{fejl af type I}) = \alpha \leq 0.05$, $P(\text{fejl af type II}) = \beta \leq 0.10$ og bagatelgrænsen $\Delta = 10$ minutter, idet man fra mange tilsvarende forsøg ved, at den fælles spredning er $\sigma = 15$ minutter.
- b) Samme spørgsmål og krav som i spørgsmål a), men nu antages, at man ikke kender spredningen, men ud fra nogle få delforsøg skønner, at den er ca. 15 minutter.
- c) Samme spørgsmål og krav som i spørgsmål b), men nu antages $P(\text{fejl af type I}) = \alpha \leq 0.01$.

LØSNING:

$$a) \text{ Af oversigt 10.3 fås } n \geq 2 \cdot \left(\frac{u_{1-\alpha} + u_{1-\beta}}{\frac{\Delta}{\sigma}} \right)^2 = 2 \cdot \left(\frac{u_{0.95} + u_{0.90}}{\frac{10}{15}} \right)^2 = 2 \cdot \left(\frac{1.645 + 1.282}{\frac{10}{15}} \right)^2 = 38,55$$

dvs. der skal udføres i alt $n = 39$ delforsøg af hver behandling

b) Idet $\frac{\Delta}{\sigma} \approx \frac{2}{3} = 0.67$ fås af tabel 3, at det største tal mindre end 0.67 er 0.66 svarende til $n \approx 40$

- c) Tabellen dækker ikke tilfældet $\alpha = 0.01$, så man må bruge Mapleprogrammet i oversigt 10.7.

Da kravet er vanskeligere at opfylde, må $n > 40$. Vi sætter forsøgsvis $n = 65$.

Gå ind i programmet under 2 variable. restart; osv. tryk på "enter" for at udføre programmet.

Ret dernæst i programmet under evalf(6.2,32.0,0.025) tosidet test til

> evalf(pt(65,0.667,0.005)) - evalf(pt(65,0.667,0.995)); #tosidet test og tryk på "enter".

Resultatet blev .1195580414. Da resultatet er lidt større end 0.1 øges n til 67.

> evalf(pt(67,0.667,0.005)) - evalf(pt(67,0.667,0.995)); #tosidet test

Vi får nu .1082259144. Vi øger nu n gradvist, og finder tilsidst, at for $n = 69$ er .09784257051Resultatet er derfor, at vi skal udføre $n = 69$ delforsøg af hver behandling.**10.3 Test og konfidensintervaller for 2 normalfordelte variable.**

Lad os igen betragte det i eksempel 10.1 og 10.2 angivne forsøg. For hver af de 2 additiver udføres en række delforsøg. Forsøgsresultaterne kan generelt skrives:

For additiv A_1 : Statistisk variabel X_1 . Stikprøve : $x_{11}, x_{12}, x_{13}, \dots, x_{1n_1}$.For additiv A_2 : Statistisk variabel X_2 . Stikprøve : $x_{21}, x_{22}, x_{23}, \dots, x_{2n_2}$.

Vi antager, at X_1 og X_2 er statistisk uafhængige normalfordelte variable med henholdsvis middelværdierne μ_1 og μ_2 og spredningerne σ_1 og σ_2 .

Sædvanligvis er man interesseret i at teste nulhypotesen $H_0: \mu_1 = \mu_2 + d$ (d er en given konstant).

Denne omskrives til $H_0: \mu_1 - \mu_2 - d = 0$, og testproceduren baseres på fordelingen af differensen $\bar{X}_1 - \bar{X}_2 - d$.

Ifølge additionssætningen (se eventuelt "Statistiske Grundbegreber side 34) er $\bar{X}_1 - \bar{X}_2 - d$ normalfordelt og fra regnereglerne fås $E(\bar{X}_1 - \bar{X}_2 - d) = E(\bar{X}_1) - E(\bar{X}_2) - d = \mu_1 - \mu_2 - d$ og

$$V(\bar{X}_1 - \bar{X}_2 - d) = V(\bar{X}_1) + V(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Heraf følger, at $U = \frac{\bar{X}_1 - \bar{X}_2 - d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ er normeret normalfordelt.

Desværre er σ_1 og σ_2 sjældent kendt eksakt. Kendes kun deres estimer s_1 og s_2 må der anvendes

andre testprocedurer. Hvis stikprøvestørrelserne er store (over 30) er det dog tilstrækkelig nøjagtigt at anvende en U - test.

Analysen deles derfor op i 4 tilfælde.

- 1) De eksakte spredninger σ_1 og σ_2 er ukendte, men antages at være ens (afsnit 10.3.1)
- 2) De eksakte spredningerne σ_1 og σ_2 er ukendte, men antages forskellige (afsnit 10.3.2).
- 3) De eksakte spredningerne σ_1 og σ_2 er kendte. (afsnit 10.3.3)
- 4) Stikprøvestørrelserne er store (over 30) (afsnit 10.3.4)

10.3.1 De eksakte spredningerne er ukendte, men antages at være ens

Forklaring af formler. I dette tilfælde må spredningerne estimeres ud fra forsøgsresultaterne. Det

ville være naturligt at betragte teststørrelsen $\frac{\bar{X}_1 - \bar{X}_2 - d}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$, men denne har desværre ikke en kendt

fordeling. Da de to variable antages at have samme spredning s_0 , beregnes et estimat for den fælles varians som et vægtet (pooled) gennemsnit s_0^2 af de to estimater for varianserne s_1^2 og s_2^2 . Disse vægtes i forhold til frihedsgraderne, dvs. $s_0^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}$. Frihedsgradstallet for s_0^2 er

$$f_0 = n_1 + n_2 - 2.$$

Teststørrelsen kan nu reduceres til $t = \frac{\bar{X}_1 - \bar{X}_2 - d}{\sqrt{\frac{s_0^2}{n_1} + \frac{s_0^2}{n_2}}} = \frac{\bar{X}_1 - \bar{X}_2 - d}{s_0 \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ som kan vises at være t - fordelt.

Testning af $H_0: \mu_1 = \mu_2 + d$ foretages derfor i følgende rækkefølge:

- 1) Først undersøges om varianserne er ens (selvom det egentlige formål er at undersøge forskelle i middelværdier).

Nulhypotesen $H_0: \sigma_1^2 = \sigma_2^2$ testes mod den alternative hypotese $H: \sigma_1^2 \neq \sigma_2^2$ ved en F - test.

Man beregner $F = \frac{s_1^2}{s_2^2}$ som er F -fordelt $F(n_1 - 1, n_2 - 1)$. Vi får klart en accept af H_0 , hvis F

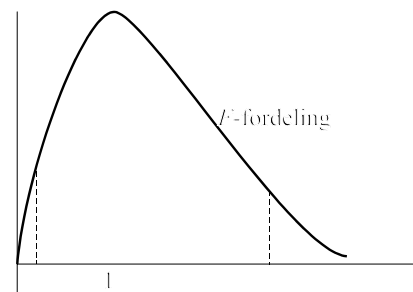
ligger "tæt ved" 1 og en forkastelse, hvis F enten ligger "tæt ved" 0 eller er "meget større" end 1 (se figuren)

I oversigt 10.4 er denne test beskrevet, og kan ses udført i eksempel 10.3.

Får man en accept af nulhypotesen, har man naturligvis ikke hermed vist at varianserne er ens, men da den følgende test af middelværdier er robust overfor mindre forskelle i varianserne, blot vi har samme antal gentagelser (mindre afvigelser kan dog accepteres), er det tilladeligt i den følgende test af middelværdierne, at antage at varianserne er ens.

Vi antager her, at vi får en accept.

Det tilfælde, hvor F - testen giver en forkastelse, bliver behandlet i det næste afsnit.



2) Nulhypotesen $H_0: \mu_1 = \mu_2$ testes på basis af teststørrelsen $t = \frac{\bar{X}_1 - \bar{X}_2 - d}{s_0 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, som er t -fordelt

med f_0 frihedsgrader (jævnfør eksempel 10.3 og oversigt 10.1).

Hvis spredningerne kan antages ens, har man endvidere mulighed for at foretage en dimensionering, dvs. vælge antal gentagelser n , så $P(\text{fejl af type II})$ holdes under β (se eventuelt eksempel 10.2).

Eksempel 10.3. Accept af ens spredning.

To katalysators indflydelse på udbyttet ved en kemisk proces skal undersøges. Den ene katalysator K_1 er billigere end den anden katalysator K_2 . Det påstås imidlertid, at K_2 i middel giver et større udbytte. Kun hvis K_2 i middel giver et udbytte som er mindst 2 enheder større end det udbytte K_1 giver, vil vi vælge katalysator K_1 .

Da forsøgene er både tidskrævende og kostbare, har man kun afsat midler til 8 forsøg med hver katalysator. Resultaterne blev:

K_1	86.41	91.92	86.27	88.99	92.59	89.38	91.40	88.70
K_2	93.65	92.97	97.96	94.01	91.56	91.85	99.99	90.99

- 1) Undersøg på basis af disse resultater, om det på et signifikansniveau på 5% kan påvises at K_2 giver et middeldudbytte der er 2 enheder større end K_1
- 2) Angiv endvidere et 95% konfidensinterval for differensen mellem de to middeldudbytter.

Løsning.

- 1) Lad X_1 = udbyttet ved anvendelse af katalysator K_1 og X_2 = udbyttet ved anvendelse af katalysator K_2 .

X_1 og X_2 antages approksimativt normalfordelte med middelværdi og spredning henholdsvis μ_1, σ_1 og μ_2, σ_2 .

Ved indtastning af forsøgsresultaterne i lommeregneren fås de tilsvarende estimater

$$\bar{x}_1 = 89.458, s_1 = 2.3873 \text{ og } \bar{x}_2 = 94.123, s_2 = 3.213.$$

- a) Først testes om varianserne er ens, dvs. $H_0: \sigma_1^2 = \sigma_2^2$ mod $H: \sigma_1^2 \neq \sigma_2^2$.

Vi anvender formelen i oversigt 10.4 række 3:

$$\text{Teststørrelsen } F = \frac{s_2^2}{s_1^2} = \frac{2.387^2}{3.213^2} = 0.5521, \text{ er } F\text{-fordelt } F(f_2, f_1) = F(7, 7).$$

Da P -værdi $= P(F < 0.5521) = FCdf(0, 0.5521, 7, 7) = 0.2257 > 0.025$ accepteres nulhypotesen, dvs. vi vil i den følgende test antage, at spredningerne er ens.

Et estimat for den fælles spredning er

$$s_0^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2} = \frac{(8 - 1) \cdot 2.387^2 + (8 - 1) \cdot 3.213^2}{14} = 8.01 \text{ med } f_0 = 14$$

- b) Vi tester nu om udbyttet ved katalysator K_2 er 2 større end ved katalysator K_1 , dvs. opstiller nulhypotesen $H_0: \mu_2 \leq \mu_1 + 2$ mod den alternative hypotese $H: \mu_2 > \mu_1 + 2$.

$$\text{Af oversigt 10.1 fås teststørrelsen } t = \frac{\bar{x}_2 - \bar{x}_1 - 2}{s_0 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{94.12 - 89.46 - 2}{\sqrt{8.01} \sqrt{\frac{1}{8} + \frac{1}{8}}} = 1.88$$

Da P -værdi $= P(T > 1.88) = tCdf(1.88, \infty, 14) = 0.0405 < 0.05$ forkastes nulhypotesen.

Konklusion: Der er (svagt) statistisk bevis for at K_2 giver et udbytte der er 2 enheder større end K_1 (vi er dog tæt på en accept).

2) 95% konfidensinterval: Af oversigt 10.1 fås

$$\begin{aligned} \bar{x}_1 - \bar{x}_2 - t_{1-\frac{\alpha}{2}}(f_0) \cdot s_0 \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} &\leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{1-\frac{\alpha}{2}}(f_0) \cdot s_0 \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \\ \Leftrightarrow 89.458 - 94.123 - t_{0.975}(14) \cdot \sqrt{8.01} \cdot \sqrt{\frac{1}{8} + \frac{1}{8}} &\leq \mu_1 - \mu_2 \leq 89.458 - 94.123 + t_{0.975}(14) \cdot \sqrt{8.01} \cdot \sqrt{\frac{1}{8} + \frac{1}{8}} \\ \Leftrightarrow -7.72 \leq \mu_1 - \mu_2 \leq -1.61 \end{aligned}$$

Opgaven løst med TI-89 og Statgraphics findes i appendix 10A og 10B



10.3.2. De eksakte spredninger er ukendte, men antages at være forskellige.

Forklaring af formler. Forkastes nulhypotesen $H_0: \sigma_1^2 = \sigma_2^2$, benyttes "Satterthwaite's approksimation" (jævnfør eksempel 10.4 og oversigt 10.2).

Det skal bemærkes, at er spredningerne tilnærmelsesvis ens, har t -testen i afsnit 10.3.1 større styrke¹ end "Satterthwaite's approksimation", og må derfor foretrækkes.

Hvis der er stor forskel på stikprøvestørrelserne, er det meget vigtigt for anvendelsen af t -testen, at forudsætningen om, at varianserne er ens er holdbar. Til gengæld er det også i sådanne tilfælde, man virkelig ser den store forbedring ved at benytte t -testen fremfor "Satterthwaite's approksimation".

Eksempel 10.4. Forkastelse af ens spredning.

Lad os antage, at vi har samme problemstilling som i eksempel 10.3, men at vi denne gang anvender to andre katalysatorer A og B. Der planlægges med 10 delforsøg men desværre mislykkedes to af forsøgene med katalysator A, så der ikke er lige mange gentagelser.

Stikprøverne giver estimaterne

$$A: \bar{x}_A = 93.92, s_A = 5.06, n_A = 8 \quad B: \bar{x}_B = 89.46, s_B = 2.29, n_B = 10.$$

Undersøg på basis af disse resultater, om det på et signifikansniveau på 5% kan påvises, at A giver et udbytte der er 2 enheder større end B.

Løsning:

1) Lad X_A = udbyttet ved anvendelse af katalysator A og

X_B = udbyttet ved anvendelse af katalysator B.

X_A og X_B antages approksimativt normalfordelte med middelværdi og spredning henholdsvis μ_A, σ_A og μ_B, σ_B .

a) Først testes om varianserne er ens, dvs. $H_0: \sigma_A^2 = \sigma_B^2$ mod $H: \sigma_A^2 \neq \sigma_B^2$.

$$\text{Vi danner teststørrelsen } F = \frac{s_A^2}{s_B^2} = \frac{5.06^2}{2.29^2} = 4.88.$$

Idet P -værdi = $\text{FCdf}(4.88, \infty, 7, 9) = 0.0157 < 0.025$ forkastes nulhypotesen, dvs. vi kan ikke tillade os at antage, at spredningerne er ens.

b) Vi tester nu om udbyttet ved katalysator A er 2 større end ved katalysator B, hvilket betyder, at vi opstiller nulhypotesen $H_0: \mu_A \leq \mu_B + 2$ mod $H: \mu_A > \mu_B + 2$.

Af oversigt 10.2 fås teststørrelsen

¹ En test har "større styrke" end en anden, hvis den med givne data og et givet signifikansniveau giver den største sandsynlighed for at forkaste nulhypotesen.

10. Hypotesetestning (2 statistiske variable)

$$t = \frac{\bar{x}_A - \bar{x}_B - 2}{\sqrt{c}} \quad \text{hvor } c = \frac{s_A^2}{n_A} + \frac{s_B^2}{n_B} = \frac{5.06^2}{8} + \frac{2.29^2}{10} = 3.7249$$

$$t = \frac{\bar{x}_A - \bar{x}_B - 2}{\sqrt{c}} = \frac{93.92 - 89.46 - 2}{\sqrt{3.7249}} = 1.275$$

Frihedsgradstallet f beregnes ved Satterthwaite's metode som det nærmeste hele tal, som er

$$\text{større end } g = \frac{c^2}{\frac{\left(\frac{s_A^2}{n_A}\right)^2}{n_A - 1} + \frac{\left(\frac{s_B^2}{n_B}\right)^2}{n_B - 1}} = \frac{3.7249^2}{\frac{\left(\frac{5.06^2}{8}\right)^2}{8 - 1} + \frac{\left(\frac{2.29^2}{10}\right)^2}{10 - 1}} = 9.29 \approx 10$$

P - værdi = $P(T > 1.275) = \text{tCdf}(1.275, \infty, 10) = 0.1156 > 0.05$ accepteres nulhypotesen.

Konklusion: Vi kan ikke på dette grundlag vise, at udbyttet ved katalysator A er 2 større end ved katalysator B.

Opgaven løst med TI-89 og Statgraphics findes i appendix 10A og 10B



10.3.3. De eksakte spredninger er kendte

Der kan i visse tilfælde ske, at man ad andre veje kender spredningerne eksakt. I disse tilfælde kan

man foretage en U - test, dvs. bruge testfunktionen $U = \frac{\bar{X}_1 - \bar{X}_2 - d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$.

Følgende eksempel illustrerer dette

Eksempel 10.5. eksakt spredning.

På basis af dimensioneringen i eksempel 10.2 udførte man 40 delforsøg af hver behandling. Efter at forsøgsrækken var afsluttet, opdagede man, at et af forsøgene var mislykket og måtte kasseres. Der var følgelig kun 39 delforsøg med additiv A_1 .

Det vides, at spredningerne eksakt er $\sigma_{A1} = 12.13$ og $\sigma_{A2} = 14.2$

1) Man beregner gennemsnittene af de to stikprøver til $\bar{x}_{A1} = 118.6$ og $\bar{x}_{A2} = 129.2$

Kan man ud fra disse data bevise på mindst signifikansniveau $\alpha = 0.05$, at malingen med additivet A_1 tilsat har en mindre middeltørretid end konkurrentens?

2) Hvad vil du anbefale virksomheden at gøre, hvis man som nævnt i eksempel 10.1 kun vil gå over til A_1 hvis tørretiden for A_1 er mindst 10 minutter kortere end for A_2 (bagatelgrænsen).

Løsning:

X_1 = tørringstiden for maling tilsat additiv A_1 .

X_2 = tørringstiden for maling tilsat additiv A_2 .

X_1 og X_2 antages at være uafhængige normalfordelte variable med henholdsvis middelværdierne μ_1 og μ_2 og spredningerne σ_1 og σ_2 .

Nulhypotese $H_0: \mu_1 \geq \mu_2$, Alternativ hypotese: $H: \mu_1 < \mu_2$

Da spredningerne er kendt eksakt anvendes en U - test (se eventuelt oversigt 10.3).

$$u = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{118.6 - 129.2}{\sqrt{\frac{12.13^2}{39} + \frac{14.2^2}{40}}} = -3.57$$

Da $\text{normCdf}(-\infty, -3.57, 0, 1) = 0.00018 < 0.001$ forkastes nulhypotesen (3-stjernet)

Konklusion: Der er et stærkt statistisk bevis for at additiv A_1 i middel har en kortere tørringstid end additiv A_2 .

2) Et 95% konfidensinterval for differensen:

$$\bar{x}_1 - \bar{x}_2 \pm u_{0,975} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 118.6 - 129.2 \pm 1.96 \sqrt{\frac{12.13^2}{39} + \frac{14.2^2}{40}} = -10.6 \pm 5.82$$

Konfidensintervallet er $[-16.43; -4.78]$. (se eventuelt oversigt 10.3)

Konklusion: Da bagatelgrænsen er 10, og næsten 50% af konfidensintervallet ligger under 10, kan det ikke på baggrund af dette materiale anbefales at gå over til det mere kostbare additiv. Bemærk, at dimensioneringen reelt kun har betydning i en acceptsituation, hvor man så med stor sikkerhed kan konkludere, at muligvis giver additiv A_1 en kortere størkningstid, men ikke så meget kortere, at det har praktisk betydning. ♦

10.3.4. Store stikprøvestørrelser (over 30)

Er stikprøvestørrelserne store, viser erfaringen, at det er tilstrækkeligt nøjagtigt at anvende en U-test fremfor de mere præcise metoder der er beskrevet i afsnit 10-3.1 og 10.3.2.

Følgende eksempel illustrerer dette

Eksempel 10.6. Stor stikprøvestørrelse (over 30)

På basis af dimensioneringen i eksempel 10.2 udførte man 40 delforsøg af hver behandling. Efter at forsøgsrækken var afsluttet, opdagede man, at et af forsøgene var mislykket og måtte kasseres. Der var følgelig kun 39 delforsøg med additiv A_1 . Man fik følgende resultater (i minutter)

A_1	117,0	130,7	117,4	142,4	106,8	115,7	116,5	86,7	124,7	127,9	103,0	133,4	103,6	114,4
	132,0	130,2	110,5	119,4	117,2	125,6	126,5	94,0	122,7	106,0	124,0	135,2	124,1	107,9
	113,4	116,2	113,9	131,4	125,6	128,0	110,6	97,2	124,4	134,1	115,7			
A_2	116,7	112,2	133,7	108,9	111,6	157,6	124,2	124,9	102,6	124,1	158,6	136,9	127,4	127,5
	138,9	122,7	120,5	133,2	122,4	143,0	135,6	135,4	136,8	121,4	116,2	135,7	128,2	122,1
	132,7	107,0	138,2	121,6	131,1	157,3	162,5	130,2	126,0	138,6	138,5	106,5		

Kan man ud fra disse data bevise på mindst signifikansniveau $\alpha = 0.05$, at malingen med additivet A_1 tilsat har en mindre middeltørretid end konkurrentens?

Løsning:

X_1 = tørringstiden for maling tilsat additiv A_1 .

X_2 = tørringstiden for maling tilsat additiv A_2 .

X_1 og X_2 antages at være uafhængige normalfordelte variable med henholdsvis middelværdierne μ_1 og μ_2 og spredningerne σ_1 og σ_2 .

For de to stikprøver har man at $\bar{x}_1 = 118.6$ og $\bar{x}_2 = 129.2$ $s_1 = 12.13$ og $s_2 = 14.2$

Vi ser, at vi har de samme værdier som i eksempel 10.5, bortset fra at spredningerne nu ikke er kendt.

Det bliver derfor samme regninger og dermed samme konklusion.

For sammenligningens skyld er i appendix 10A opgaven regnet ved en t-test, og man får ikke overraskende samme resultat.

10.4 Test og konfidensintervaller for 2 binomialfordelte variable.

Som nævnt i bind 1 kan man ofte approksimere en binomialfordeling med en normalfordeling. Det er en sådan approksimation, som formlerne i oversigt 10.5 bygger på. I praksis vil disse forudsætninger for approksimation sædvanligvis være opfyldt.

Vi belyser anvendelsen af oversigten ved følgende eksempel.

Eksempel 10.7. Binomialfordelingstest.

Ved et forsøg der skulle afgøre om C - vitamin har en forebyggende virkning mod forkølelse, fik halvdelen af en gruppe på 280 franske skiløbere C - vitamin mens de øvrige fik kalktabletter (placebobehandling). Fordelingen skete randomiseret, og forsøgspersonerne var uvidende om gruppeinddeling og hvilket medikament de fik. Efter en passende tid optaltes hvor mange af forsøgspersonerne der var forkølede. Resultaterne kan ses af følgende skema:

	Forkølet	Ikke forkølet	Total
C-vitamin	17	122	139
Kalktabletter	31	109	140

Bemærk, at en enkelt forsøgsperson gled ud af forsøget, så grupperne blev ikke helt lige store.

1) Kan det på et signifikansniveau på 5% vises, at C - vitamin har en forebyggende virkning?

2) I bekræftende fald angiv et 95% konfidensinterval for differensen mellem parametrene.

Løsning:

X_1 = antal forkølede personer der har fået C-vitamin. X_1 er binomialfordelt $b(139, p_1)$.

X_2 = antal forkølede personer der har fået Kalktabletter. X_2 er binomialfordelt $b(140, p_2)$.

1) Da vi ønsker at vise, at $p_1 < p_2$ bliver nulhypotesen $H_0: p_1 \geq p_2$ mod $H: p_1 < p_2$.

Oversigt 10.5 anvendes:

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{17}{139}, \quad \hat{p}_2 = \frac{x_2}{n_2} = \frac{31}{140} \quad \text{og} \quad \hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{17 + 31}{139 + 140} = \frac{48}{279}.$$

$$\text{Da } n_1 \cdot \hat{p} = 139 \cdot \frac{48}{279} = 23.9 \in [5; 139 - 5] \text{ og } n_2 \cdot \hat{p} = 140 \cdot \frac{48}{279} = 24.1 \in [5; 140 - 5]$$

er forudsætningerne for at approksimere med normalfordelingen opfyldt. Vi finder af formel (1)

$$u = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p} \cdot (1 - \hat{p}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\frac{17}{139} - \frac{31}{140}}{\sqrt{\frac{48}{279} \cdot \left(1 - \frac{48}{279} \right) \cdot \left(\frac{1}{139} + \frac{1}{140} \right)}} = \frac{-0.099}{0.0452} = -2.193$$

Da $P(U < -2.193) = \text{normCdf}(-\infty, -2.193, 0, 1) = 0.01415 < 0.05$ forkastes nulhypotesen (1-stjernet)

Konklusion: På signifikansniveau 5% er vist, at C-vitamin har en vis forebyggende virkning mod forkølelse.

2) Et 95% konfidensinterval for differensen $p_1 - p_2$ er ifølge oversigt 10.5:

$$\hat{p}_1 - \hat{p}_2 - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$\Leftrightarrow 0.122 - 0.221 - 1.960 \cdot \sqrt{\frac{0.122(1-0.122)}{139} + \frac{0.221(1-0.221)}{140}} \leq p_1 - p_2 \leq 0.122 - 0.221 + 1.960 \cdot \sqrt{\frac{0.122(1-0.122)}{139} + \frac{0.221(1-0.221)}{140}}$$

$$\Leftrightarrow -0.0991 - 0.08773 \leq p_1 - p_2 \leq -0.0991 + 0.08773 \Leftrightarrow \underline{\underline{-0.1869 \leq p_1 - p_2 \leq -0.0114}}$$



10.5 Test og konfidensintervaller for 2 Poissonfordelte variable.

Som nævnt i bind 1 kan man ofte approksimere en Poissonfordeling med en normalfordeling. Det er en sådan approksimation, som formlerne i oversigt 10.6 bygger på. I praksis vil disse forudsætninger for approksimation sædvanligvis være opfyldt. Vi belyser anvendelsen af oversigten ved følgende eksempel.

Eksempel 10.8. Poissonfordelingstest.

En bestemt type TV-apparat produceres på 2 fabrikker A og B. Man har mistanke om, at der er forskel på antallet af loddefejl der findes i apparater fra de to fabrikker. For at teste dette, udtages af den løbende produktion stikprøver på 5 TV-apparater, og man optalte antallet af loddefejl i de 5 apparater. Resultaterne blev:

Fabrik A: På 20 apparater fandtes i alt 12 loddefejl

Fabrik B: På 19 apparater fandtes i alt 7 loddefejl (et apparat måtte udskydes)

Test på dette grundlag, om der er forskel på fejlintensiteten på de to fabrikker.

Løsning.

X_1 = antal loddefejl pr. apparat på fabrik A. X_1 antages Poissonfordelt $p(\mu_1)$.

X_2 = antal loddefejl pr. apparat på fabrik B. X_2 antages Poissonfordelt $p(\mu_2)$.

Da vi ønsker at vise, at $\mu_1 \neq \mu_2$ bliver nulhypotesen $H_0: \mu_1 = \mu_2$ mod $H: \mu_1 \neq \mu_2$.

Oversigt 10.6 række 3 anvendes :

$$\bar{x}_1 = \frac{x_1}{n_1} = \frac{12}{20}, \bar{x}_2 = \frac{x_2}{n_2} = \frac{7}{19}, \text{ og } \bar{x} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{12 + 7}{20 + 19} = \frac{19}{39}.$$

$$\text{Da } n_1 \cdot \bar{x} = 20 \cdot \frac{19}{39} = 9.74 \geq 5 \text{ og } n_2 \cdot \bar{x} = 19 \cdot \frac{19}{39} = 9.26 \geq 5$$

er forudsætningerne for at approksimere med normalfordelingen opfyldt.

$$\text{Vi finder : } u = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\bar{x} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{\frac{12}{20} - \frac{7}{19}}{\sqrt{\frac{19}{39} \cdot \left(\frac{1}{20} + \frac{1}{19} \right)}} = \frac{0.2316}{0.2236} = 1.036$$

Da $P(U > 1.036) = \text{normCdf}(1.036, \infty, 0, 1) = 0.1501 > 0.025$ accepteres nulhypotesen.

Konklusion: Man kan ikke på det grundlag vise, at der er forskel på fejlintensiteten på de to fabrikker.



Oversigt 10.1. Test af middelværdier μ_1 og μ_2 og konfidensinterval for differens $\mu_1 - \mu_2$ for 2 normalfordelte variable. $\sigma_1 \approx \sigma_2$

X_1 og X_2 er **normalfordelte** henholdsvis $n(\mu_1, \sigma_1)$ og $n(\mu_2, \sigma_2)$.

Givet 2 stikprøver af X_1 og X_2 . med størrelse, gennemsnit og spredning henholdsvis n_1 , $\bar{x}_1$, s_1 og n_2 , $\bar{x}_2$, s_2 . Signifikansniveau er α .

Lad d være en given konstant. $t = \frac{\bar{x}_1 - \bar{x}_2 - d}{s_0 \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ hvor $s_0^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}$, $f_0 = n_1 + n_2 - 2$.

T er en statistisk variabel der er t -fordelt med frihedsgradstallet f_0 .

Forudsætning.	H_0 : Nulhypotese H : Alternativ hypotese	Beregning	H_0 forkastes	Dimensionering
σ_1 og σ_2 ukendte $H_0: \sigma_1 = \sigma_2$ accepteres ved F -test	$H_0: \mu_1 \leq \mu_2 + d$ $H: \mu_1 > \mu_2 + d$	P -værdi = $P(T \geq t)$	P -værdi $< \alpha$	$\Delta = \mu_1 - \mu_2 - d $ er den mindste ændring i μ der har praktisk interesse. $\beta = P(\text{type II fejl})$ Antal gentagelser for hver af de to variable findes af dimensioneringstabel 8 For værdier der ikke står i tabellen benyt Mapleprogrammet i oversigt 10.7
	$H_0: \mu_1 \geq \mu_2 + d$ $H: \mu_1 < \mu_2 + d$	P -værdi = $P(T \leq t)$		
	$H_0: \mu_1 = \mu_2 + d$ $H: \mu_1 \neq \mu_2 + d$	P -værdi = $P(T \geq t)$ for $\bar{x}_1 > \bar{x}_2 + d$ P -værdi = $P(T \leq t)$ for $\bar{x}_1 < \bar{x}_2 + d$	P -værdi $< \frac{1}{2}\alpha$	

$100 \cdot (1 - \alpha)\%$ **Konfidensinterval for differens** $\mu_1 - \mu_2$: $\bar{x}_1 - \bar{x}_2 - t_{1-\frac{\alpha}{2}}(f_0) \cdot s_0 \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{1-\frac{\alpha}{2}}(f_0) \cdot s_0 \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

OVERSIGT 10.2. Test af middelværdier μ_1 og μ_2 og konfidensinterval for differens $\mu_1 - \mu_2$ for 2 normalfordelte variable $\sigma_1 \neq \sigma_2$.

X_1 og X_2 er **normalfordelte** henholdsvis $n(\mu_1, \sigma_1)$ og $n(\mu_2, \sigma_2)$. Givet 2 stikprøver af X_1 og X_2 .

Størrelse, gennemsnit og spredning henholdsvis $n_1, \bar{x}_1, s_1$ og $n_2, \bar{x}_2, s_2$. Signifikansniveau er α . Lad d være en given konstant.

$$t = \frac{\bar{x}_1 - \bar{x}_2 - d}{\sqrt{c}}, \text{ hvor } c = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \quad \text{Frihedsgradstallet } f \text{ er det nærmeste hele tal, som er større end } g = \frac{c^2}{\left(\frac{s_1^2}{n_1}\right)^2 + \left(\frac{s_2^2}{n_2}\right)^2}.$$

T er en statistisk variabel der er t -fordelt med frihedsgradstallet f .

Forudsætning	H_0 : Nulhypotese H : Alternativ hypotese	Beregning	H_0 forkastes
<u>Approksimativ metode</u> (Satterthwaites test) σ_1 og σ_2 ukendte $\sigma_1 \neq \sigma_2$ dvs. $H_0: \sigma_1 = \sigma_2$ forkastes ved F -test	$H_0: \mu_1 \leq \mu_2 + d$ $H: \mu_1 > \mu_2 + d$	P -værdi = $P(T \geq t)$	P -værdi $< \alpha$
	$H_0: \mu_1 \geq \mu_2 + d$ $H: \mu_1 < \mu_2 + d$	P -værdi = $P(T \leq t)$	
	$H_0: \mu_1 = \mu_2 + d$ $H: \mu_1 \neq \mu_2 + d$	P -værdi = $P(T \geq t)$ for $\bar{x}_1 > \bar{x}_2 + d$ P -værdi = $P(T \leq t)$ for $\bar{x}_1 < \bar{x}_2 + d$	P -værdi $< \frac{1}{2}\alpha$

$$100 \cdot (1 - \alpha)\% \text{ Konfidensinterval for differens } \mu_1 - \mu_2: \bar{x}_1 - \bar{x}_2 - t_{1-\frac{\alpha}{2}}(f) \cdot \sqrt{c} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{1-\frac{\alpha}{2}}(f) \cdot \sqrt{c}$$

OVERSIGT 10.3. Test af middelværdier μ_1 og μ_2 for 2 normalfordelte variable (spredninger kendte eller stor stikprøvestørrelse)

X_1 og X_2 er **normalfordelte** henholdsvis $n(\mu_1, \sigma_1)$ og $n(\mu_2, \sigma_2)$. Givet 2 stikprøver af X_1 og X_2 . Størrelse, gennemsnit og spredning henholdsvis n_1 , $\bar{x}_1$, s_1

og n_2 , $\bar{x}_2$, s_2 . Signifikansniveau er α . Lad d være en konstant, og $u = \frac{\bar{x}_1 - \bar{x}_2 - d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$. Y en normeret normalfordelt variabel $n(0,1)$.

Forudsætn.	H_0 : Nulhypotese H : Alternativ hypotese	Beregning	H_0 forkastes	Dimensionering
σ_1 og σ_2 kendte eller $n_1 \geq 30$ og $n_2 \geq 30$ σ_1 og σ_2 erstattes så af s_1 og s_2	$H_0: \mu_1 \leq \mu_2 + d$ $H: \mu_1 > \mu_2 + d$	$P\text{-værdi} = P(Y \geq u)$	P - værdi $< \alpha$	$\Delta = \mu_1 - \mu_2 - d $ er den mindste ændring i μ der har praktisk interesse. $\beta = P(\text{type II fejl})$ $n \geq 2 \cdot \left(\frac{u_{1-\alpha} + u_{1-\beta}}{\frac{\Delta}{\sigma}} \right)^2$ $\sigma = \max\{\sigma_1, \sigma_2\}$
	$H_0: \mu_1 \geq \mu_2 + d$ $H: \mu_1 < \mu_2 + d$	$P\text{-værdi} = P(Y \leq u)$		
	$H_0: \mu_1 = \mu_2 + d$ $H: \mu_1 \neq \mu_2 + d$	$P\text{-værdi} = P(Y \geq u)$ for $\bar{x}_1 > \bar{x}_2 + d$ $P\text{-værdi} = P(Y \geq u)$ for $\bar{x}_1 < \bar{x}_2 + d$	P - værdi $< \frac{1}{2} \alpha$	Betegnelser som ovenfor: $n \geq 2 \cdot \left(\frac{u_{1-\frac{\alpha}{2}} + u_{1-\beta}}{\frac{\Delta}{\sigma}} \right)^2$

$$100 \cdot (1 - \alpha)\% \text{ konfidensinterval for differens } \mu_1 - \mu_2: \bar{x}_1 - \bar{x}_2 - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

OVERSIGT 10.4. Test af varianser σ_1^2 og σ_2^2 og konfidensinterval for $\frac{\sigma_1^2}{\sigma_2^2}$ for 2 variable.

X_1 og X_2 er **normalfordelte** henholdsvis $n(\mu_1, \sigma_1)$ og $n(\mu_2, \sigma_2)$.

Givet 2 stikprøver af X_1 og X_2 . Størrelse, gennemsnit og spredning henholdsvis $n_1, \bar{x}_1, s_1$ og $n_2, \bar{x}_2, s_2$.

Signifikansniveau er α . $F = \frac{s_1^2}{s_2^2}$, Q er F - fordelt $F(n_1 - 1, n_2 - 1)$

Forudsætn.	H_0 : Nulhypotese H : Alternativ hypotese	Beregning	H_0 forkastes
μ_1 og μ_2 ukendte	$H_0: \sigma_1^2 \leq \sigma_2^2$ $H: \sigma_1^2 > \sigma_2^2$	P - værdi = $P(Q \geq F)$	P - værdi $< \alpha$
	$H_0: \sigma_1^2 \geq \sigma_2^2$ $H: \sigma_1^2 < \sigma_2^2$	P - værdi = $P(Q \leq F)$	
	$H_0: \sigma_1^2 = \sigma_2^2$ $H: \sigma_1^2 \neq \sigma_2^2$	P - værdi = $P(Q \geq F)$ for $F > 1$ P - værdi = $P(Q \leq F)$ for $F < 1$	P - værdi $< \frac{1}{2} \alpha$

$100 \cdot (1 - \alpha)\%$ Konfidensinterval for forhold $\frac{\sigma_1^2}{\sigma_2^2}$: $\frac{F}{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq F \cdot F_{1-\frac{\alpha}{2}}(n_2-1, n_1-1)$

Specialtilfælde:

μ_1 og μ_2 kendte: $F = \frac{((n_1-1)s_1^2 + n_1 \cdot (\bar{x}_1 - \mu_1)^2) \cdot n_2}{((n_2-1)s_2^2 + n_2 \cdot (\bar{x}_2 - \mu_2)^2) \cdot n_1}$, Q er F - fordelt $F(n_1, n_2)$.

Forudsætn.	Nulhypotese H : Alternativ hypotese	Beregning	H_0 forkastes
μ_1 og μ_2 kendte	$H_0: \sigma_1^2 \leq \sigma_2^2$ $H: \sigma_1^2 > \sigma_2^2$	P - værdi = $P(Q \geq F)$	P - værdi $< \alpha$
	$H_0: \sigma_1^2 \geq \sigma_2^2$ $H: \sigma_1^2 < \sigma_2^2$	P - værdi = $P(Q \leq F)$	
	$H_0: \sigma_1^2 = \sigma_2^2$ $H: \sigma_1^2 \neq \sigma_2^2$	P - værdi = $P(Q \geq F)$ for $F > 1$ P - værdi = $P(Q \leq F)$ for $F < 1$	P - værdi $< \frac{1}{2} \alpha$

μ_1 kendt og μ_2 ukendt: $F = \frac{(n_1-1)s_1^2 + n_1 \cdot (\bar{x}_1 - \mu_1)^2}{s_2^2 \cdot n_1}$, Q er F - fordelt $F(n_1, n_2 - 1)$

Forudsætn.	Nulhypotese	Beregning	H_0 forkastes
μ_1 kendt μ_2 ukendt	$H_0: \sigma_1^2 \leq \sigma_2^2$ $H: \sigma_1^2 > \sigma_2^2$	P - værdi = $P(Q \geq F)$	P - værdi $< \alpha$
	$H_0: \sigma_1^2 \geq \sigma_2^2$ $H: \sigma_1^2 < \sigma_2^2$	P - værdi = $P(Q \leq F)$	
	$H_0: \sigma_1^2 = \sigma_2^2$ $H: \sigma_1^2 \neq \sigma_2^2$	P - værdi = $P(Q \geq F)$ for $F > 1$ P - værdi = $P(Q \leq F)$ for $F < 1$	P - værdi $< \frac{1}{2} \alpha$

OVERSIGT 10.5. Oversigt over test af parametre p_1 og p_2 for binomialfordelte variable.

X_1 og X_2 er **binomialfordelt** henholdsvis $b(n_1, p_1)$ og $b(n_2, p_2)$, hvor n_1 og n_2 er kendte og p_1 og p_2 ukendte. Observerede stikprøveværdier x_1 og x_2 .

Signifikansniveau er α . Lad $\hat{p}_1 = \frac{x_1}{n_1}$, $\hat{p}_2 = \frac{x_2}{n_2}$, $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$ og $u = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p} \cdot (1 - \hat{p}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$. Y en normeret normalfordelt variabel $N(0,1)$.

Forudsætning:	H_0 : Nulhypotese H : Alternativ hypotese	Beregning	H_0 forkastes
Approksimativ metode p_1 og p_2 ukendte. $n_1 \cdot \hat{p} \in [5; n_1 - 5]$ $n_2 \cdot \hat{p} \in [5; n_2 - 5]$	$H_0: p_1 \leq p_2$ $H: p_1 > p_2$	$P\text{-værdi} = P(Y > u)$	$P\text{-værdi} < \alpha$
	$H_0: p_1 \geq p_2$ $H: p_1 < p_2$	$P\text{-værdi} = P(Y < u)$	
	$H_0: p_1 = p_2$ $H: p_1 \neq p_2$	Hvis $\hat{p}_1 \geq \hat{p}_2$ så $P\text{-værdi} = P(Y > u)$ Hvis $\hat{p}_1 < \hat{p}_2$ så $P\text{-værdi} = P(Y < u)$	$P\text{-værdi} < \frac{1}{2}\alpha$

Forudsætning: $x_1 \in [5; n_1 - 5]$ $x_2 \in [5; n_2 - 5]$:

$$100 \cdot (1 - \alpha)\% \text{ konfidensinterval for differens } p_1 - p_2 \quad \hat{p}_1 - \hat{p}_2 - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

OVERSIGT 10.6. Oversigt over test af parameter μ_1 og μ_2 for Poissonfordelt variable .

X_1 og X_2 er Poissonfordelte variable fordelt $p(\mu_1)$ og $p(\mu_2)$ hvor μ_1 og μ_2 er ukendte. Der foreligger to stikprøver af størrelsen n_1 med gennemsnit $\bar{x}_1$ og n_2 med gennemsnit $\bar{x}_2$. Signifikansniveau er α . Lad $\bar{x} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2}{n_1 + n_2}$ og $u = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\bar{x} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$.. Y en normeret normalfordelt variabel $n(0,1)$.

Forudsætning	H_0 : Nulhypotese H : Alternativ hypotese	Beregning	H_0 forkastes
Approksimativ metode μ_1 og μ_2 ukendte. $n_1 \cdot \bar{x} \geq 5$ og $n_2 \cdot \bar{x} \geq 5$	$H_0: \mu_1 \leq \mu_2$ $H: \mu_1 > \mu_2$	$P\text{-værdi} = P(Y > u)$	$P\text{-værdi} < \alpha$
	$H_0: \mu_1 \geq \mu_2$ $H: \mu_1 < \mu_2$	$P\text{-værdi} = P(Y < u)$	
	$H_0: \mu_1 = \mu_2$ $H: \mu_1 \neq \mu_2$	Hvis $\bar{x}_1 \geq \bar{x}_2$ så $P\text{-værdi} = P(Y > u)$ Hvis $\bar{x}_1 < \bar{x}_2$ så $P\text{-værdi} = P(Y < u)$	$P\text{-værdi} < \frac{1}{2} \alpha$

Forudsætning: $n_1 \cdot \bar{x}_1 \geq 10$ og $n_2 \cdot \bar{x}_2 \geq 10$:

$$100 \cdot (1 - \alpha)\% \text{ konfidensinterval for differens } \mu_1 - \mu_2 \quad \bar{x}_1 - \bar{x}_2 - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{x}_1}{n_1} + \frac{\bar{x}_2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{x}_1}{n_1} + \frac{\bar{x}_2}{n_2}}$$

OVERSIGT 10.7. Mapleprogram til dimensionering i 1 og 2 variable.

```

> restart;
> with(stats);
    MAPLE angiver her funktioner i "stats"
> pt:=proc(n,c,a)
    # 1 variabel (ensidet test) med samme stikprøvestørrelse og samme (ukendte) spredning
> local b,d,f,t,cf,gam,x,y;
> f:=n-1;
> gam:=x->int(z^(x-1)*exp(-z),z=0..infinity);
> cf:=evalf(1/(gam(f/2)*2^((f-2)/2)));t:=statevalf[icdf,studentst[f]](1-a);
> y:=cf*int(statevalf [cdf,normald](t/sqrt(f)*u-c*sqrt(n))*u^(f-1)*exp(-(u^2)/2),u=0..infinity);
> y
> end proc;
>
MAPLE GENTAGER HER PROGRAM
> evalf(pt(6,1.87,0.05)); #ensidet test
>
.01288518088
> evalf(pt(6,1.87,0.025))-evalf(pt(6,2.1,0.975)); #tosidet test
0.04047268301
> restart;# 2 variable (ensidet test) med samme stikprøvestørrelse og samme(ukendte) spredning
> with(stats);
> pt:=proc(n,c,a)
> local b,d,f,t,cf,gam,x,y,n1;
> n1:=2*n-1;f:=2*n-2;
> gam:=x->int(z^(x-1)*exp(-z),z=0..infinity);
> cf:=evalf(1/(gam(f/2)*2^((f-2)/2)));t:=statevalf[icdf,studentst[f]](1-a);
> y:=cf*int(statevalf [cdf,normald](t/sqrt(f)*u-c*sqrt(n/2))*u^(f-1)*exp(-(u^2)/2),u=0..infinity);
> y
> end proc;
> MAPLE GENTAGER HER PROGRAM

> evalf(pt(6,2.05,0.05));# ensidet test
0.04927524649
> evalf(pt(6,2.32,0.025))-evalf(pt(6,2.32,0.975)); #tosidet test
0.04947586837

```

Appendix 10A Eksempler regnet på TI - 89.

1. Indledning.

Det forudsættes, at man kender de grundlæggende operationer på lommeregneren.

I Appendix "Vejledning i TI - 89" er beskrevet, hvorledes man beregner sandsynligheden for forskellige fordelinger, beregner gennemsnit og spredning, samt hvorledes man tester og beregner konfidensintervaller for funktion af 1 variabel. Dette forudsættes ligeledes bekendt.

2. Hypotesetest og konfidensintervaller for 2 variable.

Eksempel 10.3. Accept af ens spredning.

Udbyttet ved anvendelse af 2 katalysatorer sammenlignes. Resultaterne blev:

K_1	86.41	91.92	86.27	88.99	92.59	89.38	91.40	88.70
K_2	93.65	92.97	97.96	94.01	91.56	91.85	99.99	90.99

1) Undersøg, om det på et signifikansniveau på 5% kan påvises, at K_2 giver et udbytte der er 2 enheder større end K_1

2) Angiv endvidere et 95% konfidensinterval for differensen mellem de to middeludbytter.

Løsning.

APPS, STAT/LIST, indtast data i list1 og list 2 eller navngiv to kolonner tilhøjre for list 6 med eksempelvis k1 og k2. og indtast data her. Det sidste antages gjort.

1) a) $H_0: \sigma_1^2 = \sigma_2^2$ mod $H: \sigma_1^2 \neq \sigma_2^2$

F6,9:2 - SampFtest, ENTER,

I menuen vælg Data Input Method= Data, ENTER

I menu for "list 1" skrives k1 og for "list2" skrives k2. (benyt evt. VAR-LINK)

"alternative Hyp = $\sigma_1 \neq \sigma_2$ ", ENTER

I udskrift findes P - værdi = 0.4514

Da P - værdi = 0.4514 > 0.05 ¹ accepteres H_0 , dvs. vi vil i den følgende test antage, at spredningerne er ens.

b) $H_0: \mu_2 \leq \mu_1 + 2$ mod $H: \mu_2 > \mu_1 + 2$

F6, 4:2 - SampTtest .ENTER

I menu for "list 1" skrives k1+2, for "alternative Hyp" $\mu_1 < \mu_2$ og pooled til "YES"

I udskrift findes P - værdi = 0.04031

Da P - værdi = 0.04 < 0.05 forkastes nulhypotesen (svagt)

Konklusion: Der er (svagt) statistisk bevis for at K_2 giver et udbytte der er 2 enheder større end K_1 (vi er dog tæt på en accept).

2) Konfidensinterval: F7, 3: 2-SampTInt

Menu er selvforklarende

95% konfidensinterval for differens : C-int=[-7.72; -1.61]

¹ Ved tosidde test beregnes summen af begge "haler" og derfor skal sammenlignes med α og ikke med $\alpha/2$

Eksempel 10.4. Forkastelse af ens spredning.

Der laves forsøg med 2 katalysatorer A og B. Stikprøverne giver estimerne

$$A: \bar{x}_A = 93.92, s_A = 5.06, n_A = 8 \quad B: \bar{x}_B = 89.46, s_B = 2.29, n_B = 10.$$

Undersøg på basis af disse resultater, om det på et signifikansniveau på 5% kan påvises, at A giver et udbytte der er 2 enheder større end B.

Løsning:

APPS, STAT/LIST

1) a) $H_0: \sigma_1^2 = \sigma_2^2$ mod $H: \sigma_1^2 \neq \sigma_2^2$

F6,9:2 - SampFtest, ENTER,

I menuen vælg Data Input Method= Stats, ENTER

Udfyld menu $s_1 = 5.06, n_1 = 8, s_2 = 2.29, n_2 = 10$ "alternative Hyp = $\sigma_1 \neq \sigma_2$ ", ENTER

I udskrift findes P -værdi = 0.0313

Da P -værdi = 0.0313 < 0.05² forkastes H_0 , dvs. vi vil i den følgende test antage, at spredningerne ikke er ens.

b) $H_0: \mu_A \leq \mu_B + 2$ mod $H: \mu_A > \mu_B + 2$.

F6, 4:2 - SampTtest .Data Input Method= Stats., ENTER,

Udfyld menu: $\bar{x}_1 = 93.92, s_1 = 5.06, n = 8, \bar{x}_2 = 89.46 + 2, s_B = 2.29, n_B = 10$, "alternative

Hyp" $\mu_1 > \mu_2$, pooled = no

I udskrift findes P -værdi = 0.1167

Da P -værdi = 0.1167 > 0.05 accepteres nulhypotesen.

Konklusion: Vi kan ikke på dette grundlag vise, at udbyttet ved katalysator A er 2 større end ved katalysator B. 

Eksempel 10.5. eksakt spredning.

Der udføres 39 forsøg med additiv A og 40 forsøg med additiv B.

Det vides, at spredningerne eksakt er $\sigma_{A1} = 12.13$ og $\sigma_{A2} = 14.2$

1) Man beregner gennemsnittene af de to stikprøver til $\bar{x}_{A1} = 118.6$ og $\bar{x}_{A2} = 129.2$

Kan man på mindst signifikansniveau $\alpha = 0.05$ bevise, at malingen med additivet A_1 tilsat har en mindre middeltørretid end konkurrentens?

2) Hvad vil du anbefale virksomheden at gøre, hvis man som nævnt i eksempel 10.1 kun vil gå over til A_1 hvis tørretiden for A_1 er mindst 10 minutter kortere end for A_2 (bagatelgrænsen)

Løsning:

1) Nulhypotese $H_0: \mu_1 \geq \mu_2$, Alternativ hypotese: $H: \mu_1 < \mu_2$

F6, 3:2 - SampZtest .Data Input Method= Stats., ENTER,

Udfyld menu: $\sigma_1 = 12.13, \sigma_2 = 14.2, \bar{x}_1 = 118.6, n_1 = 39, \bar{x}_2 = 129.2, n_B = 40$,

"alternative Hyp" $\mu_1 < \mu_2$

I udskrift findes P -værdi = 0.000178

Da P -værdi = 0.000178 < 0.001 forkastes nulhypotesen (stærkt)

Konklusion: Der er et stærkt statistisk bevis for at additiv A_1 i middel har en kortere tørringstid end additiv A_2 .

² Ved tosidde test beregnes summen af begge "haler" og derfor skal sammenlignes med α og ikke med $\alpha/2$

2) F7, 3:2 - SampZInt . Udfyld menu (er selvforklarende)

C-int=[16.4; -4.77]

Konklusion: Da bagatelgrænsen er 10, og næsten 50% af konsidensintervallet ligger under 10, kan det ikke på baggrund af dette materiale anbefales at gå over til det mere kostbare additiv.

Eksempel 10.6. Stor stikprøvestørrelse (over 30)

Problem som eksempel 10.5, men nu er data opgivet. Vi regner derfor som en sædvanlig t-test, hvor spredningerne ikke er eksakt kendt.

Løsning.

a) $H_0: \sigma_1^2 = \sigma_2^2$ mod $H: \sigma_1^2 \neq \sigma_2^2$

F6,9:2 - SampFtest, Data Input Method= Stats., Menu udfyldes

P - værdi = 0.7297

Da P - værdi = 0.7297 > 0.05 accepteres H_0 , dvs. vi vil i den følgende test antage, at spredningerne er ens.

b) Nulhypotese $H_0: \mu_1 \geq \mu_2$, Alternativ hypotese: $H: \mu_1 < \mu_2$

F6, 4:2 - SampTtest .Data Input Method= Stats., Menu udfyldes herunder pooled=yes

Resultat:

P - værdi = 0.000316

Da P - værdi = 0.0003 < 0.001 forkastes nulhypotesen (stærkt)

Ved sammenligning med resultatet i eksempel 10.5 ses, som forventet, at selv om man anvender en u - test fremfor en t - test, giver det ingen større ændring i P-værdien, da $n \geq 30$.

Konklusion: Der er et stærkt statistisk bevis for at additiv A_1 i middel har en kortere tørringstid end additiv A_2 . ♦

Eksempel 10.7. Binomialfordelingstest.

Halvdelen af en gruppe på 280 franske skiløbere fik C - vitamin mens de øvrige fik kalktabletter .Efter en passende tid optaltes hvor mange af forsøgspersonerne der var forkølede. Resultaterne kan ses af følgende skema:

	Forkølet	Ikke forkølet	Total
C-vitamin	17	122	139
Kalktabletter	31	109	140

1) Kan det på et signifikansniveau på 5% vises, at C - vitamin har en forebyggende virkning?

2) I bekræftende fald angiv er 95% konsidensinterval for differensen mellem parametrene.

Løsning:

1) $H_0: p_1 \geq p_2$ mod $H: p_1 < p_2$.

Forudsætningerne for at approksimere med normalfordelingen er opfyldt (se tidligere)

F6, 6: 2-Prop-ZTest

Udfylder menu: Succes $x_1=17$, $n_1= 139$, succes $x_2=31$, $n_2 = 140$, Alt. hyp : $p_1 < p_2$

Udskrift giver P- værdi=0.0141

Da P - værdi = 0.0141 < 0.05 forkastes nulhypotesen (svagt)

Konklusion: På signifikansniveau 5% er vist, at C-vitamin har en vis forebyggende virkning mod forkølelse.

2) 95% konsidensinterval;

F7, 6: 2-Prop-ZInt , menu udfyldes som under punkt 1.

Udskrift viser C_int=[-0.187 ; -0.0111]

Eksempel 10.8. Poissonfordelingstest. TI - 89 har ikke dette program

Appendix 10B Eksempler regnet på Statgraphics.

1. Indledning

I "Grundlæggende begreber Appendix A" er beskrevet hvorledes man beregner sandsynligheden for forskellige fordelinger og beregner gennemsnit og spredning. Dette forudsættes bekendt.

2. Hypotesetest og konfidensintervaller for 2 variable.

Eksempel 10.3. Accept af ens spredning.

Udbyttet ved anvendelse af 2 katalysatorer sammenlignes. Resultaterne blev:

K ₁	86.41	91.92	86.27	88.99	92.59	89.38	91.40	88.70
K ₂	93.65	92.97	97.96	94.01	91.56	91.85	99.99	90.99

- 1) Undersøg, om det på et signifikansniveau på 5% kan påvises, at K₂ giver et udbytte der er 2 enheder større end K₁
- 2) Angiv endvidere et 95% konfidensinterval for differensen mellem de to middeludbytter.

Løsning:

- 1) a) $H_0: \sigma_1^2 = \sigma_2^2$ $H: \sigma_1^2 \neq \sigma_2^2$

Vælg (Compare\ Two Samples\ Two Samples Comparison).

I den fremkomne tabel: Vælg (K1\tryk på ► ved "Sample 1"\K2\tryk på ► ved "Sample 2"\ OK)

Vælg(gul ikon=Tabular options\Comparison of Standard Deviations\OK)

Resultat:

Comparison of Standard Deviations

	K1	K2
Standard deviation	2,38733	3,37667
Variance	5,69936	11,4019
Df	7	7
Ratio of Variances	= 0,499861	

95,0% Confidence Intervals

Standard deviation of K1: [1,57844;4,85887]

Standard deviation of K2: [2,23257;6,87244]

Ratio of Variances: [0,100074;2,49676]

F-test to Compare Standard Deviations

Null hypothesis: sigma1 = sigma2

Alt. hypothesis: sigma1 NE sigma2

F = 0,499861 P-value = 0,380525

Da P -value = 0.4514 > 0.05¹ accepteres nulhypotesen, dvs. man kan i den følgende analyse antage, at spredningerne er ens.

- b) $H_0: \mu_2 \leq \mu_1 + 2$ mod $H: \mu_2 > \mu_1 + 2$.

eller hvis H skrives $H: \mu_2 > \mu_1 + 2 \Leftrightarrow \mu_1 - \mu_2 < -2$.

Vælg (Compare\ Two Samples\ Two Samples Comparison).

I den fremkomne tabel: Vælg (K1\tryk på ► ved "Sample 1"\K2\tryk på ► ved "Sample 2"\ OK)

Vælg (gul ikon = Tabular Options\Comparison of Means\ OK)

Vi får en udskrift, som ikke er relevant.

Vælg (Cursoren på udskrift, højre musetast\ Pane options\Null Hypothesis =- 2\ alt. hypothesis" til "Lesser than"\ OK).

Markering ved "Asume Equal Sigmas"

Comparison of Means

```
-----
95,0% confidence bound for mean of K1: 89,4575 + 1,59912    [91,0566]
95,0% confidence bound for mean of K2: 94,1225 + 2,15215    [96,2747]
95,0% confidence bound for the difference between the means
    assuming equal variances: -4,665 + 2,49262    [-2,17238]
t test to compare means
Null hypothesis: mean1 - mean2 = -2,0
Alt. hypothesis: less than
    assuming equal variances: t = -1,88312    P-value = 0,0403142
```

Da P - værdi = 0.04 < 0.05 forkastes nulhypotesen (svagt)

Konklusion: Der er (svagt) statistisk bevis for at K_2 giver et udbytte der er 2 enheder større end K_1 , (vi er dog tæt på en accept).

2) 95% konfidensinterval:

Her beregnes (se ovenstående udskrift) kun et ensidet interval, hvoraf vi kan slutte, at $\mu_1 - \mu_2 < -2.17238 \Leftrightarrow \mu_2 > \mu_1 + 2.17238$

Ved ensidede test er et 95% konfidensinterval for differensen også "ensidet" $[g; \infty[$ hvor den nedre grænse g er bestemt ved 95% fraktilen:

$$g = \bar{x}_2 - \bar{x}_1 - t_{0,95}(14) \cdot s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 94.12 - 89.46 - 1.76 \cdot \sqrt{8.01} \cdot \sqrt{\frac{1}{8} + \frac{1}{8}} = 4.67 - 2.49 = 2.18$$

At dette er mere korrekt, ses også af, at dette konfidensinterval ikke indeholder 2



Eksempel 10.4. Forkastelse af ens spredning.

Der laves forsøg med 2 katalysatorer A og B. Stikprøverne giver estimerne

A: $\bar{x}_A = 93.92$, $s_A = 5.06$ $n_A = 8$ og B: $\bar{x}_B = 89.46$, $s_B = 2.29$, $n_B = 10$.

Undersøg på basis af disse resultater, om det på et signifikansniveau på 5% kan påvises, at A giver et udbytte der er 2 enheder større end B.

Løsning:

a) Vælg (Compare\ Two Samples\ Hypothesis tests).

I tabel Vælg (Normal Sigmas\Sample 1 Sigma = 5,06\Sample 2 Sigma = 2,29 \ Sample 1 Size = 8, Sample 2 Size = 10\ OK) Udskrift:

Hypothesis Tests

```
-----
Sample standard deviations = 5,06 and 2,29
Sample sizes = 8 and 10
95,0% confidence interval for ratio of variances: [1,16329;23,5487]
Null Hypothesis: ratio of variances = 1,0
Alternative: not equal
Computed F statistic = 4,88236,    P-Value = 0,031381
Reject the null hypothesis for alpha = 0,05
```

Nulhypotesen forkastes

b) Test $H: \mu_A > \mu_B + 2$.

Vælg (Compare\ Two Samples\ Hypothesis tests).

Vælg (rød ikon = Input Dialog\ I tabel vælg "Normal Means"\Null Hypothesis for Difference = 2, Sample 1 Mean = 93,92\ osv.\ OK)

Vi får en udskrift, som ikke er relevant.

Vælg (Cursoren på udskrift, højre musetast\Analysis Options\ alt. hypothesis" til "Greater than"\Ved "Asume Equal Sigmas" afhænger markeringen af resultatet af en test, her antages de er forskellige \ OK).

Vi får:

```
Hypothesis Tests
-----
Sample means = 93,92 and 89,46
Sample standard deviations = 5,06 and 2,29
Sample sizes = 8 and 10
Approximate 95,0% lower confidence bound for difference between means:
4,46 - 3,52545 [0,93455]
Null Hypothesis: difference between means = 2,0
Alternative: greater than
Computed t statistic = 1,27462
P-Value = 0,116704
Do not reject the null hypothesis for alpha = 0,05.
(Equal variances not assumed).
```

Konklusion: Vi kan ikke på dette grundlag vise, at udbyttet ved katalysator A er 2 større end ved katalysator B.



Eksempel 10.5. eksakt spredning.

Hvis begge spredninger er eksakt kendt kan Statgraphics ikke lave den ønskede u - test eller 95% konfidensinterval,

Eksempel 10.7. Binomialfordelingstest.

Halvdelen af en gruppe på 280 franske skiløbere fik C - vitamin mens de øvrige fik kalktabletter .Efter en passende tid optaltes hvor mange af forsøgspersonerne der var forkølede. Resultaterne kan ses af følgende skema:

	Forkølet	Ikke forkølet	Total
C-vitamin	17	122	139
Kalktabletter	31	109	140

- 1) Kan det på et signifikansniveau på 5% vises, at C - vitamin har en forebyggende virkning?
- 2) I bekræftende fald angiv er 95% konfidensinterval for differensen mellem parametrene.

Løsning:

$$1) H_0: p_1 \geq p_2 \text{ mod } H: p_1 < p_2 .$$

Forudsætningerne for at approksimere med normalfordelingen er opfyldt (se tidligere)

$$\text{Idet estimerer er } \hat{p}_1 = \frac{17}{139} = 0.1223 \text{ og } \hat{p}_2 = \frac{31}{140} = 0.2214 \text{ fås:}$$

Vælg (Compare\Two Samples \Hypotesis tests\Binomial Proportions\Sample 1 Proportions = 0.1223\ Sample 2 Proportions = 0.2214\ Sample 1 Size = 139\ Sample 2 Size = 140\OK)

Vi får en udskrift, som ikke er relevant.

Vælg (Cursoren på udskrift, højre musetast\ Pane options\ alt. hypothesis" til "less than"| OK)

Der fremkommer nu udskriften:

```
Hypothesis Tests
-----
Sample proportions = 0,122302 and 0,221429
Sample sizes = 139 and 140
Approximate 95,0% upper confidence bound for difference between proportions:
[-0,0254986]
Null Hypothesis: difference between proportions = 0,0
Alternative: less than
Computed z statistic = -2,19349
P-Value = 0,0141358
Reject the null hypothesis for alpha = 0,05.
```

Da P -værdi = 0.0141 < 0.05 forkastes nulhypotesen (svagt)

Konklusion: På signifikansniveau 5% er vist, at C-vitamin har en vis forebyggende virkning mod forkølelse.

2) Ved en ensidet test er konfidensintervallet ensidet. $p_1 - p_2 < -0.0245 \Leftrightarrow p_1 < p_2 - 0.0245$

Eksempel 10.8. Poissonfordelingstest.

En bestemt type TV-apparat produceres på 2 fabrikker A og B. Man har mistanke om, at der er forskel på antallet af loddefejl der findes i apparater fra de to fabrikker. For at teste dette, udtages af den løbende produktion stikprøver på 5 TV-apparater, og man optalte antallet af loddefejl i de 5 apparater. Resultaterne blev:

Fabrik A: På 20 apparater fandtes i alt 12 loddefejl

Fabrik B: På 19 apparater fandtes i alt 7 loddefejl (et apparat måtte udskydes)

Test på dette grundlag, om der er forskel på fejlintensiteten på de to fabrikker.

Løsning.

X_1 antages Poissonfordelt $p(\mu_1)$. Estimat $\tilde{\mu}_1 = \frac{12}{20} = 0.6$

X_2 antages Poissonfordelt $p(\mu_2)$. Estimat $\tilde{\mu}_2 = \frac{7}{19} = 0.368$

Test nulhypotesen $H_0: \mu_1 = \mu_2$ mod den alternative hypotese $H: \mu_1 \neq \mu_2$:

Løsning:

Vælg (Compare\Two Samples \Hypothesis tests\Poisson Rates\Sample 1 Rate = 0,6\ Sample 2 Rate = 0,368\ Sample 1 Size = 20\ Sample 2 Size = 19\OK)

Der fremkommer nu udskriften:

Hypothesis Tests

Sample rates = 0,6 and 0,368

Sample sizes = 20 and 19

Approximate 95,0% confidence interval for difference between rates:
[-0,203485;0,667485]

Null Hypothesis: difference between rates = 0,0

Alternative: not equal

Computed z statistic = 1,03775

P-Value = 0,299383

Do not reject the null hypothesis for alpha = 0,05.

Warning: normal approximation may not be appropriate for small sample sizes.

Da P-Value = 0,299383 > 0.05 accepteres nulhypotesen

Konklusion: Man kan ikke på det grundlag vise, at der er forskel på fejlintensiteten på de to fabrikker. 

OPGAVER

Opgave 10.1

Det påstås at modstanden i en tråd af type A er større end modstanden i en tråd af type B. Til afklaring af denne påstand udtages ved et fuldstændigt randomiseret forsøg tilfældigt n tråde af hver type og deres modstande måles.

Find det mindste antal n

- 1) hvis man ønsker at $P(\text{fejl af type I}) = \alpha \leq 0.05$, $P(\text{fejl af type II}) = \beta \leq 0.05$ og bagatelgrænsen er $\Delta = 0.1$ ohm, og man ved, at spredningen $\sigma = 0.1$ ohm.
- 2) hvis man ønsker at $P(\text{fejl af type I}) = \alpha \leq 0.05$, $P(\text{fejl af type II}) = \beta \leq 0.05$ og bagatelgrænsen er $\Delta = 0.1$ ohm, og man har en forhåndsformodning om, at spredningen er ca. $\sigma = 0.1$ ohm.
- 3) Hvilke konklusioner vedrørende behandlingernes virkning kan gøres, såfremt man ved testning af forsøgsresultaterne finder
 - a) signifikans
 - b) ingen signifikans
- 4) Hvilke yderligere analyser af forsøgsresultaterne bør foretages, såfremt testningen
 - a) viser signifikans
 - b) ikke viser signifikans.

Opgave 10.2

I et forsøg ønsker man at sammenligne udbyttet ved benyttelse af 2 reaktortyper.

Man ønsker at kunne påvise eventuelle forskelle i middeludbytte ned til ca. $\Delta = 6.0$.

Find den mindste værdi af $n = \text{“antal delforsøg med hver reaktortype”}$, for hvilken

$$P(\text{fejl af type I}) = \alpha \leq 0.05, \quad P(\text{fejl af type II}) = \beta \leq 0.10.$$

Man kender ikke spredningen eksakt, men mener, den højst er ca 7 enheder.

Opgave 10.3^T

Et levnedsmiddelfirma havde udviklet en diæt, som har lavt indhold af fedt, kulhydrater og kolesterol. Diæten er udviklet med henblik på patienter med hjerteproblemer, men firmaet ønsker nu at undersøge diættens virkning på folk med vægtproblemer.

To stikprøver på hver 100 personer med vægtproblemer blev udtaget tilfældigt. Gruppe A fik den nye diæt, mens gruppe B fik den diæt, man normalt gav. For hver person blev registreret størrelsen af vægttabet i en 3 ugers periode.

Man fandt følgende værdier for gennemsnit og spredning:

$$\text{Gruppe A: } \bar{x}_A = 9.31 \text{ kg}, \quad s_A = 4.67$$

$$\text{Gruppe B: } \bar{x}_B = 7.40 \text{ kg}, \quad s_B = 4.04.$$

- 1) Undersøg om vægttabet for gruppe A er signifikant større end for gruppe B. Signifikansniveau $\alpha = 5\%$.
- 2) Beregn et 95% konfidensinterval for differensen mellem de to gruppers middelværdier.

Opgave 10.4

På et laboratorium undersøgte filtreringstiden for en opløsning af et bestemt gødningsstof ved benyttelsen af to forskellige filtertyper (F_1) og (F_2). Følgende stikprøveværdier observeredes:

(F_1)	8	10	12	13	13	9	14	4
(F_2)	9	10	10	7	9			

Det antages, at filtreringstiderne X_1 og X_2 er normalfordelte $n(\mu_1, \sigma_1)$ og $n(\mu_2, \sigma_2)$

1) Test, om det kan antages, at $\sigma_1 = \sigma_2$.

2) Test under hensyntagen til det i 1) fundne, om det kan antages, at $\mu_1 = \mu_2$.

3) Opstil et 95% -konfidensinterval for:

a) $\mu_1 - \mu_2$, b) $\frac{\sigma_1}{\sigma_2}$.

Opgave 10.5

Det påstås at modstanden i en tråd af type A er større end modstanden i en tråd af type B. Til afklaring af denne påstand udtages tilfældigt 6 tråde af hver type og deres modstande måles. Følgende resultater fandtes:

Modstand i tråd A (i ohm)	0.140	0.138	0.143	0.142	0.144	0.137
Modstand i tråd B (i ohm)	0.135	0.140	0.142	0.136	0.138	0.140

Hvilke konklusioner kan drages med hensyn til påstanden?

Opgave 10.6

I et laboratorium foretoges 15 uafhængige bestemmelser af furfurols kogepunkt, idet 8 af bestemmelserne foretoges af én kemiingeniør, de resterende bestemmelser af en anden kemiingeniør. Resultaterne var ($^{\circ}\text{C}$):

1. ingeniør	162.2	161.3	161.9	161.2	163.4	162.4	162.5	162.0
2. ingeniør	163.3	162.6	161.8	163.8	163.0	163.2	164.1	

Undersøg, om de to ingeniørers resultater i middel er ens.

Opgave 10.7

Med henblik på at sammenligne de farmakologiske virkninger af stofferne morphin og nalbuphin foretoges et fuldstændigt randomiseret forsøg, hvorved man på 10 forsøgspersoner målte ændringen i pupildiameter (millimeter) efter indsprøjtning af en standarddosis af en opløsning af morphin (M) eller nalbuphin (N).

Forsøgsplan og forsøgsresultater var:

M: 1.0	N: 0.0	M: 1.9	M: 2.0	N: 0.8	M: 0.8	M: 0.1	N: - 0.3	N: 0.4	N: 0.2
--------	--------	--------	--------	--------	--------	--------	----------	--------	--------

Analyser forsøgsresultaterne og opstil et 95%-konfidensinterval for differensen $\mu(M) - \mu(N)$ mellem de to middelværdier.

Opgave 10.8

En produktion af plastikvarer må omlægges på grund af bestemmelser i en ny miljølov.

Ved den fremtidige produktion kan inden for miljølovens rammer vælges mellem 2 produktionsmetoder I og II. Metode I er den dyreste, og fabrikanten har regnet ud, at det (kun) kan betale sig at benytte metode I, såfremt den giver et middeludbytte, som er mindst 10 måleenheder (udbytteprocenter) større end udbyttet ved benyttelse af metode II.

1) Find det mindste antal n , hvis man ønsker

at $P(\text{fejl af type I}) = \alpha \leq 0.05$, $P(\text{fejl af type II}) = \beta \leq 0.05$ og bagatelgrænsen er $\Delta = 10$ enheder.

Man kender ikke σ , men har en forhåndsformodning om, at spredningen er højst 6 enheder.

2) Ved et fuldstændigt randomiseret forsøg fandtes følgende måleresultater:

Metode I	35.2	38.1	37.6	37.6	34.9	37.9	36.5	40.0	36.2	37.4	37.2	37.9
Metode II	26.2	22.2	24.3	24.5	22.0	27.6	23.8	22.8	23.4	20.8		

Fabrikanten valgte herefter at benytte metode I.

Foretag en undersøgelse af, om valget var statistisk velmotiveret.

Opstil et 95% - konfidensinterval for differensen mellem middeludbytteerne ved benyttelse af metoderne I og II.

Opgave 10.9

To sjællandske fabrikker producerer begge en bestemt type kvægfoder, for hvilken det ønskes, at proteinindholdet i færdigvaren skal være 26%. På de 2 fabrikkers driftslaboratorier foretoges følgende målinger af proteinindholdet i en uges produktion:

Fabrik 1	27.3	26.1	26.9	24.8	26.2	25.7	26.5
----------	------	------	------	------	------	------	------

Fabrik 2	26.0	26.7	25.6	26.1	26.2	25.5	26.0	26.1	26.2	25.9
----------	------	------	------	------	------	------	------	------	------	------

Foretag en statistisk vurdering af, om de to produktioner kan antages i middel at give kvægfoder med samme proteinindhold.

Opgave 10.10

Måling af intelligenskvotient på 16 tilfældigt udvalgte studerende ved en diplom-retning (med mere end 200 studerende) viste et gennemsnit på $\bar{x}_1 = 107$ og en empirisk varians på $s_1^2 = 100$, medens en tilsvarende måling på 14 tilfældigt udvalgte studerende fra en anden diplomretning viste et gennemsnit på $\bar{x}_2 = 112$ og en empirisk varians på $s_2^2 = 64$.

Tyder disse tal på en forskel på studentermaterialet på de to retninger?

Opgave 10.11

- 1) 100 studerende, 52 piger og 48 drenge, indstillede sig til en prøve, ved hvilken 39 piger og 27 drenge bestod.
Undersøg, om det anførte tyder på, at resultatet ved den pågældende prøve afhænger af deltagerens køn.
- 2) Det oplyses supplerende, at pigerne ved ovennævnte prøve opnåede et gennemsnit på 64% med en empirisk spredning på 10%, medens drengenes gennemsnit var 59% med en empirisk spredning på 8%.
Undersøg, om det anførte kan tages som vidnesbyrd om, at piger i almindelighed klarer sig bedre end drenge ved den omhandlede prøve.

Opgave 10.12

To sjællandske fabrikker producerer begge en bestemt type kvægfoder, for hvilken det ønskes, at proteinindholdet i færdigvaren skal være 26%. For den omhandlede produktion er der fastsat en øvre og en nedre tolerancegrænse for proteinindholdet. Partier med et proteinindhold uden for toleranceintervallet klassificeres som "dumpere".

I en 3-måneders periode havde fabrik 1 af en produktion på 60 foderstofpartier 5 dumpere, medens fabrik 2 af en produktion på 100 foderstofpartier havde 12 dumpere. Kan det heraf statistisk konkluderes, at dumpeprocenten i middel har været størst for fabrik 2?

Opgave 10.13

To virksomheder A og B fremstiller dåser med nominelt 100 g rejeost. 10 tilfældigt udtagne dåser fra A's produktion og 20 tilfældigt udtagne dåser fra B's produktion viste følgende resultater:

	Virksomhed	
	A	B
Totalt antal rejer	81	216
Gennemsnittet $\bar{x}$ af nettoindhold	101.2 g	98.3 g
Empirisk spredning s af nettoindhold	1.0 g	2.7 g

- 1) Test, om det gennemsnitlige antal rejer pr. dåse kan antages at være det samme for virksomhedernes produktion.
- 2) Test, om det gennemsnitlige nettoindhold i en dåse kan antages at være det samme for virksomhedernes produktion.

Opgave 10.14

- 1) Mange forbrugere tror, at såkaldte "mandagsbiler", dvs. biler produceret om mandagen, har flere alvorlige fejl end biler produceret på ugens øvrige arbejdsdage.
For at undersøge, om der er noget grundlag for denne tro, udtog man på en bilfabrik tilfældigt 100 "mandagsbiler" og undersøgte dem for fejl. Man fandt at 8 biler havde alvorlige fejl. Tilsvarende udtog man tilfældigt 200 biler, der var produceret på ugens øvrige arbejdsdage, og man fandt 12 biler, der havde alvorlige fejl.
Giver denne undersøgelse støtte til formodningen om, at "mandagsbiler" er af dårligere kvalitet end andre biler.
- 2) De 100 ovennævnte "mandagsbiler" havde i alt 1030 konstaterede større eller mindre enkeltfejl, medens de 200 ovennævnte andre biler i alt havde 1899 konstaterede fejl. Tyder dette på, at der er forskel i fejlintensiteten på bilerne i de to grupper?

Opgave 10.15

Ved en undersøgelse af en eventuel sammenhæng mellem luftforurening og forekomsten af lungecancer sammenlignedes bl.a. sygdommens forekomst i byen X - købing inden for den gamle bygrænse (i nærheden af byens industrivirksomheder) med dens forekomst i samme bys forstadsområde (villakvarter):

	Antal tilfælde af lungecancer	Samlet indbyggerantal
Indre by	30	9000
Forstadsområde	40	27000

- 1) Det ses, at den relative hyppighed af cancertilfælde i den indre by afviger fra den relative hyppighed i forstadsområdet. Kan dette forklares som et tilfældigt udsving? Den opstillede nulhypotese, som testes, ønskes specificeret med angivelse af den alternative hypotese.
- 2) Diskuter muligheden for at drage årsagsmæssige konklusioner ud fra det fundne testresultat.

11 VARIANSANALYSE

11.1 Indledning

I kapitel 10 sammenlignede vi 2 middelværdier. I dette kapitel sammenlignes flere end to middelværdier. Det karakteristiske er, at de forekommende faktorer er kvalitative, dvs. har niveauer, som ikke er karakteriseret ved en målelig egenskab. Dette illustreres i det følgende eksempel.

Eksempel 11.1 (én faktor). *Virkningerne af 4 tilsætningsstoffer T_1, T_2, T_3, T_4 på mængden af urenheder ved en kemisk proces ønskes sammenlignet. For hvert tilsætningsstof måles mængden af “uønsket stof” 3 gange.*

Forsøgsresultaterne blev følgende:

Tilsætningsstof	T_1	T_2	T_3	T_4
Mængde urenhed	108	105	108	117
	110	110	111	119
	112	109	113	112

Der ønskes fundet det tilsætningsstof der i middel giver den mindste urenhed. ◆

Faktoren “tilsætningsstof” siges at være en **kvalitativ** faktor på 4 **niveauer**.

Havde man eksempelvis i stedet på 4 tidspunkter målt mængden af uønsket stof

Tid [i minutter fra starttidspunkt]	0	10	20	30
Mængde urenhed	108	105	108	117
	110	110	111	119
	112	109	113	112

siges faktoren “tid” at være en **kvantitativ** faktor.

En kvantitativ faktor er altså en “talfaktor”, hvor det også har mening at spørge om mængde urenhed for mellemliggende værdier. Kvalitative faktorer er derimod ikke talbestemmende, og hvor det naturligvis ikke har mening at se på mellemliggende værdier (såsom tilsætningsstof nr. $T_{1.53}$).

Problemer, hvor faktorerne er kvalitative analyseres ved en “variansanalyse. Er faktorerne alle kvantitative, vil en metode kaldet “regressionsanalyse” være at foretrække.

Hvis nogle faktorer er kvantitative og nogle er kvalitative, kan man dog godt analysere problemet med variansanalyseteknikken, men da findes der mere effektive metoder, som dog ikke behandles i dette notat.

11.2 Ensided variansanalyse

Vi vil i dette afsnit behandle problemer, af den type, som er vist i eksempel 11.1, dvs. med én faktor på mere end 2 niveauer. I eksemplet hedder faktoren "Tilsætningsstof" og den er på 4 niveauer.

Det vil i sådanne tilfælde være af interesse at teste om de til niveauerne svarende middelværdier afviger fra hinanden og i bekræftende fald hvilket niveau der giver den største/mindste værdi. I eksempel 11.1 ønskes det således at finde det stof, der giver den mindste "middelurenhed".

Umiddelbart kunne man synes, at så foretager vi blot de samme parvise sammenligninger som i kapitel 10, hvor vi så på differenser mellem 2 middelværdier. Problemet er imidlertid, at selv om de forskellige tilsætningsstofferne giver samme udbytte, så ville støjen i forsøget bevirke, at de mange gennemsnit fordeler sig "klokkeformet" (normalfordelt), og det vil sige, at den største og den mindste værdi let vil ligge så langt fra hinanden, at man ved at teste på deres differens fejlagtigt slutter, at der er forskel, selv om det faktisk ikke er tilfældet (fejl af type 1).

For at undgå dette, skal man derfor **altid** starte med at foretage den i det følgende beskrevne variansanalyse. Giver den, at der ikke er signifikant forskel på middelværdierne, så skal man rette sig efter det, og ikke derefter begynde at se på konfidensintervaller.

Giver analysen, at der er en signifikant forskel, så ved man, at der i hvert fald er en signifikant forskel mellem det største og mindste middelværdi. Man kan så ved hjælp af passende konfidensintervaller forsøge at finde ud af om der også er en signifikant forskel mellem den største og næststørste værdi osv.

I afsnit 11.2.1 gives en forståelse for den teoretiske baggrund for variansanalyser. Forklaringen understøttes af et eksempel, hvor regningerne foretages med lommeregner uden et egentligt statistikprogram.

En oversigt over de anvendte formler findes i oversigt 11.1.

I samme eksempel vises dog også, hvorledes man ved hjælp af TI-89 og Statgraphics let kan finde de samme størrelser.

11.2.1.Forklaring af metode og formler

Vi betragter igen eksempel 11.1.

Opstilling af nulhypotese.

Lad X_i = mængden af uønsket stof ved tilsætning af stof T_i , hvor $i \in \{1,2,3,4\}$

Idet de 4 variables middelværdier kaldes μ_1, μ_2, μ_3 og μ_4 ønsker vi at teste nulhypotesen

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$, mod H : "mindst én middelværdierne er forskellig fra de øvrige".

Forsøgets udførelse. Forsøget skal udføres som et **fuldstændigt randomiseret forsøg**. (jævnfør kapitel 10 hvor et sådant forsøg er beskrevet). Derved sikrer vi, at der udføres et "statistisk gyldigt" forsøg. Hvis vi derfor, efter at have foretaget en ensidet variansanalyse, konkluderer, at der er forskel på tilsætningsstofferne, så er det "korrekt", idet det ville være helt tilfældigt, hvis én af tilsætningsstofferne har været begunstiget med særlig gode forsøgseenheder.

Beregning af gennemsnit og spredning. For at få et skøn for mængden af urenheder, udregnes gennemsnittene for hvert tilsætningsstof. Disse er angivet i nedenstående skema. Umiddelbart ud fra gennemsnit synes T_4 at adskille sig fra de tre øvrige, men hvis der er stor spredning, kan det måske blot være et tilfælde. Det er derfor naturligt at udregne spredningerne, hvilket derfor også er anført i skemaet.

	T_1	T_2	T_3	T_4
Gennemsnit	110.0	108.0	110.67	116.0
Spredning	2.000	2.6458	2.5166	3.6056

Forudsætninger.

- 1) De 4 variable T_1, T_2, T_3 og T_4 skal være statistisk uafhængige.
En måling af mængden af urenhed eksempelvis med tilsætningsstoffet T_3 må ikke afhænge af hvilke målinger der inden da er sket.
Kravet må opfyldes ved en hensigtsmæssig forsøgsplan, bl.a. ved at randomisere.
- 2) De 4 variable T_1, T_2, T_3 og T_4 skal være tilnærmelsesvis normalfordelte.
- 3) De 4 variable T_1, T_2, T_3 og T_4 skal have samme varians (kort: der skal være varianshomogenitet)
Kravene i punkt 2 og 3 er sædvanligvis opfyldt i praksis, da det heldigvis er sådan, at analysen stadig er gyldig selv om der foreligger mindre afvigelser, hvis blot antallet af gentagelser for hvert niveau er den samme. Dette gælder i eksempel 11.1, hvor der er 3 gentagelser for hvert tilsætningsstof.
Man siger kort, at analysen er robust overfor afvigelser fra normalitet og varianshomogenitet, hvis blot antallet af gentagelser er den samme.
Vi vil derfor sædvanligvis antage, at forudsætningerne er opfyldt.

Er der derimod forskelligt antal gentagelser, bliver analysen følsom overfor afvigelser.

Er man i tvivl om kravet om normalitet er rimeligt opfyldt, kan man få et grafisk indtryk af, om der er alvorlige afvigelser, ved at tegne såkaldte normalfordelingsplot.

Der findes en lang række test som undersøger om kravet om varianshomogenitet er rimeligt opfyldt. Her kan nævnes ("simplificerede F - test", "Bartlett's test", "Levines test") som alle er beskrevet i oversigt 11.1 og 11.2).

Er forudsætningerne ikke opfyldt kan man eventuelt udføre en "rangtest" (se eventuelt kapitel 15)

Vi vil i det følgende antage, at forudsætningerne er opfyldte.

Pooling. Da de 4 varianser antages at være nogenlunde ens, beregnes et vægtet gennemsnit, i forhold til frihedsgraderne (man foretager en "pooling").

Lad $N = n_1 + n_2 + n_3 + n_4 (=12)$ og $r = \text{antal niveauer} (=4)$

$$s_e^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2 + (n_3 - 1) \cdot s_3^2 + (n_4 - 1) \cdot s_4^2}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1) + (n_4 - 1)} = \frac{(3 - 1) \cdot 2.000^2 + (3 - 1) \cdot 2.6458^2 + (3 - 1) \cdot 2.5166^2 + (3 - 1) \cdot 3.6056^2}{N - r}$$

$$= \frac{60.6667}{8} = \frac{SAK_0}{8} = 7.583$$

s_e^2 er variansen for "forsøgsfejlen" eller på engelsk "error".

s_e^2 har $N - r = 8$ frihedsgrader. Det kan også ses af, at da hver af varianserne s_i^2 er baseret på $n = 3$ målinger har de hver 2 frihedsgrader ($f = n - 1 = 3 - 1$). s_e^2 har derfor $2 \cdot 4 = 8$ frihedsgrader.

Beregning af F - test. Antages nulhypotesen at være sand, dvs. udbyttet fra de 4 tilsætningsstoffer har samme middelværdi, er den eneste grund til, at vi ikke får samme gennemsnit i de 4 tilfælde, den ukontrollable “støj” (forsøgsvariablens variation) som forekommer ved forsøgets udførelse.

Indtastes de fire gennemsnit i en lommeregner findes $s_{\bar{x}}^2 = 11.667$.

Et gennemsnit af n tal har en varians, der er n gange mindre end variansen på den enkelte måling. I dette tilfælde er $n = 3$.

Et estimat for støjens varians forudsat nulhypotesen er sand er derfor

$$s_R^2 = 3 \cdot s_{\bar{x}}^2 = 3 \cdot 11.667 = 35.00. \text{ Frihedsgradstallet er } f_R = \text{antal niveauer} - 1 = 4 - 1 = 3.$$

Hvis nulhypotesen er sand burde $s_R^2 \approx s_e^2$ eller $\frac{s_R^2}{s_e^2} \approx 1$, mens hvis nulhypotesen er falsk

(middelværdierne er forskellige) er $s_R^2 > s_e^2$ dvs. forholdet $F = \frac{s_R^2}{s_e^2}$ være signifikant større end 1.

Da $F = \frac{s_R^2}{s_e^2} = \frac{35}{7.583} = 4.62$ er spørgsmålet derfor, om dette tal er signifikant større end 1.

Da forholdet mellem de to varianser (som sædvanlig) er F - fordelt med $f_R = 3$ frihedsgrader i tælleren og $f_e = 8$ i nævneren kan vi afgøre dette ved at regne P - værdien ud.

Konfidensintervaller. Disse beregnes **kun** hvis vi får en forkastelse af nulhypotesen, og dermed ved, at den største og den mindste middelværdi er signifikant forskellige. Om nogle af de øvrige middelværdier er “lige så gode” som den optimale vil ofte være af interesse.

De “sædvanlige” konfidensintervaller for hvert niveau bestemmes ved (jævnfør oversigt 11.1):

$$\bar{x}_i \pm r_{kon} \text{ hvor } r_{kon} = t_{1-\frac{\alpha}{2}}(N-r) \cdot \sqrt{\frac{s_e^2}{n}}$$

$$\text{I vort tilfælde er } r_{kon} = t_{1-\frac{\alpha}{2}}(N-r) \cdot \sqrt{\frac{s_e^2}{n}} = t_{0.975}(8) \sqrt{\frac{7.583}{3}} = 2.31 \cdot 1.589 = 3.673.$$

og konfidensintervallet $\bar{x}_i \pm r_{kon} = \bar{x}_i \pm 3.673$

Konfidensintervallerne for 2 middelværdier overlapper derfor ikke, hvis afstanden er større end $2r_{kon}$. Imidlertid vil disse intervaller være lidt for “brede”, dvs. selv om der faktisk er en forskel på 2 middelværdier, så overlapper intervallerne hinanden, så man opdager det ikke.

Man kunne så foretage parvise sammenligninger svarende til de konfidensintervaller vi fandt i kapitel 10. Her vil man kunne opdage en forskel, hvis middelværdiernes afstand er større end $r_{kon} \sqrt{2}$.

Problemet er imidlertid her, at hvis vi har n middelværdier, så vil der være $\frac{n(n-1)}{2}$ parvise sammenligninger. For hver af disse sammenligninger er der jo en vis sandsynlighed for at begå en fejl af type 1, dvs. påstå der er en forskel som reelt ikke er der. Sådanne fejl vil jo hobe sig op, hvis man foretager mange sammenligninger, så sandsynligheden for at begå en fejl af type 1 kunne blive betragtelig.

Konfidensintervaller, der beror på parvise sammenligninger kaldes LSD intervaller (Least

Significance Difference), og er bestemt ved, at deres radius er $r_{lsd} = r_{kon} \frac{\sqrt{2}}{2} = \frac{r_{kon}}{\sqrt{2}}$

Vi har derfor $\bar{x}_i \pm r_{lsd} = \bar{x}_i \pm \bar{x}_i \pm \frac{3.673}{\sqrt{2}} = \bar{x}_i \pm 2.597$

De fleste statistikprogrammer har en række andre metoder til beregning af konfidensintervaller, som søger at formindske sandsynligheden for at begå fejl af type 1 og type 2.

Statgraphics har således indbygget Tukey, Scheffe og Bonferroni's metoder.

Vi vil i dette notat kun se på ovennævnte to typer, og hvis vi har få middelværdier stole mest på LSD-intervallerne.

11.2.2. Beregning af ensidet variansanalyse

Vi vil regne eksempel 11.1 ved at benytte formlerne i oversigt 11.1

I appendix 11A og 11B regnes det tilsvarende eksempel ved benyttelse af henholdsvis TI - 89 og Statgraphics.

Eksempel 11.2 (én faktor). *Virkningerne af 4 tilsætningsstoffer T_1, T_2, T_3, T_4 på mængden af urenheder ved en kemisk proces ønskes sammenlignet. For hvert tilsætningsstof måles mængden af "uønsket stof" 3 gange.*

Forsøgsresultaterne blev følgende:

Tilsætningsstof	T_1	T_2	T_3	T_4
Mængde	108	105	108	117
urenhed	110	110	111	119
	112	109	113	112

Der ønskes fundet det tilsætningsstof der i middel giver den mindste urenhed.

Løsning:

X_i = mængden af uønsket stof ved tilsætning af stof T_i hvor $i \in \{1,2,3,4\}$

X_i antages approksimativt normalfordelt med middelværdien μ_i og spredning σ_i .

Vi ønsker at teste

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$, mod H : "mindst én middelværdierne er forskellig fra de øvrige".

Antal forsøg $N = 12$. Antal niveauer $r = 4$. Antal gentagelser $n = 3$.

Gennemsnit og spredning beregnes:

	T_1	T_2	T_3	T_4
Gennemsnit	110.0	108.0	110.67	116.0
Spredning	2.000	2.6458	2.5166	3.6056

$$s_e^2 = \frac{2.000^2 + 2.6458^2 + 2.5166^2 + 3.6056^2}{4} = \frac{60.667}{8} = 7.583. \text{ Frihedsgrad } f_e = N - r = 12 - 4 = 8$$

Spredningen s_x^2 på de 4 gennemsnit beregnes: $s_x^2 = 11.667$

$$s_R^2 = n \cdot s_x^2 = 3 \cdot 11.667 = 35.00 \text{ Frihedsgrad } f_R = r - 1 = 3$$

$$F = \frac{s_R^2}{s_e^2} = \frac{35}{7.583} = 4.62. P\text{-værdi} = P(F > 4.62) = \text{FCdf}(4.62, \infty, 3, 8) = 0.037$$

Da $P\text{-værdi} = 0.037 < 0.05$ forkastes nullhypotesen (svagt).

Konklusion: De fire tilsætningsstoffer har ikke samme virkning.

Mindst 2 af middelurenhederne er forskellige.

Konfidensintervaller:

Ud fra gennemsnittene ses, at T_2 er signifikant mindre end T_4 . Om de øvrige tilsætningsstoffer giver lige så lidt urenhed som T_2 må afgøres ved opstilling af konfidensintervaller.

Som nævnt i indledningen, må man kun foretage en vurdering ud fra konfidensintervaller, hvis ovennævnte F - test har vist, at der er en signifikant forskel.

De "sædvanlige" konfidensintervaller for hvert niveau (jævnfør oversigt 11.1):

$$\bar{x}_i \pm r_{kon} \text{ hvor } r_{kon} = t_{1-\frac{\alpha}{2}}(N-r) \cdot \sqrt{\frac{s_e^2}{n}} = t_{0.975}(8) \sqrt{\frac{7.583}{3}} = 2.31 \cdot 1.589 = 3.673$$

$$\bar{x}_i \pm r_{kon} = \bar{x}_i \pm 3.673$$

LSD konfidensinterval:

$$r_{lsd} = r_{kon} \frac{\sqrt{2}}{2} = \frac{r_{kon}}{\sqrt{2}}, \quad \bar{x}_i \pm r_{lsd} = \bar{x}_i \pm \frac{3.673}{\sqrt{2}} = \bar{x}_i \pm 2.597$$

	T_1	T_2	T_3	T_4
Gennemsnit	110.0	108.0	110.67	116.0
95% konfidensinterval	[106.33;113.67]	[104.33;111.67]	[107.00;114.33]	[112.33;119.67]
95% LSD-konfidensinterval	[107.40;112.60]	[105.40;110.60]	[108.07;113.26]	[113.40;118.60]

Konklusion: Begge metoder viser, at man får den mindste urenhed, hvis man vælger enten T_1 , T_2 eller T_3 (de kan ikke adskilles).



11.3 Tosidet variansanalyse

11.3.1 Indledning.

Har man 2 kvalitative faktorer vil det også være naturligt at udføre en variansanalyse, men da man her kan risikere, at de to faktorer "spiller sammen" på en uventet måde, bliver forholdene noget mere kompliceret. Til gengæld kan begreberne her så umiddelbart generaliseres til forsøg med mere end 2 faktorer.

11.3.2 Planlægning af forsøg.

I dette afsnit benyttes følgende eksempel som illustration af begreberne.

Eksempel 11.3. En bilfabrikant ønsker at finde ud af hvorledes 3 olieblandinger O_1 , O_2 , og O_3 , og 2 karburatortyper K_1 og K_2 påvirker benzinforbruget.

Vi har et forsøg med 2 **kvalitative faktorer**: olieblanding og karburator.

Faktoren "olieblanding" er på 3 **niveauer** O_1 , O_2 , og O_3 , mens faktoren "karburator" har 2 niveauer nemlig K_1 og K_2 .

Man har (ved en dimensionering) fundet, at for at få de tilstrækkelig relevante oplysninger skal hvert niveau gentages mindst 4 gange.

11.3.2.1 Een faktor ad gangen

I mange forsøgsvejledninger står, at man bør kun variere en faktor ad gangen. Alle andre faktorer end den udvalgte fastholdes på et bestemt niveau.

En forsøgsplan efter disse retningslinier kunne eksempelvis være som skitseret nedenfor, hvor hvert delforsøg er markeret med et $\times$:

		Karburator	
		K_1	K_2
Olieblanding	O_1	$\times \times \times \times$	
	O_2	$\times \times \times \times$	$\times \times \times \times$
	O_3	$\times \times \times \times$	

I dette eksempel, hvor der kun er 2 faktorer, vælger vi først at variere olieblandingen, mens den anden faktor fastholdes.

Idet vi har valgt først at fastholde karburatoren på niveauet K_1 , kan forsøget udføres således:

12 af de 16 biler, som skal anvendes, udstyres med karburator K_1 , og derefter (randomiseret) får 4 af disse biler olieblanding O_1 , 4 andre biler olieblanding O_2 , og de sidste 4 biler olieblanding O_3 .

Efter at have kørt en udvalgt strækning måles benzinforbruget.

Derefter varieres den anden faktor (her karburator), mens olieblandingen fastholdes på O_2 , dvs. de sidste 4 biler udstyres med karburator K_2 og olieblanding O_2 .

Igen gennemkøres den udvalgte strækning, og benzinforbruget måles.

Det er vigtigt, at **hver behandling har lige mange gentagelser**.

Da hvert af de 4 niveauer skal gentages mindst 4 gange, så må hver behandling gentages 4 gange, så der er i alt 16 delforsøg.

Indtegnes for hver karburator det gennemsnitlige benzinforbrug mod olie-blandingen kunne vi eksempelvis få tegningen på fig. 11.1.

Umiddelbart ses, at K_1 giver lavest benzinforbrug, og O_1 (eller O_3) skal foretrækkes.

Hvad nu med benzinforbruget i karburator K_2 , hvis vi anvender olieblending O_1 eller O_3 ?

Kan man slutte, at benzinforbruget ved olieblending O_1 og O_3 er lavere, når man bruger karburator K_1 , end når man bruger karburator K_2 ?

Kun, hvis man ud fra tekniske eller andre grunde mener at vide, at "karburatorkurven" for K_2 er parallel med kurven for K_1 , så er forsøgsplanen anvendelig, men ikke den bedste.

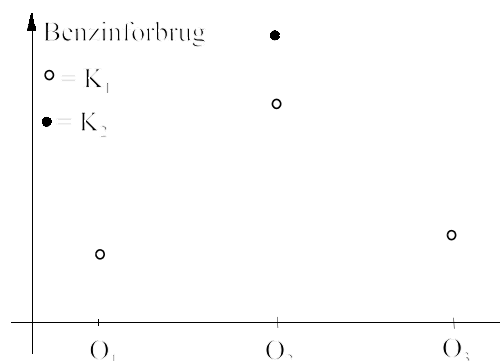


Fig 11.1 Skitse af benzinforbrug

En statistisk set bedre forsøgsplan som endda ofte er mindre ressourcekrævende, er følgende:

11.3.2.2 Fuldstændig faktorstruktur

Denne plan består i, at hvert niveau af den ene faktor kombineres med ethvert niveau af den anden. Planen kan skitseres således:

		Karburator	
		K_1	K_2
Olieblanding	O_1	× ×	× ×
	O_2	× ×	× ×
	O_3	× ×	× ×

Her er hver af de 6 behandlinger gentaget 2 gange, dvs. i alt er der udført 12 delforsøg. Hermed er kravet opfyldt om at hvert niveau skal gentages mindst 4 gange.

I "en faktor ad gangen" var vi tvunget til at udføre 16 delforsøg, mens vi kun skal lave 12 delforsøg i det "fuldstændige faktorforsøg".

Vi kan altså nøjes med færre delforsøg, når vi laver et fuldstændigt faktorforsøg.

Indtegnes for hver karburator det gennemsnitlige benzinforbrug mod olie-blandingen, kan det eksempelvis vise sig, at man får figur 11.2.

Vi ser, i modstrid med hvad vi antog ud fra "en faktor ad gangen forsøget", at kombinationen af katalysator K_2 og olieblending O_1 giver det laveste benzinforbrug.

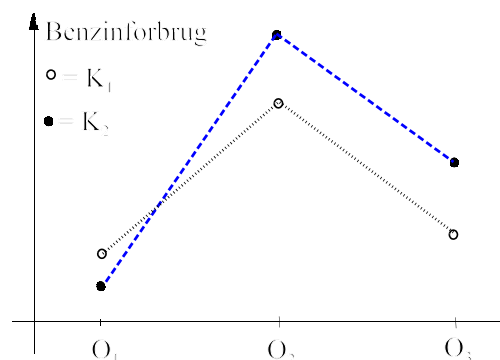


Fig 11.2. Vekselvirkning

Det ses, at de to kurver **ikke er parallelle**. Dette kunne være tilfældigt og blot skyldes forsøgets "støj", men det kunne også være signifikant, og derfor være udtryk for en såkaldt "**vekselvirkning**".

En model uden vekselvirkning (kurverne tilnærmelsesvis parallelle) siges at være **additiv**.

11.3.3. Formler og metode

Vi vil i det følgende kun analysere forsøg med en fuldstændig faktorstruktur, og hvor hver behandling har lige mange gentagelser (samme antal delforsøg i hver "celle").

Et eksempel på et sådant forsøg er

Eksempel 11.4. En bilfabrikant ønsker at finde ud af, hvorledes 3 olieblandinger O_1 , O_2 , og O_3 , og 2 karburatorer K_1 og K_2 påvirker benzinforbruget.

Forsøgsresultaterne er følgende:

		Karburator			
		K_1		K_2	
Olieblanding	O_1	830	860	810	840
	O_2	940	990	1050	1020
	O_3	855	815	930	910

Angiv hvilke kombinationer af karburator og olieblanding der giver det laveste forbrug, og giv et estimat for dette forbrug.

Symbolik:

Lad os kalde rækkefaktoren for R , antal rækkenniveauer r , søjlefaktoren for C , antal søjleniveauer for q og antal gentagelser af hver behandling n . Det totale antal delforsøg er følgelig $N = r \cdot q \cdot n$.

I eksempel 11.4 er R = "olieblanding", $r = 3$, C = "karburator", $q = 2$, $n = 2$ og $N = 12$.

Forudsætninger

Disse er de samme som ved den ensidede variansanalyse. Analysen er også her robust overfor afvigelser fra normalitet og varianshomogenitet, blot antallet af gentagelser i hver celle er den samme.

Opstilling af nulhypotese

Det første man skal teste må være, at undersøge om modellen er additiv, dvs. om den er uden en signifikant vekselvirkning.

Nulhypotesen skrives så kort $H_0: R \cdot C = 0$ (faktorerne vekselvirker ikke)

og den alternative hypotese $H: R \cdot C \neq 0$ (faktorerne vekselvirker)

Der er nu to muligheder:

- 1) H_0 forkastes, dvs. faktorerne vekselvirker

Ved hjælp af konfidensintervaller for alle $r \cdot q$ celler søger man at finde den optimale kombination af faktorer.

- 2) H_0 accepteres, dvs. modellen er additiv.

Man tester nu nulhypoteserne

$H_0: R = 0$ (rækkefaktoren har en virkning)

$H_0: C = 0$ (søjlefaktor har en virkning)

Finder man eksempelvis at $H_0: R = 0$ forkastes, mens $H_0: C = 0$ accepteres, konkluderes, at kun rækkefaktoren har en virkning, og man opstiller eventuelt konfidensintervaller til bestemmelse af det optimale niveau.

Beregninger: Hvorledes man foretager testen ved hjælp af en lommeregner er beskrevet i oversigt 11.2. Da specielt beregningerne af SAK'erne er temmelig omfattende, og næppe giver en dybere forståelse, vil vi dog sædvanligvis benytte TI-89 eller Statgraphics hertil. Programmerne dækker dog ikke alle muligheder, så i enkelte tilfælde må man derfor benytte formlerne for eksempelvis konfidensintervaller.

Anskuelig forklaring på hvorledes man kan beregne vekselvirkning: Her gives kun en kort forklaring, som kan tjene til at forstå baggrunden for beregningerne, der i øvrigt med fordel kan foretages af et statistikprogram.

I nedenstående skema er skitseret et forsøg med 2 faktorer R og C. R er på 3 niveauer, og C er på 4 niveauer. Der er 2 gentagelser af hver "behandling" (treatment).

	C ₁	C ₂	C ₃	C ₄
R ₁	1 3	2 6	1 5	6 8
R ₂	7 11	8 14	8 12	13 15
R ₃	2 6	5 7	3 7	8 10

For hver af de 12 celler kan man udregne et skøn for spredningen. Hvis man forudsætter at spredningen er nogenlunde den samme i alle 12 tilfælde, kan man poole de 12 s^2 sammen til et fælles skøn s_0 for spredningen på forsøgsfejlen (støjen). Den vil have 12 frihedsgrader, da hvert enkelt s har 1 frihedsgrad.

I nedenstående skema er beregnet gennemsnit for hver celle, hver række, hver søjle og totalt.

	C ₁	C ₂	C ₃	C ₄	Gennemsnit
R ₁	2	4	3	7	4
R ₂	9	11	10	14	11
R ₃	4	6	5	9	6
Gennemsnit	5	7	6	10	7

Tallene er konstrueret således, at vi har en helt præcis model uden vekselvirkning ($R_2 = R_1 + 7$, $R_3 = R_1 + 2$). For en sådan model gælder helt præcist, at resultatet i celle (i, j) fås af formelen $RC_{ij} = R_i + C_j$ - totale gennemsnit.

Eksempel: $RC_{2,3} = 10$ og $R_2 + C_3$ - totale gennemsnit = $11 + 6 - 7 = 10$.

I praksis vil dette naturligvis aldrig være tilfældet på grund af den tilfældige variation (støj), men udregnes kvadratet på afvigelse (SAK), og disse afvigelser ikke er større end hvad er rimeligt i forhold til støjen (S_0), vil vi kunne konkludere at der ikke kan konstateres nogen vekselvirkning.

11.3.4 Beregning af tosidet variansanalyse

Som nævnt er der 2 hovedtilfælde nemlig om der konstateres vekselvirkning eller ej.

Vi vil derfor dele i det følgende regne to eksempler som illustrerer hver sin situation.

11.3.4.1. Model med vekselvirkning.

Eksempel 11.4. En bilfabrikant ønsker at finde ud af, hvorledes 3 olieblandinger O_1 , O_2 , og O_3 , og 2 karburator typer K_1 og K_2 påvirker benzinforbruget.

Forsøgsresultaterne er følgende:

		Karburator			
		K ₁		K ₂	
Olieblanding	O ₁	830	860	810	840
	O ₂	940	990	1050	1020
	O ₃	855	815	930	910

Angiv hvilke kombinationer af karburator og olieblanding der giver det laveste forbrug, og giv et estimat for dette forbrug.

Løsning:

I oversigt 11.2 er de såkaldte SAK'er beregnet, og ud fra dem er det nemt at opstille en såkaldt variansanalysetabel

Man finder

Variation	SAK=SS	f=df	$s^2 = \frac{SAK}{f}$	F
Rækkefaktor R : Olieblanding	58716.67	$f_R = r - 1 = 2$	29358.33	
Søjlefaktor C : Karburator	6075.00	$f_C = q - 1 = 1$	6075.00	
Vekselvirkning R*C	6450.00	$f_{RC} = (r - 1)(q - 1) = 2$	3225.00	$F_{RC} = \frac{s_{RC}^2}{s_0^2} = 5.38$
Gentagelser (Residual, Error)	3600.00	$f_e = r \cdot q \cdot (n - 1) = 6$	600.00	
Total	74841.67	$f_{total} = N - 1 = 11$		

- 1) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning)

P - værdi = $P(F > F_{RC} = 5.38) = \text{FCdf}(5.38, \infty, 2, 6) = 0.04588$

Da P - værdi = $0.0459 < 0.05$ forkastes H_0 (svagt).

Konklusion: Begge faktorer har en virkning i form af en vekselvirkning.

- 2) Beregning af 95% konfidensintervaller:

Ifølge oversigt 11.2 er konfidensintervaller for hver celle:

$[\bar{x}_{ij} - r_{kon}; \bar{x}_{ij} + r_{kon}]$, hvor radius er

$$r_{kon} = t_{1-\frac{\alpha}{2}}(r \cdot q \cdot (n - 1)) \frac{s_e}{\sqrt{n}} = t_{0.975}(3 \cdot 2 \cdot (2 - 1)) \frac{\sqrt{600}}{\sqrt{2}} = 2.45 \cdot \sqrt{300} = 42.44$$

Vi finder gennemsnittene i hver celle

Gennemsnit		Karburator	
		K ₁	K ₂
Olieblanding	O ₁	845	825
	O ₂	965	1035
	O ₃	835	920

95% konfidensinterval		Karburator	
		K ₁	K ₂
Olieblanding	O ₁	[802.6 ; 887.3]	[782.6 ; 867.4]
	O ₂	[922.6 ; 1007.4]	[992.6 ; 1077.4]
	O ₃	[792.6 ; 877.4]	[877.6 ; 962.4]

Det ses, at umiddelbart giver K_2 O₁ det laveste benzinforbrug, men af konfidensintervallerne ses, at der ingen signifikant forskel er mellem K_2 O₁, K_1 O₃ og K_1 O₁.



11.3.4.2. Additiv model: Model uden vekselvirkning**Eksempel 11.5**

I forbindelse med nogle brudstyrkebestemmelser for Portland-cement udføres et fuldstændigt randomiseret forsøg til undersøgelse af middelbrudstyrkens afhængighed af cementblandere og cementknusere.

Med hver af 4 cementblandere udstøbtes efter blanding med vand 12 cementterninger, som efter en uges lagring underkastedes en brudstyrkeprøve ved hjælp af en af 3 cementknusere.

Forsøgsresultaterne var:

		Cementknusere											
		1			2			3			4		
Cementblandere	1	147	175	130	99	85	75	67	23	35	215	97	180
	2	211	145	163	131	100	145	75	45	71	151	157	167
	3	123	85	153	137	143	82	67	25	83	135	91	129

- 1) Angiv hvilke kombinationer af cementblander og cementknuser, der giver den største brudstyrke, og giv et estimat og et 95% konfidensinterval for denne største middelbrudstyrke.
- 2) Da cementknuser 2 og cementblander 1 er de billigste ønskes opstillet under hensyntagen til resultatet af 1) et 95% - konfidensinterval for middelværdien af brudstyrken, når disse anvendes.

Løsning:

Lad os antage, at SAK'erne er beregnet (se evt. oversigt 11.2). Ud fra dem er det nemt at opstille en såkaldt variansanalyse tabel

Man finder

Variation	SAK=SS	f=df	$s^2 = \frac{SAK}{f}$	F
Rækkefaktor R : Cementblander	$SAK_R = 4299.29$	$f_R = r - 1 = 2$	$s_R^2 = 2149.69$	
Søjlefaktor C : Cementknuser	$SAK_C = 51995.2$	$f_C = q - 1 = 3$	$s_C^2 = 17331.7$	
Vekselvirkning R*C	$SAK_{RC} = 7122.61$	$f_{RC} = (r-1)(q-1) = 6$	$s_{RC}^2 = 1187.1$	$F_{RC} = \frac{s_{RC}^2}{s_e^2} = 1.33$
Gentagelser (residual, error)	$SAK_e = 21270.7$	$f_e = r \cdot q \cdot (n-1) = 24$	$s_e^2 = 886.28$	
Total	84687.9	$f_{total} = N - 1 = 35$		

- 1) a) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning)

$$P\text{-værdi} = P(F > F_{RC} = 1.33) = \text{FCdf}(1.33, \infty, 6, 24) = 0.2787$$

Da $P\text{-værdi} = 0.2787 > 0.05$ accepteres H_0 .

Konklusion: Vi antager i det følgende, at vekselvirkningen er forsvindende.

b) Vi antager nu, at $s_{RC}^2 = 1187.1$ også er et estimat for "støjens" varians.

Inden vi tager stilling til om faktorerne har en hovedvirkning, "pooles" derfor s_{RC}^2 med s_e^2 til

$$\text{et nyt estimat } s_m^2 = \frac{SAK_e + SAK_{RC}}{f_e + f_{RC}} = \frac{28393.3}{30} = 946.443.$$

$$F_{knuser} = \frac{17331.7}{946.44} = 18.31. \quad P\text{-værdi} = P(F > 18.31) = \text{FCdf}(18.31, \infty, 3, 30) = 6.09 \cdot 10^{-7}$$

$$F_{blander} = \frac{2149.69}{946.44} = 2.271. \quad P\text{-værdi} = P(F > 2.271) = \text{FCdf}(2.271, \infty, 2, 30) = 0.121$$

Det giver følgende tabel:

Variation	SAK=SS	f = df	s^2	F	P - værdi
Cementblander : R	4299.39	2	2149.69	2.271	0.121
Cementknuser : C	51995.2	3	17331.7	18.31	0.0000006
Residual	28393.3	30	946.443		

H_0 : C = 0 (Cementknuser har ingen virkning) forkastes, da P - værdi = 0.000 < 0.05

H_0 : R = 0 (Cementblander har ingen virkning) accepteres, da P - værdi = 0.121 > 0.05

Konklusion: Cementknuserne har en (stærk) virkning

Cementblandere har ingen virkning

Beregning af 95% konfidensintervaller:

Da cementblandere ingen virkning har pooles s_R^2 med s_m^2 (se oversigt 11.2 punkt 4)

$$s_{m_2}^2 = \frac{SAK_R + SAK_m}{f_R + f_m} = \frac{32692.7}{32} = 1021.65 \text{ med frihedsgradstal } f_{m_2} = 32$$

$[\bar{x}_{ij} - r_{kon}; \bar{x}_{ij} + r_{kon}]$, hvor radius er

$$r_{kon} = t_{0.975}(f_{m_2}) \frac{s_{m_2}}{\sqrt{n \cdot r}} = t_{0.975}(32) \frac{\sqrt{102165}}{\sqrt{3 \cdot 3}} = 2.26 \cdot 10.65 = 24.07$$

Vi finder gennemsnittene i hver søjle

	Cementknusere			
	1	2	3	4
Gennemsnit	148.0	110.8	54.56	146.89
Konfidensinterval	[126.3; 169.7]	[89.1; 132.5]	[32.9; 76.3]	[125.2; 168.6]

Konklusion: Cementknuser 1 og 4 må foretrækkes, selv om man ikke helt kan afvise at cementknuser 2 kan være lige så god.

Et estimat for største middelbrudstyrke: 148 95% konfidensinterval [126.3; 169.7]

2) Cementknuser 2: Middelbrudstyrke: 110.8 95% konfidensinterval [89,1; 132.5]



11.4 Fuldstændigt randomiseret blokforsøg.

I forbindelse med planlægningen af et forsøg, kan man blive tvunget til at benytte forsøgsenheder, som er ret uensartede. Derved får den tilfældige forsøgsfejl en relativ stor spredning (stor "støj"). Dette kan bevirke, at man skal op på et urealistisk stort antal gentagelser for at kunne opnå den ønskede information. For at "dæmpe støjen" kan man inddеле forsøgsenhederne i grupper (blokke), hvor de forsøgsenheder der ligger i samme blok er væsentlig mere ensartede end forsøgsenhederne i forskellige blokke. Man siger, at man har et fuldstændigt randomiseret blokforsøg, hvis hver behandling forekommer det samme antal gange (sædvanligvis netop én gang) i hver blok.

Til illustration heraf, så betragter vi igen forsøget beskrevet i eksempel 11.2.

Eksempel 11.6 (randomiseret blokforsøg). En bilfabrikant ønsker at finde ud af, hvorledes 3 olieblandinger O_1 , O_2 , og O_3 , og 2 karburatortyper K_1 og K_2 påvirker benzinforbruget. Forsøget planlægges som et fuldstændigt faktorforsøg idet hvert niveau skal gentages mindst 4 gange. Dette betyder at der skal udføres 12 delforsøg.

Et delforsøg med én bil tager 1 dag. (1 tank = 40 liter: Kører ca. 15 km/l så 40 liter = 600 km, hvilket giver ca. 7 timer med 80 km/time). Af tidsmæssige grunde kan man ikke benytte 12 dage til forsøget. Der benyttes 2 biler med tilhørende chauffør, hvilket forkorter forsøgstiden til 6 dage.

Da de to biler (med tilhørende chauffør) kan frygtes at give systematisk forskellige resultater, ønskes foretaget et **randomiseret blokforsøg** med biler som blokke.

- 1) Angiv fordele og ulemper ved at foretage et randomiseret blokforsøg fremfor et fuldstændigt randomiseret forsøg.
- 2) Beskriv hvorledes en randomisering kunne tænkes at foregå.
- 3) Skitser udseendet af en variansanalysetabel med angivelse af frihedsgrader.

Løsning:

- 1) Fordele: Begrundelsen for ikke at foretage et fuldstændigt randomiseret forsøg er, at to biler frygtes at give så stor spredning, at selv betydelige forskelle ikke kan påvises.

Ved blokforsøget er støjen, der skyldes eventuelle forskelle mellem biler elimineret

Større mulighed for stabilt vejr i 6 dage end i 12 dage, hvilket også formindsker spredningen.

Ret få delforsøg,

Ulempe: Selv på 6 dage kan vejret skifte og give anledning til stor spredning.

- 2) Randomisering: To dåser mærkes henholdsvis bil 1 og bil 2. Behandlingen O_1K_1 skrives på 2 sedler som anbringes i hver sin dåse, behandlingen O_1K_2 skrives på 2 sedler som anbringes i hver sin dåse osv. (se figuren).

Man trækker nu først de 6 sedler fra dåse med mærket bil 1.

Lad den første seddel der trækkes være O_1K_2 . Det betyder nu, at bil 1 skal forsynes med karburator 2 og olieblending 1 og køre dag 1. Lad den næste seddel der trækkes være O_3K_2 . Det betyder tilsvarende at bil 1 skal forsynes med karburator 2 og olieblending 3 og køre dag 2. Således fortsættes indtil alle 6 sedler er udtrukket

Resultatet blev:

bil 1	dag 1	dag 2	dag 3	dag 4	dag 5	dag 6
	O_1K_2	O_3K_2	O_2K_2	O_1K_1	O_2K_1	O_3K_1

O_1K_1	O_1K_2	O_3K_2
O_3K_1	O_2K_1	O_2K_2
O_1K_2	O_1K_1	O_3K_1

Bil 2

Derefter fortsættes med at trække sedler fra dåsen med mærket bil 2. Resultatet blev:

bil 2	dag 1	dag 2	dag 3	dag 4	dag 5	dag 6
	O ₁ K ₁	O ₃ K ₂	O ₁ K ₂	O ₂ K ₁	O ₃ K ₁	O ₂ K ₂

3) Analyse: Tresidet variansanalyse:

Bemærk: Vi antager altid, at **blokke ikke vekselvirker med faktorerne**, idet vi forudsætter, at den ene blok (eksempelvis bil 1) bidrager med en systematisk højere resultat end den anden blok (eksempelvis at bil 1 på alle dage giver et større benzinforbrug end bil 2).

		K ₁	K ₂	Variansanalyse	SAK	f
Bil 1	O ₁	-	-	Blokke (biler)		1
	O ₂	-	-	Olieblanding		2
	O ₃	-	-	Karburator		1
Bil 2	O ₁	-	-	Olie * karburator		2
	O ₂	-	-	Residual		5
	O ₃	-	-	Total		11

Bemærk: Selv om analysen viser, at blokkene mod forventning ikke kan antages at have betydning, må man **ikke poole blokkene ned**, da det svarer til, at man analyserer forsøget som om det var et fuldstændigt randomiseret forsøg. ◆

Eksempel 11.7 (randomiseret blokforsøg)

I nedenstående tabel er anført resultaterne af et fodringsforsøg med svin. Formålet med forsøget var at undersøge, hvorvidt en ændring af vitaminindholdet i foderet gav en forskel i svinenes vægtforøgelse. Vægtforøgelsen afhænger imidlertid også af det enkelte individs genetiske egenskaber. Et fuldstændigt randomiseret forsøg vil derfor sandsynligvis kunne bevirke, at forsøgsfejls spredning bliver så stor, at intet kan påvises (forsøget drukner i støj). Da grise fra samme kuld må forventes at være mere ensartede, vælger man at lave et randomiseret blokforsøg med kuld som blokfaktor.

Fra hvert af 4 forskellige kuld grise udtages 3 grise, der bliver fodret med hver sin af tre fodertyper A, B og C med forskelligt vitaminindhold.

Forsøgsresultaterne (vægtforøgelse i kg) var

		Fodertype		
		A	B	C
Kuld	1	7.0	14.0	8.5
	2	16.0	15.5	16.5
	3	10.5	15.0	9.5
	4	13.5	21.0	13.5

Test, om der er nogen væsentlig virkning af ændringen i foderets vitaminindhold.

Løsning:

H_0 : Foder har ingen virkning.

Lad os antage, at SAK'erne er beregnet (se evt. oversigt 11.2). Ud fra dem er det nemt at opstille en såkaldt variansanalysetabel

Man finder

Variation	SAK=SS	f = df	$s^2 = \frac{SAK}{f}$	F
Faktor :Foder	54.125	$f_R = r - 1 = 2$	27.06	$F_{foder} = \frac{s_{foder}^2}{s_{residual}^2} = 5.756$
Blok: Kuld	87.7729	$f_C = q - 1 = 3$	29.24	
Residual	28.2083	$f_{residual} = f_{total} - f_{foder} - f_{kuld} = 6$	4.701	
Total	170.063	$f_{total} = N - 1 = 11$		

$$P\text{-værdi} = P(F > 5.756) = \text{FCdf}(5.756, \infty, 2, 6) = 0.0402$$

Da $P\text{-værdi} = 0.0402 < 0.05$ forkastes H_0 , dvs.

Konklusion: Der sker en væsentlig ændring i vægtforøgelsen ved at ændre foderblanding.

Konfidensintervaller kunne vise hvilken foderblanding der skal foretrækkes.

Da der ser ud til at være en blokvirkning, vil vægtforøgelsen afhænge af hvilket kuld der er det bedste, så konfidensintervaller kan kun bruges til relative sammenligninger



OVERSIGT 11.1. Formler til beregning af ensidet variansanalyse

I denne oversigt vises hvorledes man kan beregne en ensidet variansanalyse, blot man har en lommeregner der kan beregne gennemsnit og spredning.

For hvert observationssæt udregnes gennemsnit og spredning.

Faktor	Observationer	Gennemsnit	Spredning
R_1	$x_{11}, x_{12}, x_{13}, \dots, x_{1n}$	$\bar{x}_1.$	s_1
R_2	$x_{21}, x_{22}, x_{23}, \dots, x_{2n}$	$\bar{x}_2.$	s_2
R_3	$x_{31}, x_{32}, x_{33}, \dots, x_{3n}$	$\bar{x}_3.$	s_3
R_r	$x_{r1}, x_{r2}, x_{r3}, \dots, x_{rn}$	$\bar{x}_r.$	s_r

Forudsætning: x_{ij} - værdierne er uafhængige observationer af statistisk uafhængig normalfordelte variable X_i med middelværdi μ_i og samme varians σ^2 .

For hver af de r - faktorniveauer er der lige mange gentagelser n af x - værdier, dvs. i alt $N = n \cdot r$ observationer.

Beregninger:

Man beregner $s_e^2 = \frac{s_1^2 + s_2^2 + \dots + s_r^2}{r}$. s_e^2 har $r(n-1) = N - r$ frihedsgrader.

Man indtaster de r gennemsnit $\bar{x}_1.$, $\bar{x}_2.$, $\dots$, $\bar{x}_r.$ i lommeregneren og finder spredningen $s_{\bar{x}}^2$.

$$s_R^2 = n \cdot s_{\bar{x}}^2. F_R = \frac{s_R^2}{s_e^2}$$

Testprocedure.

Nulhypotese: $H_0: \mu_1 = \mu_2 = \dots = \mu_r \Leftrightarrow H_0: R = 0$

Lad α være signifikansniveau..

H_0 forkastes, hvis P - værdi = $P(Z > F_R) < \alpha$, hvor Z er F - fordelt $(f_T, f_N) = (r-1, N-r)$.

Konfidensintervaller: Lad $r_{kon} = t_{1-\frac{\alpha}{2}}(N-r) \cdot \sqrt{\frac{s_e^2}{n}}$

Konfidensinterval for μ_i : $\left[\bar{x}_i. - r_{kon} ; \bar{x}_i. + r_{kon} \right]$

LSD Konfidensinterval for μ_i : $\left[\bar{x}_i. - \frac{r_{kon}}{\sqrt{2}} ; \bar{x}_i. + \frac{r_{kon}}{\sqrt{2}} \right]$

Variansanalysetabel: (ANOVA = ANalysis Of VAriance)

Variation (Source)	SAK (SS)	f (df)	$s^2 = \frac{SAK}{f}$	F	P - værdi
Behandlinger (Between groups)	SAK_R	$r - 1$	$s_R^2 = \frac{SAK_R}{r - 1}$	$F_R = \frac{s_R^2}{s_0^2}$	
Gentagelser (Within groups) (error)	SAK_e	$N - r$	$s_e^2 = \frac{SAK_0}{N - r}$		
Total	SAK_{total}	$N - 1$			

Der gælder: $SAK_{total} = SAK_R + SAK_e$

Model.

Lad α være signifikansniveau.

Lad $\bar{\mu} = \frac{\mu_1 + \mu_2 + \dots + \mu_n}{r}$.

Vi kan da skrive $\mu_i = \bar{\mu} + (\mu_i - \bar{\mu}) = \bar{\mu} + R_i = \bar{\mu} + \text{Rækkevirkning}$

Varianshomogenitet. $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ Test for, at de variable Y_i har samme varians σ^2

a) Simplificeret F-test.

Lad den største værdi af de k estimerede varianser være $s_{\max}^2$ og den mindste være $s_{\min}^2$.

Beregn teststørrelsen $F = \frac{s_{\max}^2}{s_{\min}^2}$.

Lad Y være F - fordelt med frihedsgraderne $f_{tæller} = f_{nævner} = n - 1$

H_0 forkastes, hvis P - værdi $= P(Y > F) < \frac{\alpha}{2}$.

Hvis nulhypotesen accepteres, så antages kravet om varianshomogenitet at være opfyldt.

Hvis nulhypotesen forkastes, må anvendes en test med større styrke såsom Bartlett's test eller Levines test.

b) Bartlett's test. Denne test er beregningsmæssigt vanskelig, og har den svaghed, at den er særdeles følsom overfor afvigelser fra normalitet.

$$(N - k) \cdot \ln \left(\frac{\sum_{i=1}^k (n_i - 1) \cdot s_i^2}{N - k} \right) - \sum (n_i - 1) \cdot \ln(s_i^2)$$

Beregn teststørrelsen $\chi^2 = \frac{(N - k) \cdot \ln \left(\frac{\sum_{i=1}^k (n_i - 1) \cdot s_i^2}{N - k} \right) - \sum (n_i - 1) \cdot \ln(s_i^2)}{1 - \frac{1}{N - k} - \sum_{i=1}^k \frac{1}{n_i - 1}}$

Lad Y være χ^2 - fordelt med frihedsgrade $k - 1$. H_0 forkastes, hvis P - værdi $= P(Y > \chi^2) < \alpha$.

b) Levines test. God test, som imidlertid kræver mere end 2 gentagelser.

Lad $d_{ij} = |y_{ij} - \hat{y}_i|$ hvor $\begin{cases} i = 1, 2, \dots, k \\ j = 1, 2, \dots, n_i \end{cases}$, hvor $\hat{y}_i$ er medianen af de n_i gentagelser af i 'te behandling.

Man udfører en sædvanlig ensidet variansanalyse på tallene d_{ij}

Median af en række tal Tallene ordnes i voksende rækkefølge:

Ulige antal tal: median = midtertal blandt de ordnede tal,

Lige antal tal: median = gennemsnit af de to midterste blandt de ordnede tal

Eksempel: Tal fra eksempel 11.1.

		$\hat{y}_i$	d_{ij}	En ensidet variansanalyse på d_{ij} giver $F = 0.15$, og dermed P -værdi = 0.9285, dvs. en accept af nulhypotesen.
T_1	108, 110, 112	110	2, 0, 2	
T_2	105, 110, 109	109	4, 1, 0	
T_3	108, 111, 113	111	3, 0, 2	
T_4	117, 119, 112	117	0, 2, 5	

Forklaring på konstruktion af normalfordelingsplot.

Et koordinatsystem har en lodret akse, hvor inddelingen er "normalfordelt", dvs fordelingsfunktionen for en normeret normalfordeling vil i dette koordinatsystem blive en ret linie.

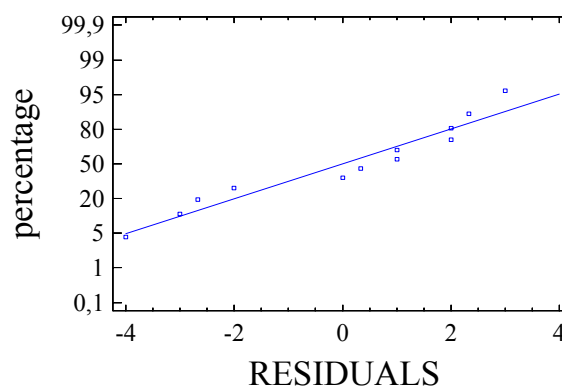
I dette koordinatsystem placeres residualerne som vist:

Lad residualerne (fra eksempel 11.2) -2, 0, 2, -3, 2, 1, -2.7, 0.3, 2.3, 1, 3, -4

De ordnes i rækkefølge og man beregner deres kumulative frekvens

i	Residualer x	$y = \frac{i - 0.5}{12} 100 \%$
1	-4	4.1
2	-3	12.5
3	-2.7	20.8
4	-2	29.1
5	0	37.5
6	0.3	45.8
7	1	54.2
8	1	62.5
9	2	70.8
10	2	79.2
11	2.3	87.5
12	3	95.8

Normal Probability Plot for RESIDUALS



Hvis residualerne er aproksimativt normalfordelt burde punkterne (x,y) afsat i koordinatsystemet tilnærmelsesvis de ligge på en ret linie.

OVERSIGT 11.2. Formler til beregning af tosidet variansanalyse

I denne oversigt vises hvorledes man kan beregne en tosidet variansanalyse, blot man har en lommeregner med gennemsnit og spredning.

Som taleksempel benyttes eksempel 11.2.

Forsøgsresultaterne er følgende:

		Karburator	
		K ₁	K ₂
Olieblanding	O ₁	830 860	810 840
	O ₂	940 990	1050 1020
	O ₃	855 815	930 910

Beregning af gennemsnit.		Karburator		
		K ₁	K ₂	Rækkegennemsnit
Olieblanding	O ₁	845	825	835
	O ₂	965	1035	1000
	O ₃	835	920	877.5
	Søjlesum	881.667	926.667	

Antal rækker $r = 3$, Antal søjler $q = 2$, Antal delforsøg i celler $n = 2$

Antal delforsøg i række = $n \cdot q = 2 \cdot 2 = 4$. Antal delforsøg i søjle $n \cdot r = 2 \cdot 3 = 6$

Antal celler $r \cdot q = 3 \cdot 2 = 6$, Totalt antal forsøg $N = r \cdot q \cdot n = 3 \cdot 2 \cdot 2 = 12$.

Spredning på de r rækkegennemsnit: $s_{xr} = 85.6714$

Spredning på de q søjlegennemsnit: $s_{xq} = 31.8198$

Spredning på de $r \cdot q$ cellegennemsnit: $s_{celler} = 84.4048$

Beregninger:

$$SAK_{\text{rækker}} = (N - n \cdot q) s_{xr}^2 = (12 - 4) \cdot 85.6714^2 = 58716.67, \quad f_R = r - 1 = 2$$

$$SAK_{\text{søjler}} = (N - n \cdot r) \cdot s_{xq}^2 = (12 - 6) \cdot 31.8198^2 = 6075.00, \quad f_C = q - 1 = 1$$

$$SAK_{\text{celler}} = (N - n) \cdot s_{celler}^2 = (12 - 2) \cdot 84.4048^2 = 71241.66, \quad f_{\text{celler}} = r \cdot q - 1 = 5$$

$$SAK_{\text{vekselvirkning}} = SAK_{\text{celler}} - SAK_{\text{rækker}} - SAK_{\text{søjle}} = 6450.00, \quad f_{RC} = f_{\text{celler}} - f_R - f_C = 2$$

$$SAK_{\text{total}} = (N - 1) s_{\text{total}}^2 = (12 - 1) \cdot 82.485^2 = 74841.66, \quad f_{\text{total}} = N - 1 = 11$$

$$SAK_{e(\text{error}=\text{residual})} = SAK_{\text{total}} - SAK_{\text{celler}} = 600.00 \quad f_0 = f_{\text{total}} - f_{\text{celler}} = 6$$

$$(\text{alternativt: } SAK_{\text{rækker}} = n \cdot q \cdot SAK_{\bar{x} \text{ rækker}}, \text{ hvor } SAK_{\bar{x} \text{ rækker}} = (r - 1) \cdot s_{xr}^2$$

$$SAK_{\text{søjler}} = n \cdot r \cdot SAK_{\bar{x} \text{ søjler}}, \text{ hvor } SAK_{\bar{x} \text{ søjler}} = (q - 1) \cdot s_{xq}^2$$

$$SAK_{\text{celler}} = n \cdot SAK_{\bar{x} \text{ celler}}, \text{ hvor } SAK_{\bar{x} \text{ celler}} = (r \cdot q - 1) \cdot s_{\text{celler}}^2)$$

Opstilling af variansanalysetabel:

Variation	SAK=SS	f	$s^2 = \frac{SAK}{f}$	F
Rækkefaktor R : Olieblanding	$SAK_R = 58716.67$	$f_R = r - 1 = 2$	$s_R^2 = 29358.33$	
Søjlefaktor C : Karburator	$SAK_C = 6075.00$	$f_C = q - 1 = 1$	$s_C^2 = 6075.00$	
Vekselvirkning R*C	$SAK_{RC} = 6450.00$	$f_{RC} = (r - 1)(q - 1) = 2$	$s_{RC}^2 = 3225.00$	$F_{RC} = \frac{s_{RC}^2}{s_e^2} = 5.38$
Gentagelser (residual, error)	$SAK_e = 3600.00$	$f_e = r \cdot q \cdot (n - 1) = 6$	$s_e^2 = 600.00$	
Total	74841.67	$f_{total} = N - 1 = 11$		

Test:

Lad α være signifikansniveau.

1) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning)

H_0 forkastes, hvis P -værdi $= P(Z > F_{RC}) < \alpha$, hvor Z er F -fordelt $(f_T, f_N) = (f_{RC}, f_e)$.

2a) Hvis H_0 forkastes, så opstilles konfidensintervaller til nærmere vurdering af faktorerens virkning.

2b) Hvis H_0 accepteres, antages, at der ikke er nogen signifikant vekselvirkning, og man pooler de to varianser sammen, til et nyt estimat for forsøgsfejlens variation (støjen).

$$s_m^2 = \frac{SAK_{RC} + SAK_e}{f_{RC} + f_e} \quad \text{med } f_m = f_{RC} + f_e$$

Dette estimat benyttes så til en samtidig vurdering af hovedvirkningerne.

2b.1) $H_0: R = 0$ (Ingen signifikant virkning af rækkefaktor)

$$\text{Lad } F_R = \frac{s_R^2}{s_m^2}$$

H_0 forkastes hvis P -værdi $= P(Z > F_R) < \alpha$, hvor Z er F -fordelt $(f_T, f_N) = (f_R, f_m)$.

Hvis H_0 forkastes, så opstilles konfidensintervaller til vurdering af faktorerens virkning.

2b.2) $H_0: C = 0$ (Ingen signifikant virkning af søjlefaktor)

$$\text{Lad } F_C = \frac{s_C^2}{s_m^2}$$

H_0 forkastes, hvis P -værdi $= P(Z > F_C) < \alpha$, hvor Z er F -fordelt $(f_T, f_N) = (f_C, f_m)$.

Hvis H_0 forkastes, så opstilles konfidensintervaller til vurdering af faktorerens virkning.

Opstilling af konfidensintervaller og drage konklusion.

Lad $\bar{x}_{ij}$ være gennemsnittet af værdierne i cellen i i 'te række og j 'te søjle.

Lad $\bar{x}_{i.}$ være gennemsnittet af værdierne i den i 'te række.

Lad $\bar{x}_{.j}$ være gennemsnittet af værdierne i den j 'te søjle.

1) $R * C \neq 0$ ¹.

Konfidensintervaller for hver celle: $\left[\bar{x}_{ij} - t_{1-\frac{\alpha}{2}}(r \cdot q \cdot (n-1)) \frac{s_0}{\sqrt{n}}; \bar{x}_{ij} + t_{1-\frac{\alpha}{2}}(r \cdot q \cdot (n-1)) \frac{s_0}{\sqrt{n}} \right],$

2) $R * C = 0$ ² $R \neq 0 \wedge C \neq 0$:

For celle i i 'te række og j 'te søjle er den estimerede middelværdi

$$\tilde{\mu}_{ij} = \bar{x}_{i.} + \bar{x}_{.j} - \bar{x}_{..} \text{ (jævnfør betragtningerne i afsnit 11.3.2.2 side 59.)}$$

Konfidensintervaller for hver celle:

$$\left[\tilde{\mu}_{ij} - t_{1-\frac{\alpha}{2}}(N-r-q+1) \sqrt{\frac{(r+q-1)}{N}} s_m; \tilde{\mu}_{ij} + t_{1-\frac{\alpha}{2}}(N-r-q+1) \sqrt{\frac{(r+q-1)}{N}} s_m \right]$$

Det giver et bedre overblik, hvis man udregner de marginale konfidensintervaller:

$$\text{Konfidensintervaller for hver række: } \left[\bar{x}_{i.} - t_{1-\frac{\alpha}{2}}(N-r-q+1) \frac{s_m}{\sqrt{n \cdot q}}; \bar{x}_{i.} + t_{1-\frac{\alpha}{2}}(N-r-q+1) \frac{s_m}{\sqrt{n \cdot q}} \right]$$

$$\text{Konfidensintervaller for hver søjle: } \left[\bar{x}_{.j} - t_{1-\frac{\alpha}{2}}(N-r-q+1) \frac{s_m}{\sqrt{n \cdot r}}; \bar{x}_{.j} + t_{1-\frac{\alpha}{2}}(N-r-q+1) \frac{s_m}{\sqrt{n \cdot r}} \right]$$

3) $R * C = 0$, $R \neq 0 \wedge C = 0$:

For hver række i beregnes et rækkegennemsnit $\bar{x}_{i.}$

$$s_{m1}^2 = \frac{SAK_e + SAK_{RC} + SAK_C}{f_0 + f_{RC} + f_C}, \quad f_{m1} = f_e + f_{RC} + f_C = N - r$$

$$\text{Konfidensintervaller for hver række: } \left[\bar{x}_{i.} - t_{1-\frac{\alpha}{2}}(N-r) \frac{s_{m1}}{\sqrt{n \cdot q}}; \bar{x}_{i.} + t_{1-\frac{\alpha}{2}}(N-r) \frac{s_{m1}}{\sqrt{n \cdot q}} \right]$$

4) $R * C = 0$, $R = 0 \wedge C \neq 0$:

For hver søjle j beregnes et søjlegennemsnit $\bar{x}_{.j}$

$$s_{m2}^2 = \frac{SAK_e + SAK_{RC} + SAK_R}{f_0 + f_{RC} + f_R}, \quad f_{m2} = f_e + f_{RC} + f_R = N - q$$

$$\text{Konfidensintervaller for hver søjle: } \left[\bar{x}_{.j} - t_{1-\frac{\alpha}{2}}(N-q) \frac{s_{m2}}{\sqrt{n \cdot r}}; \bar{x}_{.j} + t_{1-\frac{\alpha}{2}}(N-q) \frac{s_{m2}}{\sqrt{n \cdot r}} \right]$$

¹Kort skrivemåde for, at $H_0: R * C = 0$ forkastes.

²Kort skrivemåde for, at $H_0: R * C = 0$ accepteres

Appendix 11A Eksempler regnet på TI - 89.

1. Indledning.

Det forudsættes, at man kender de grundlæggende operationer på lommeregneren.

I Appendix "Vejledning i TI - 89" er beskrevet, hvorledes man beregner sandsynligheden for forskellige fordelinger, beregner gennemsnit og spredning, samt hvorledes man tester og beregner konfidensintervaller for funktion af 1 variabel. Dette forudsættes ligeledes bekendt.

2. Variansanalyse

2.1 Ensidede variansanalyse

Eksempel 11.2. For hvert af 4 tilsætningsstoffer T_1, T_2, T_3, T_4 måles 3 gange mængden af urenheder ved en kemisk proces. Forsøgsresultaterne blev følgende:

T_1	T_2	T_3	T_4
108	105	108	117
110	110	111	119
112	109	113	112

Der ønskes fundet det tilsætningsstof der giver den mindste urenhed.

Løsning:

APPS, STAT/LIST hvorefter data indtastes i list1, list 2, list3 og list4.

F6, C:ANOVA, Antal grupper = 4, ENTER, Udfyld listnavne (VAR-Link osv.) ENTER, ENTER.

Der fremkommer nu en række resultater, der stort set svarer til resultaterne i en sædvanlig varianstabel. En sådan er udfyldt nedenfor:

Variation	SS	df	MS = s^2	F	P-Value
Factor: Tilsætningsstof	105.0	3	35.0	4.61538	0.03712
Error: Gentagelser	60.6667	8	7.58333		

Man finder en P -værdi på $=0.03712$.

Da P -værdi $= 0.0372 < 0.05$ forkastes nulhypotesen $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ (svagt)

Konklusion: De fire tilsætningsstoffer har ikke samme virkning.

2) De sædvanlige konfidensintervaller findes som ekstra søjler efter list6

xbar	lowlist	uplist
110	106.33	113.67
108	104.33	111.67
110.67	107.0	114.33
116	112.33	119.67

Ønskes beregnet LSD-intervaller må man udnytte, at $r_{kon} = \frac{113.67 - 103.33}{2}$ og

$$r_{lsd} = \frac{r_{kon}}{\sqrt{2}} = \frac{113.67 - 106.33}{2\sqrt{2}} = 2.592$$

	T_1	T_2	T_3	T_4
Gennemsnit	110.0	108.0	110.67	116.0
95% konfidensinterval	[106.33;113.67]	[104.33;111.67]	[107.00;114.33]	[112.33;119.67]
95% LSD-konfidensinterval	[107.40;112.60]	[105.40;110.60]	[108.07;113.26]	[113.40;118.60]

Konklusion: Begge metoder viser, at man får den mindste urenhed, hvis man vælger enten T_1 , T_2 eller T_3 (de kan ikke adskilles).

2.2 Tosidet variansanalyse

Eksempel 11.4. Model med vekselvirkning.

Man undersøger, hvorledes 3 olieblandinger O_1 , O_2 , O_3 , og 2 karburatortyper K_1 og K_2 påvirker benzinforbruget. Man fandt:

		Karburator			
		K_1		K_2	
Olieblanding	O_1	830	860	810	840
	O_2	940	990	1050	1020
	O_3	855	815	930	910

Angiv hvilke kombinationer af karburator og olieblanding der giver det laveste forbrug, og giv et estimat for dette forbrug.

Løsning:

APPS, STAT/LIST hvorefter data indtastes med første søjle (K_1 søjlen) i list1, 2 søjle (K_2) i list 2.

List 1	List 2
830	810
860	840
940	1050
990	1020
855	930
815	910

F6, ANOVA2-Way, ENTER

DESIGN=2 Factor,EqReps, Levl of Col Factor =2,Levl of Row Factor =3, ENTER

Næste skema udfyldes med List1 og List 2, ENTER

Resultatet kan umiddelbart aflæses:

Nedenfor er resultaterne angivet i den sædvanlige variansanalysetabel

Variation	SAK=SS	df	MS= s^2	F	P-værdi
Column:Søjlefaktor	6075	1	6075	10.125	0.019028
Row: Rækkefaktor	58716.7	2	29358.3	48.93	0.000193
Interaction:Vekselvirkning	6450	2	3225	5.375	0.044963
Error:Residual	3600	6	600		

- 1) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning)

For "interaction" findes P -værdi = 0.04596.

Idet vi som sædvanlig antager at signifikansniveauet er 5 % fås, at da P -værdi = 0.0460 < 0.05 forkastes H_0 (svagt).

Konklusion: Begge faktorer har en virkning i form af en vekselvirkning.

- 2) Beregning af 95% konfidensintervaller må ske ved anvendelse af formlerne.

Ifølge oversigt 11.2 er konfidensintervaller for hver celle:

$[\bar{x}_{ij} - r_{kon}; \bar{x}_{ij} + r_{kon}]$, hvor radius er

$$r_{kon} = t_{1-\frac{\alpha}{2}}(r \cdot q \cdot (n-1)) \frac{s_{\text{residual}}}{\sqrt{n}} = t_{0.975}(3 \cdot 2 \cdot (2-1)) \frac{\sqrt{600}}{\sqrt{2}} = 2.45 \cdot \sqrt{300} = 42.44$$

Vi finder gennemsnittene i hver celle

Gennemsnit		Karburator	
		K_1	K_2
Olieblanding	O_1	845	825
	O_2	965	1035
	O_3	835	920

95% konfidensinterval		Karburator	
		K_1	K_2
Olieblanding	O_1	[802.6 ; 887.3]	[782.6; 867.4]
	O_2	[922.6 ; 1007.4]	[992.6 ; 1077.4]
	O_3	[792.6 ; 877.4]	[877.6 ; 962.4]

Det ses, at umiddelbart giver K_2 O_1 det laveste benzinforbrug, men af konfidensintervallerne ses, at der ingen signifikant forskel er mellem K_2 O_1 , K_1 O_3 og K_1 O_1 .

Eksempel 11.5 Additiv model: Model uden vekselvirkning

Med hver af 4 cementblandere udstøbtes efter blanding med vand 12 cementterninger, som efter en uges lagring underkastedes en brudstyrkeprøve ved hjælp af en af 3 cementknusere.

Forsøgsresultaterne var:

		Cementknusere											
		1			2			3			4		
Cementblandere	1	147	175	130	99	85	75	67	23	35	215	97	180
	2	211	145	163	131	100	145	75	45	71	151	157	167
	3	123	85	153	137	143	82	67	25	83	135	91	129

- 1) Angiv hvilke kombinationer af cementblander og cementknuser, der giver den største brudstyrke, og giv et estimat og et 95% konfidensinterval for denne største middelbrudstyrke.
- 2) Da cementknuser 2 og cementblander 1 er de billigste ønskes opstillet under hensyntagen til resultatet af 1) et 95% - konfidensinterval for middelværdien af brudstyrken, når disse anvendes.

Løsning:

APPS, STAT/LIST hvorefter data indtastes med første søjle (cementknuser 1) i list1, 2 søjle (cementknuser 2) i list 2 osv.

F6, ANOVA2-Way, ENTER

DESIGN=2 Factor, EqReps, Levels of Col Factor =4, Levels of Row Factor =3, ENTER

Næste skema udfyldes med List1, List 2, List3 og List 4, ENTER

Resultatet kan umiddelbart aflæses:

Nedenfor er resultaterne angivet i den sædvanlige variansanalysetabel

Variation	SAK=SS	df	MS= s^2	F	P - værdi
Column:Cementknuser : C	51995.2	3	17331.7	19.5557	0.000001
Row: Cementblander : R	4299.39	2	2149.69	2.42553	0.109798
Interaction: R*C	7122.61	6	1187.1	1.33942	0.278669
Error	21270.7	24	886.278		

- 1) a) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning)

For “interaction” findes P - værdi = 0.2787.

Da P - værdi = 0.2787 > 0.05 accepteres H_0 .

Konklusion: Vi antager i det følgende, at vekselvirkningen er forsvindende.

- b) Inden vi tager stilling til om faktorerne har en hovedvirkning, “pooles” interaction med “error”.

Det kan TI-89 ikke gøre (mærkeligt), så det gøres manuelt.

$$SAK_{\text{residual}} = SAK_{\text{error}} + SAK_{\text{vekselvirkning}} = 28393.3. f_{\text{residual}} = f_{\text{error}} + f_{\text{vekselvirkning}} = 30$$

$$s_{\text{residual}}^2 = \frac{28393.3}{30} = 946.443$$

$$F_{\text{knuser}} = \frac{17331.7}{946.44} = 18.31. P - værdi = P(F > 18.31) = FCdf(18.31, \infty, 3, 30) = 6.09 \cdot 10^{-7}$$

$$F_{blander} = \frac{2149.69}{946.44} = 2.271 \quad P\text{-værdi} = P(F > 2.271) = \text{FCdf}(2.271, \infty, 2, 30) = 0.121$$

Det giver følgende tabel:

Variation	SAK=SS	df	MS= s^2	F	P - værdi
Column:Cementknuser : C	51995.2	3	17331.7	18.31	0.000000
Row: Cementblander : R	4299.39	2	2149.69	2.271	0.121
Residual	28393.3	30	946.443		

H_0 : C = 0 (Cementknuser har ingen virkning) forkastes, da $P\text{-værdi} = 0.000 < 0.05$

H_0 : R = 0 (Cementblander har ingen virkning) accepteres, da $P\text{-værdi} = 0.121 > 0.05$

Konklusion: Cementknuserne har en (stærk) virkning

Cementblandere har ingen virkning

For at finde hvilken cementknuser der giver den største middelbrudstyrke kunne vi nu beregnes gennemsnit og konfidensinterval for de 4 knusere ved at benytte de i oversigt 11.2 angivne formler.

Lettere er det at udnytte, at da vi nu kun har en faktor tilbage, så udnytte programmet for ensidet variansanalyse, med cementknusere som faktor på 4 niveauer..

Konfidensintervaller findes som ekstra søjler efter list6

xbar	lowlist	uplist
148	126.3	169.7
110.78	89.08	132.48
54.56	32.85	76.26
146.89	125.19	168.59

Konklusion: Cementknuser 1 og 4 må foretrækkes, selv om man ikke helt kan afvise at cementknuser 2 kan være lige så god.

Et estimat for største middelbrudstyrke: 148 95% konfidensinterval [126.3 ; 169.7]

2) Cementknuser 2: Middelbrudstyrke: 110.78 95% konfidensinterval [89,08 ; 132.5]

Eksempel 11.7. Blokforsøg:

I nedenstående tabel er anført resultaterne af et fodringsforsøg med svin. Da grise fra samme kuld må forventes at være mere ensartede, vælger man at lave et randomiseret blokforsøg med kuld som blokfaktor.

Fra hvert af 4 forskellige kuld grise udtages 3 grise, der bliver fodret med hver sin af tre fodertyper A, B og C med forskelligt vitaminindhold.

Forsøgsresultaterne (vægtforøgelse i kg) var

		Fodertype		
		A	B	C
Kuld	1	7.0	14.0	8.5
	2	16.0	15.5	16.5
	3	10.5	15.0	9.5
	4	13.5	21.0	13.5

Test, om der er nogen væsentlig virkning af ændringen i foderets vitaminindhold.

Løsning:

APPS, STAT/LIST hvorefter data indtastes med første søjle (A) i list1, 2 søjle (B) i list 2 osv.

List 1	List 2	LIST 3
7	14	8.5
16	15.5	16.5
10.5	15	9.5
13.5	21	13.5

F6, ANOVA2-Way, ENTER

DESIGN=Block, Levels of Col Factor =3, ENTER

Næste skema udfyldes med List1 , List 2 og LIST3, ENTER

Resultatet kan umiddelbart aflæses: Nedenfor er resultaterne angivet i den sædvanlige variansanalysetabel

Variation	SAK=SS	df	MS= s^2	F	P-værdi
Factor: foder	54.125	2	27.0625	5.756	0.0402
Block: Kuld	87.7729	3	29.2431	6.22	0.0284
Error	28.2083	6	4.70139		

H_0 : Foder = 0 (Foder har ingen virkning) forkastes, da P-værdi = 0.0402 < 0.05

Konklusion: Der sker en væsentlig ændring i vægtforøgelsen ved at ændre foderblanding. Konfidensintervaller kunne vise hvilken foderblanding der skal foretrækkes.



Appendix 11B Eksempler regnet på Statgraphics.

1. Indledning

I "Grundlæggende begreber Appendix A" er beskrevet hvorledes man beregner sandsynligheden for forskellige fordelinger og beregner gennemsnit og spredning. Dette forudsættes bekendt.

2 Variansanalyse

2.1. Ensided variansanalyse.

Eksempel 11.2. For hvert af 4 tilsætningsstoffer T_1, T_2, T_3, T_4 måles 3 gange mængden af urenheder ved en kemisk proces. Forsøgsresultaterne blev følgende:

T_1	T_2	T_3	T_4
108	105	108	117
110	110	111	119
112	109	113	112

Der ønskes fundet det tilsætningsstof der giver den mindste urenhed.

Løsning:

Data indtastes

Lad starten af indtastningen i regnearket være

Stof	Urenhed
T1	108
T1	110
T1	112
T2	105
osv.	

Vælg (Compare\ Analysis of Variance\ One-Way Anova\Klik på "Urenhed"\klik på pilen ved "Dependent variable"\Klik på "stof"\klik på pilen "Factor"\OK)

Der fremkommer en "Analysis Summary" med nogle statistiske betragtninger.

Endvidere fremkommer et "scatterplot". Scatterplottet kan man om ønsket fjerne ved med cursoren på tabellen hurtigt 2 gange at trykke på venstre musetast.

Vælg (gul ikon = Tabular Options\Anova Tables\ OK) (fjern evt. krydset ved "Analysis summary").

Der fremkommer følgende udskrift

ANOVA Table for udbytte by tilsætningsstof
Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Between groups	105,0	3	35,0	4,62	0,0372
Within groups	60,6667	8	7,58333		
Total (Corr.)	165,667	11			

Forklaring på de enkelte størrelser kan man finde i oversigt 11.1.

Da P -værdi = $0.0372 < 0.05$ forkastes nulhypotesen $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ (svagt)

Konklusion: De fire tilsætningsstoffer har ikke samme virkning.

Konfidensintervaller.

Vælg(Tabular options \Tables of Means \ OK),

Den fremkomne tabel er over 95% "LSD" konfidensintervaller.

Ønskes de sædvanlige 95% konfidensintervaller så

Vælg (Med cursor på udskrift over LSD-intervaller, tryk på højre musetast\Pane options\Confidence intervals (pooled s)\OK).

De "individuelle" 95% konfidensintervaller er

Table of Means for udbytte by tilsaetningsstof with 95,0 percent confidence intervals

tilsaetningsstof	Count	Mean	Stnd. error (pooled s)	Lower limit	Upper limit
T1	3	110,0	1,5899	106,334	113,666
T2	3	108,0	1,5899	104,334	111,666
T3	3	110,667	1,5899	107,0	114,333
T4	3	116,0	1,5899	112,334	119,666
Total	12	111,167			

Man får følgende LSD-intervaller.

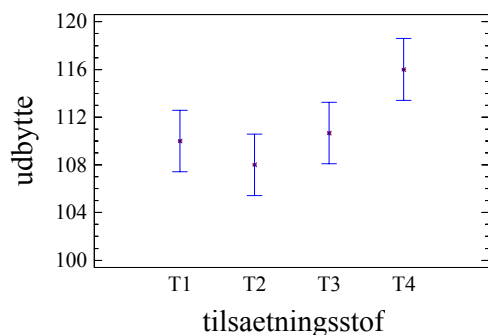
Table of Means for udbytte by tilsaetningsstof with 95,0 percent LSD intervals

tilsaetningsstof	Count	Mean	Stnd. error (pooled s)	Lower limit	Upper limit
T1	3	110,0	1,5899	107,408	112,592
T2	3	108,0	1,5899	105,408	110,592
T3	3	110,667	1,5899	108,074	113,259
T4	3	116,0	1,5899	113,408	118,592
Total	12	111,167			

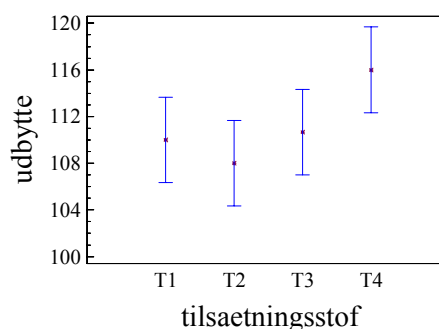
Plot af konfidensintervallerne:

Vælg (blå ikon = Graphics Options\ Means Plot\Cursor på figur\ højre musetast\ Pane options\vælg type " Confidence intervals"OK)

Means and 95,0 Percent LSD Intervals



Means and 95,0 Percent Confidence Intervals (pooled)

Heraf ses, at T_4 er signifikant større end T_2 , mens T_1 , T_2 og T_3 ikke kan adskilles.Konklusion: Man skal ikke vælge tilsaetningsstof T_4 , hvorimod de tre øvrige tilsaetningsstoffer giver nogenlunde den samme mængde urenhed.**Kontrol af forudsætninger:**

Det er let i Statgraphics at kontrollere forudsætningerne men derfor behøver man ikke nødvendigvis at gøre det.

Varianshomogenitet:

Vælg (gul ikon=Tabular options \ Variance check \ OK).

Variance Check

Cochran's C test: 0,428571 P-Value = 0,746356

Bartlett's test: 1,0944 P-Value = 0,897063

Hartley's test: 3,25

Da P -værdi = 0.897 (eller 0.0746) $>$ 0.05 accepteres nulhypotesen H_0 : De 4 varianser er ens.

Normalfordelingsplot:

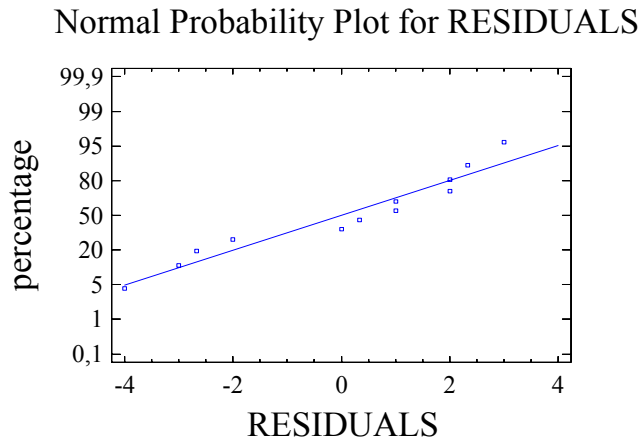
Vælg (sort ikon = Save Results\Save Residuals\ OK).

Residualerne bliver nu gemt som en søjle i data under navnet "RESIDUALS", da vi ikke har ændret navnet under "Target Variables"

Vælg (Describe\Numerical Data\One Variable Analysis\RESIDUALS\Pilen Data\ OK)

Vælg (blå ikon = Graphics options\Normal Probability Plot\ OK).

Vælg (Med cursor på tegning, højre musetast\Pane Options \Using Least Squares\OK)



Residualerne synes at ligge nogenlunde på en ret linie , så man må antage at de er approksimativt normalfordelte.

2.2 Tosidet variansanalyse.

Data indtastes på sædvanlig måde .

Karburator	oliebland	benzinforb
k1	o1	830
k1	o1	860
k1	o2	940
k1	o2	990
k1	o3	855
k1	o3	815
k2	o1	810

OSV.

Vælg (Compare\ Analysis of Variance\ Multifactor ANOVA\ OK).

Der fremkommer nu en tabel der skal udfyldes:

Vælg ("brudstyrke"\pilen "Dependent variable"\ "cementblandere"\pilen "Factors"\cementknusere\pilen "Factors\OK). Der fremkommer en "Analysis Summary " og et "Scatterplot".

Vælg (gul ikon = Tabular options\ Anova Tables\OK).

Der fremkommer en variansanalysetabel uden vekselvirkning.

For at få en variansanalysetabel med vekselvirkning:

Vælg (Cursor i tabellen\ højre musetast\ Analysis options\ Maximum Order Interaction til 2\ OK)

Vi får følgende udskrift:

Appendix 11B Eksempler regnet på Statgraphics

Analysis of Variance for benzinform - Type III Sums of Squares

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
MAIN EFFECTS					
A:karburator	6075,0	1	6075,0	10,12	0,0190
B:oliebland	58716,7	2	29358,3	48,93	0,0002
INTERACTIONS					
AB	6450,0	2	3225,0	5,37	0,0460
RESIDUAL	3600,0	6	600,0		
TOTAL (CORRECTED)	74841,7	11			

All F-ratios are based on the residual mean square error.

1) $H_0: A * B = 0$ (Ingen signifikant vekselvirkning)

For "interaction" findes P - værdi = 0.0460.

Da P - værdi = 0.0460 < 0.05 forkastes H_0 (svagt).

Konklusion: Begge faktorer har en virkning i form af en vekselvirkning.

2) Konfidensintervaller

Vælg (gul ikon = Tabular options\ Tables of Means\ OK).

Vi får følgende udskrift :

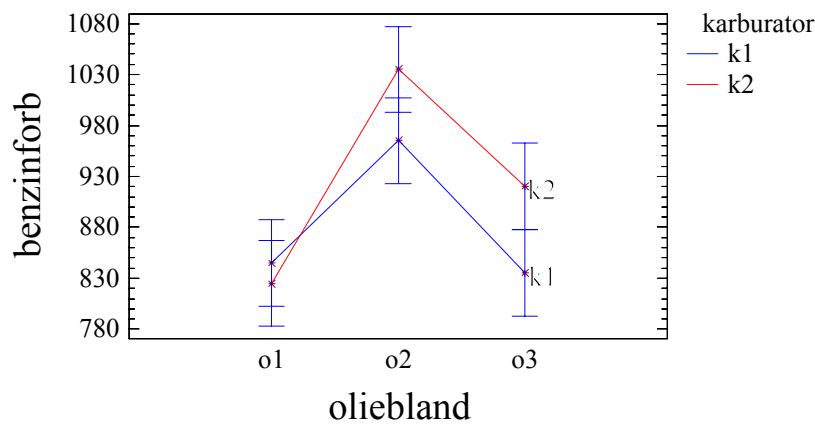
Table of Least Squares Means for benzinform with 95,0 Percent Confidence Intervals

Level	Count	Mean	Std. Error	Lower Limit	Upper Limit
GRAND MEAN					
karburator	12	904,167			
k1	6	881,667	10,0	857,197	906,136
k2	6	926,667	10,0	902,197	951,136
oliebland					
o1	4	835,0	12,2474	805,031	864,969
o2	4	1000,0	12,2474	970,031	1029,97
o3	4	877,5	12,2474	847,531	907,469
karburator by oliebland					
k1 o1	2	845,0	17,3205	802,618	887,382
k1 o2	2	965,0	17,3205	922,618	1007,38
k1 o3	2	835,0	17,3205	792,618	877,382
k2 o1	2	825,0	17,3205	782,618	867,382
k2 o2	2	1035,0	17,3205	992,618	1077,38
k2 o3	2	920,0	17,3205	877,618	962,382

Mere overskueligt er det, at få tegnet et "interaction-plot" med afsatte konfidensintervaller: Sædvanligvis fås det mest overskuelige billede, hvis man afsætter den faktor med flest niveauer (altså olieblending) ud af den vandrette akse.

Vælg (blå ikon = Graphics options | Interaction Plot | Cursoren på figur, højre musetast | Pane options | Confidence intervals | "Second Factor" | OK).

Interactions and 95,0 Percent Confidence Intervals



Konklusion: Vi ser af tabel og figur, at man ikke bør vælge olieblending O_2 .

Umiddelbart giver kombinationen $K_2 O_1$ det laveste benzineforbrug (825), men af konfidensintervallerne ses, at der ingen signifikant forskel er mellem $K_2 O_1$, $K_1 O_3$ og $K_1 O_1$.



Eksempel 11.5 Additiv model: Model uden vekselvirkning

Med hver af 4 cementblandere udstøbtes efter blanding med vand 12 cementterninger, som efter en uges lagring underkastedes en brudstyrkeprøve ved hjælp af en af 3 cementknusere.

Forsøgsresultaterne var:

		Cementknusere											
		1			2			3			4		
Cementblandere	1	147	175	130	99	85	75	67	23	35	215	97	180
	2	211	145	163	131	100	145	75	45	71	151	157	167
	3	123	85	153	137	143	82	67	25	83	135	91	129

- Angiv hvilke kombinationer af cementblander og cementknuser, der giver den største brudstyrke, og giv et estimat og et 95% konfidensinterval for denne største middelbrudstyrke.
- Da cementknuser 2 og cementblander 1 er de billigste ønskes opstillet under hensyntagen til resultatet af 1) et 95% - konfidensinterval for middelværdien af brudstyrken, når disse anvendes.

Løsning:

Lad starten af indtastningen i regnearket være

cementblandere	cementknusere	brudstyrke
1	1	147
1	1	175
1	1	130
1	2	99
1	2	85
osv.		

1) Variansanalysetabel opstilles.

Vælg (Compare\ Analysis of Variance\ Multifactor ANOVA\ OK).

Der fremkommer nu en tabel der skal udfyldes:

Vælg ("brudstyrke"\pilen "Dependent variable"\ "cementblandere"\pilen "Factors"\cementknusere\pilen "Factors\OK)

Der fremkommer en "Analysis Summary " og et "Scatterplot".

Vælg (gul ikon = Tabular options\ Anova Tables\OK).

Der fremkommer en variansanalysetabel uden vekselvirkning.

For at få en variansanalysetabel med vekselvirkning:

Vælg (Cursor i tabellen\ højre musetast\ Analysis options\ Maximum Order Interaction til 2\ OK)

Analysis of Variance for brudstyrke - Type III Sums of Squares

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
MAIN EFFECTS					
A:cementblandere	4299,39	2	2149,69	2,43	0,1098
B:cementknusere	51995,2	3	17331,7	19,56	0,0000
INTERACTIONS					
AB	7122,61	6	1187,1	1,34	0,2787
RESIDUAL	21270,7	24	886,278		
TOTAL (CORRECTED)					
	84687,9	35			

All F-ratios are based on the residual mean square error.

a) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning)

For "interaction" findes P - værdi = 0.2787.

Da P - værdi = 0.2787 > 0.05 accepteres H_0 .

Konklusion: Vi antager i det følgende, at vekselvirkningen er forsvindende.

b) AB "pooles" ned i Residualen".

Vælg (Cursor i tabellen\ højre musetast\ Analysis options\ Maximum Order Interaction til 1\ OK)

Vi får følgende tabel:

Analysis of Variance for brudstyrke - Type III Sums of Squares

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
MAIN EFFECTS					
A:cementblandere	4299,39	2	2149,69	2,27	0,1206
B:cementknusere	51995,2	3	17331,7	18,31	0,0000
RESIDUAL	28393,3	30	946,443		
TOTAL (CORRECTED)					
	84687,9	35			

$H_0: C = 0$ (Cementknuser har ingen virkning) forkastes, da P -værdi = 0.000 < 0.05

$H_0: R = 0$ (Cementblander har ingen virkning) accepteres, da P -værdi = 0.121 > 0.05

Konklusion: Cementknuserne har en stærk signifikant virkning.
Cementblanderne ikke har en signifikant virkning.

Da cementblandere ikke mere indgår i modellen slettes de af modellen.

Vælg (rød ikon = Input dialog\ Slet cementblandere\ OK).

Dette har den lidt kedelige virkning, at der nu sker en "omdøbning", så det der før blev kaldt A nu kaldes B osv.

Analysis of Variance for brudstyrke - Type III Sums of Squares

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
MAIN EFFECTS					
A:cementknusere	51995,2	3	17331,7	16,96	0,0000
RESIDUAL	32692,7	32	1021,65		
TOTAL (CORRECTED)	84687,9	35			

Vælg(Tabular options \Tables of Means \ OK),

Der fremkommer følgende tabel over 95% konfidensintervaller.

Table of Least Squares Means for brudstyrke with 95 Percent Confidence Intervals

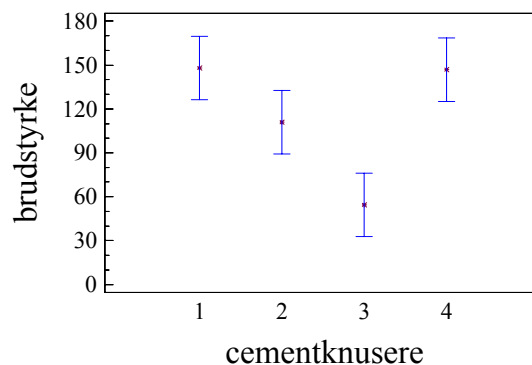
Level	Count	Mean	Std. Error	Lower Limit	Upper Limit
GRAND MEAN	36	115,056			
cementknusere					
1	9	148,0	10,6544	126,298	169,702
2	9	110,778	10,6544	89,0754	132,48
3	9	54,5556	10,6544	32,8532	76,2579
4	9	146,889	10,6544	125,187	168,591

Grafisk fremstilling:

LSD-intervaller: Vælg (blå ikon = Graphics Options\ Means Plot OK)

Sædvanlige konfidensintervaller: Vælg (Med Cursor på figur\ højre musetast\ Pane options\ Confidence intervals og vælg faktor, der ønskes undersøgt\OK)

Means and 95,0 Percent Confidence Intervals



Konklusion: Cementknuser 1 og 4 må foretrækkes, selv om man ikke helt kan afvise at cementknuser 2 kan være lige så god.

Et estimat for største middelbrudstyrke: 148 95% konfidensinterval [126.3 ; 169.7]

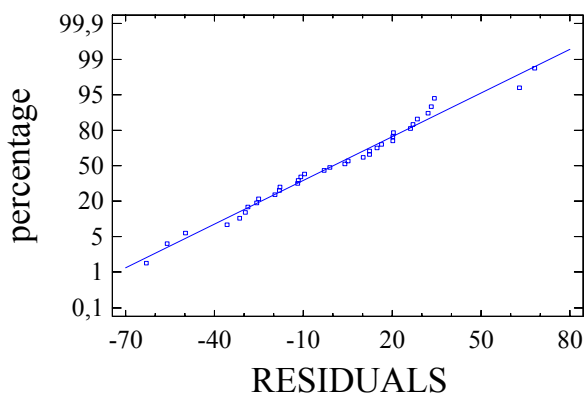
2) Cementknuser 2: Middelbrudstyrke: 110.78 95% konfidensinterval [89,08 ; 132.5]

3) **Forudsætninger.** Det er let i Statgraphics at kontrollere forudsætningerne men derfor behøver man ikke nødvendigvis at gøre det, da de sædvanligvis er opfyldt.

Grafisk kontrol af kravet om normalitet

Udføres med et normalfordelingsplot.(se under ensidet variansanalyse)

Normal Probability Plot for RESIDUALS



Residualerne synes at fordele sig nogenlunde på en ret linie, så kravet om normalitet synes rimeligt opfyldt.

Varianshomogenitet.

Man går ind i regnearket og danne en ekstra søjle "behandlinger". Da der er 12 behandlinger (celler) med 3 tal i hver bliver søjlen:

cementblandere	cementknusere	brudstyrke	behandlinger
1	1	147	1
1	1	175	1
1	1	130	1
1	2	99	2
1	2	85	2
osv.			

Vælg (Compare\ Analysis of Variance\ One Way ANOVA\OK).

Vælg (Klik på "brudstyrke" og pil ved "Dependent variable"Klik på "behandlinger" og på pil ved "Factor"\OK)

Vælg (gul ikon = Tabular options\ Variance check\ OK). Fjern evt. krydset ved "Analysis summary

Variance Check

Cochran's C test: 0,345358 P-Value = 0,113559

Bartlett's test: 1,57303 P-Value = 0,602582

Hartley's test: 56,2194

Det ses, at vi får en accept af en nulhypotese om at varianserne er ens, så kravet er rimeligt opfyldt. ◆

Eksempel 11.7. Blokforsøg:

I nedenstående tabel er anført resultaterne af et fodringsforsøg med svin. Da grise fra samme kuld må forventes at være mere ensartede, vælger man at lave et randomiseret blokforsøg med kuld som blokfaktor.

Fra hvert af 4 forskellige kuld grise udtages 3 grise, der bliver fodret med hver sin af tre fodertyper A, B og C med forskelligt vitaminindhold.

Forsøgsresultaterne (vægtforøgelse i kg) var

		Fodertype		
		A	B	C
Kuld	1	7.0	14.0	8.5
	2	16.0	15.5	16.5
	3	10.5	15.0	9.5
	4	13.5	21.0	13.5

Test, om der er nogen væsentlig virkning af ændringen i foderets vitaminindhold.

Løsning.Indtastning af data

fodertype	kuld	vægtforøgelse
A	1	7
A	2	16
A	3	10,5
A	4	13,5
B	1	14
B	2	15,5
B	3	15
B	4	21
C	1	8,5
C	2	16,5
C	3	9,5
C	4	13,5

Vælg (Compare\ Analysis of Variance\ Multifactor ANOVA\ OK).

Der fremkommer nu en tabel der skal udfyldes:

Vælg ("vægtforøgelse" \pilen "Dependent variable" \ "fodertype" \ pilen "Factors" \ kuld \ pilen "Factors" \ OK)

Der fremkommer en "Analysis Summary " og et "Scatterplot".

Vælg (gul ikon = Tabular options\ Anova Tables\ OK).

Der fremkommer en variansanalysetabel uden vekselvirkning.

Analysis of Variance for vægtforøgelse - Type III Sums of Squares

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value

MAIN EFFECTS					
A:fodertype	54,125	2	27,0625	5,76	0,0402
B:kuld	87,7292	3	29,2431	6,22	0,0285

RESIDUAL	28,2083	6	4,70139		

TOTAL (CORRECTED)	170,063	11			

All F-ratios are based on the residual mean square error.

Vi ser , at der på et signifikansniveau på 5 % er signifikant forskel på fodertyperne (mindst én afviger fra de øvrige).

Vi ser endvidere, at det var fornuftigt at dele op i kuld, da der også er signifikans for kuld.

Vi er imidlertid ikke interesseret i at finde ud af hvilket kuld der er det bedste, da vi jo blot har taget nogle tilfældige kuld ud.

For at finde den fodertype, der giver den højeste vægtforøgelse opstilles konfidensintervaller. Bemærk: Da der er en blokvirkning vil vægtforøgelsen jo afhænge af hvilket kuld man betragter. Konfidensintervaller kan derfor kun anvendes til relative sammenligninger.

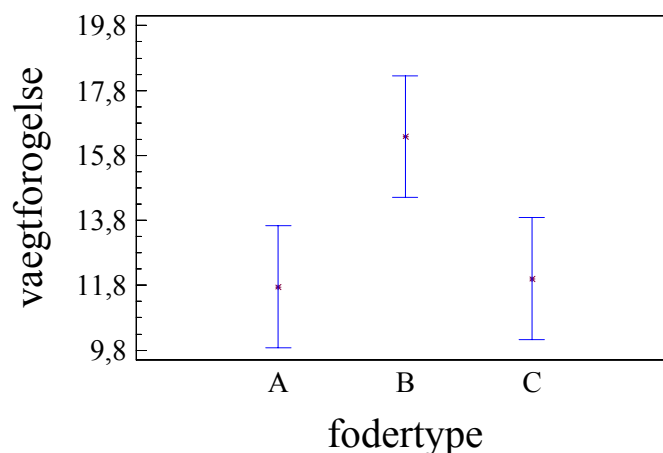
Vælg (gul ikon = Tabular options\ Tables of Means\ OK).

Table of Least Squares Means for vægtforøgelse with 95,0 Percent Confidence Intervals

Level	Count	Mean	Std. Error	Lower Limit	Upper Limit
GRAND MEAN	12	13,375			
fodertype					
A	4	11,75	1,08413	9,09721	14,4028
B	4	16,375	1,08413	13,7222	19,0278
C	4	12,0	1,08413	9,34721	14,6528
kuld					
1	3	9,83333	1,25185	6,77016	12,8965
2	3	16,0	1,25185	12,9368	19,0632
3	3	11,6667	1,25185	8,60349	14,7298
4	3	16,0	1,25185	12,9368	19,0632

En tegning af 95% LSD - konfidensintervallerne :

Means and 95,0 Percent LSD Intervals



Konfidensintervallerne viser ganske vist et svagt overlap, men det gør LSD-intervallerne ikke, og da variansanalysen har vist at der er en signifikant forskel, må der gælde, at fodertype B giver den største vægtforøgelse.



OPGAVER

Opgave 11.1

Fire forskellige typer teknik til blanding af cement ønskes undersøgt med hensyn til resultatets trykstyrke. Følgende data blev opnået:

Blandingsteknik	Trykstyrke (psi)			
B1	3129	3000	2865	2890
B2	3200	3300	2975	3250
B3	2800	2900	2985	3050
B4	2600	2700	2600	2765

Undersøg om forskellen i blandingsteknik har betydning for trykstyrken, og angiv i bekræftende fald den (de) blandingsteknik(er) der har størst trykstyrke.

Opgave 11.2

I følgende tabel er angivet resultaterne af gentagne bestemmelser af blodets alkoholkoncentration (i promille) hos 6 forskellige personer efter indtagelsen af 4 cl. alkohol.

Person					
1	2	3	4	5	6
0.76	0.84	0.83	1.00	0.88	0.86
0.82	0.79	0.78	0.90	0.90	0.89
0.79	0.82	0.97	0.92	1.03	0.87
0.86	0.79	0.88	0.88	0.87	0.84

Vurdér på grundlag af dette materiale en antagelse om, at alkoholkoncentrationen i blodet ikke afhænger af andre faktorer end den indtagne alkoholemængde.

Opgave 11.3

Modstanden af 5 spoler målt for at kontrollere, om spolerne har samme elektriske modstand. For hver spole målt 4 uafhængige observationer:

Man fandt for hver spole følgende gennemsnit og varians:

Spole nr.	Antal gentagelser	Gennemsnit	Varians
1	4	15.2	0.035
2	4	14.95	0.015
3	4	14.8	0.0216667
4	4	15.15	0.0333333
5	4	14.7375	0.065625

- 1) Undersøg om det kan antages, at de 5 spolers modstande er ens.
- 2) På alle 5 spoler er angivet, at modstanden er 15.0 Ohm. Undersøg under hensyntagen til besvarelsen af spørgsmål 1) ved opstilling af et eller flere konfidensintervaller, om nogle af spolerne kan antages at have modstanden 15.0 Ohm og i bekræftende fald hvilke.

Opgave 11.4

I et forsøg undersøgte, om det kemiske udbytte af en proces afhænger af hvilken af 2 katalysatorer, der anvendes. Endvidere kan man benytte 3 forskellige apparater, og de kunne også tænkes at have indflydelse på resultatet.

Der fandtes følgende udbytter:

	Katalysator K_1	Katalysator K_2
Apparat A_1	69 72	65 66
Apparat A_2	72 71	70 69
Apparat A_3	70 71	72 73

- 1) Undersøg, om det kan antages, at udbyttet fra de 6 behandlinger har samme spredning.
- 2) Idet det antages at forudsætningerne for at udføre en variansanalyse er tilstede, skal der udføres en test til vurdering af, om middeludbyttets (eventuelle) afhængighed af de benyttede katalysatorer og apparater kan beskrives ved en additiv model.
- 3a) Hvis man af økonomiske grund vælger apparat A_1 hvilken katalysator skal man så vælge?
Samme spørgsmål vedrørende apparat 2 og 3.
- 3b) Hvis man af økonomiske grunde vælger katalysator 1 hvilket apparat skal man så vælge?
Samme spørgsmål vedrørende katalysator 2
- 3c) Hvilken (hvilke) kombinationer af apparat og katalysator giver det største udbytte.

Opgave 11.5^T

Man ønsker at undersøge den virkning som 2 faktorer (typen af glas og fosfor) har på skarphe den af billedet på en TV-skærm. Responsvariablen er den strøm (i microampere) som er nødvendig for at opnå et specifik skarpheds niveau.

Data er vist i nedenstående tabel:

		Fosfortype		
		1	2	3
Glastype	1	280	300	290
		290	310	285
		285	295	290
	2	230	260	220
		235	240	225
		240	235	230

Spørgsmål 1: Undersøg om forudsætningen om varianshomogenitet er opfyldt

Idet de sædvanlige variansanalyseforudsætninger antages opfyldt, ønskes følgende spørgsmål belyst:

Spørgsmål 2: Har glastype og fosfortype indflydelse på skarphe den?

Spørgsmål 3: Ud fra svaret i spørgsmål 1 skal angives, hvilken glastype og fosfortype der giver den største skarphed (giver den mindste respons)

Opgave 11.6

På en fabrik for glasvarer ønsker man at undersøge hvilken blandt 3 typer lim, der er bedst ved sammenlimning af 3 forskellige glastyper. Forsøget foregik ved, at man limede to glasplader sammen, og efter en passende tid undersøgte, hvor stor en kraft der skulle til for at trække pladerne fra hinanden. Man valgte at lave et fuldstændigt faktorforsøg med 5 gentagelser af hver behandling. Resultatet af forsøget var:

	Glastype A	Glastype B	Glastype C
LIM I	20 18 23 22 23	21 27 24 20 18	17 23 18 21 25
LIM II	30 25 28 27 28	28 24 16 25 21	28 25 29 27 28
LIM III	31 32 18 30 21	18 30 18 32 31	23 24 19 22 24

- 1) Angiv hvilke faktorer der har en virkning.
- 2) Angiv den eller de kombinationer af type lim og type glas, der har den største sammenhængskraft. Angiv et 95% konfidensinterval for de pågældende kombinationer.

Opgave 11.7

Fabrikationen af et kemikalium baseres på en bestemt kemisk proces, som forudsætter tilsætning af katalysator og en PH - værdi på ca. 5.

Som led i en laboratoriemæssig undersøgelse af mulighederne for at forøge procesudbyttet foretoges bl.a. et forsøg, hvor man dels sammenlignede virkningen af tilsætning af 3 forskellige katalysatorer, dels undersøgte, om udbyttet afhang af, om den nødvendige PH - værdi opnåedes ved tilsætning af HCl i stedet for som hidtil H_2SO_4 .

Forsøgsresultaterne var (udbytteprocenter):

		Tilsat syre	
		HCl	H_2SO_4
Katalysatorer	1	27.0 27.7	30.1 29.1
	2	25.5 27.0	30.6 28.9
	3	26.5 25.0	25.8 28.0

- 1) Foretag en statistisk analyse af forsøgsresultaterne og drag konklusioner.
- 2) Estimer under hensyn til resultatet af den under punkt 1) foretagne analyse procesudbyttet ved benyttelse af katalysator 2 under tilsætning af HCl og opstil et 95% - konfidensinterval for dette udbytte.

Opgave 11.8

Hver af tre laboranter har bestemt hydroquinons smeltepunkt ($^{\circ}$ Celcius) med (de samme) 4 termometre. Resultaterne var:

		Termometre			
		1	2	3	4
Laboranter	1	174.0	173.0	171.5	173.5
	2	173.0	172.0	171.0	171.0
	3	173.5	173.0	173.0	172.5

Det antages, at de nødvendige variansanalyseforudsætninger er opfyldt, og at termometre og laboranter ikke vekselvirker.

Følgende 2 spørgsmål ønskes belyst:

1. Aflæser laboranterne termometrene på samme måde?
2. Viser termometrene ens.
3. Vurder ved et residualplot, om forudsætningen om at termometre og laboranter ikke vekselvirker er rimelig.

Opgave 11.9

På en ingeniørskole ønsker man at sammenligne effektiviteten af undervisningen, når man underviser efter tre forskellige undervisningsmaterialer. En række studerende meldte sig frivilligt til forsøget. I det følgende er angivet 12 studerende ordnet efter studentereksamensgennemsnit.

Navn	JK	AL	TS	BS	DT	HN	MO	FD	PJ	KM	SR	RA
Snit	6.3	6.8	7.3	7.3	7.9	8.2	8.4	8.5	9.0	10.2	11.1	11.2

- 1) Hvordan ville du opdele disse studenter på tre hold med 4 på hver hold?
- 2) Hvordan ville du gøre det, hvis karaktererne gik fra 7.8 til 8.2 ?

Opgave 11.10.

Følgende resultater blev opnået fra et eksperiment, hvor man ville undersøge om der var forskel på de resultater, som 5 analyseapparater gav, når man analyserede kvælstofindholdet i jordprøver.

På hver af 3 dage blev en portion jord udvalgt og delt i 5 dele, som ved lodtrækning blev givet til analyse i hver sin maskine. Resultaterne var:

		Maskiner				
		P	Q	R	T	U
	Tirsdag	376	379	399	373	376
	Onsdag	372	374	409	387	386
	Torsdag	332	339	365	350	342

Undersøg på dette grundlag om der er forskel mellem analyseapparaterne, og angiv i bekræftende fald hvilke der er forskellige.

Mener du, at det i denne situation var en god ide at foretage forsøget som et blokforsøg?

Opgave 11.11.

Fire forskellige produktionsmetoder P, Q, R, og T ønskes sammenlignet med hensyn til det procentiske udbytte ved udvinding af et metal fra et bestemt mineral. Da man ved forsøget er nødt til at benytte forskellige råvarepartier, og er bange for, at det vil give stor spredning, vælger man at lave et fuldstændigt randomiseret blokforsøg med råvarepartier som blokke. Nedenstående skema angiver resultatet af dette forsøg.

	Metode P	Metode Q	Metode R	Metode T
Råvareparti 1	2.5 2.7	4.7 3.3	2.8 3.3	5.5 5.0
Råvareparti 2	4.6 4.3	7.9 5.9	5.1 6.9	7.2 6.8
Råvareparti 3	4.7 3.9	4.7 4.4	4.4 3.7	6.4 5.7

Undersøg på grundlag af disse oplysninger, om der er forskel på metoderne.

12 Flere end 2 binomial- og Poissonfordelte variable.

12.1. Indledning

De faktorer der forekommer i dette kapitel vil være binomial-eller Poissonfordelte. Vi vil dels analysere forsøg med 1 faktor der har mere end 2 niveauer, dels forsøg med 2 faktorer i et fuldstændigt faktorforsøg.

12.2. Én faktor på mere end 2 niveauer.

12.2.1 Binomialfordelt variabel

Ved analysen anvendes formlerne i oversigt 12.1.

Eksempel 12.1 (binomialfordelt variabel).

For hver af 6 leverancer af billige legetøjsbiler udtages en tilfældig prøve på 100 biler, og antallet af defekte biler taltes. Følgende resultater fandtes:

Leverance	1	2	3	4	5	6	7	8
Antal defekte biler	6	14	8	4	7	3	13	7

Foretag en statistisk analyse af, om procenten af defekte biler i de 8 leverancer kan antages at være den samme.

LØSNING:

Lad X_i være antallet af defekte biler i leverance i .

Det antages, at X_i er binomialfordelt $b(100, p_i)$.

$$H_0: p_1 = p_2 = \dots = p_8$$

$$\hat{p}_1 = \frac{6}{100}, \hat{p}_2 = \frac{14}{100}, \dots, \hat{p}_8 = \frac{7}{100}; \hat{p} = \frac{6+14+\dots+7}{800} = \frac{64}{800} = 0.08$$

Da $n_i \cdot \hat{p} = 8 \in [5; 95]$ er forudsætningen for at benyttet oversigt 12.1 opfyldt

$$\chi^2 = \frac{1}{0.08 \cdot (1 - 0.08)} 100((0.06 - 0.08)^2 + (0.14 - 0.08)^2 + \dots + (0.07 - 0.08)^2) = \frac{1.08}{0.08 \cdot 0.92} = 14.67$$

χ^2 er χ^2 -fordelt med frihedsgradstallet $f = n - 1 = 7$

Da P -værdi $= P(\chi^2 > 14.67) = \text{chi2Cdf}(14.67, \infty, 7) = 0.0405 < 0.05$ forkastes nulhypotesen (svagt), dvs. vi har et (svagt) statistisk bevis for, at procenten af defekte biler i leverancen ikke er den samme.



TI -89 og Statgraphics har ikke dette program

12.2.2 Poissonfordelt variabel

Ved analysen anvendes formlerne i oversigt 12.2.

Eksempel 12.2 (Poissonfordelt variabel)

Ved en optælling af hvide blodlegemer i en blodprøve med voluminet v fandtes for 6 personer antallene 14, 28, 18, 23, 15 og 22.

Viser disse resultater, at den gennemsnitlige antal blodlegemer pr. volumenenhed er forskelligt for de tre personer?

LØSNING:

Lad X_1 være antallet af hvide blodlegemer i en blodprøve for person 1

Lad X_2 være antallet af hvide blodlegemer i en blodprøve for person 2

...

Lad X_6 være antallet af hvide blodlegemer i en blodprøve for person 6.

X_i antages at være Poissonfordelt med middelværdi μ_i .

Begrundelse: Benyttes en kanyle til udtagning af blodprøven ankommer de hvide blodlegemer "tilfældigt" i tiden. Det mulige antal blodlegemer er næsten ubegrænset.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$$

Antal elementer i hver stikprøve er 1, dvs. i oversigt 12.4 er $n_1 = n_2 = \dots = n_6 = 1$ og

$$\bar{x}_1 = x_1 = 14, \bar{x}_2 = x_2 = 28, \dots, \bar{x}_6 = x_6 = 22$$

$$\text{Vi får } \bar{x} = \frac{14 + 28 + \dots + 22}{6} = 20.0.$$

Heraf ses, at $n_i \bar{x} \geq 5$, dvs. forudsætningen for at benytte oversigten er opfyldt.

$$\chi^2 = \frac{1 \cdot (14 - 20)^2 + 1 \cdot (28 - 20)^2 + \dots + 1 \cdot (22 - 20)^2}{20} = 7.1$$

χ^2 er χ^2 -fordelt med frihedsgradstallet $f = n - 1 = 5$

Da P -værdi $= P(\chi^2 > 7.1) = \text{chi2Cdf}(7.1, \infty, 5) = 0.2133 > 0.05$ accepteres H_0 , det vil sige, at det ikke er påvist, at det gennemsnitlige antal hvide blodlegemer pr. volumenenhed er forskelligt for de 6 personer.

TI-89 og Statgraphics har ikke dette program



12.3. To faktorer i et fuldstændigt faktorforsøg.

Har man 2 faktorer i en fuldstændig faktorstruktur, og de statistiske variable er enten binomialfordelte eller Poissonfordelte, kan man ikke bruge variansanalyseteknikken, da den kræver, at de variable er normalfordelte. Transformerer data som angivet i oversigt 12.3 er det imidlertid tilladeligt at bruge variansanalyseteknikken på de transformerede data (bevise ikke her). Endvidere får man så den fordel, at man får en eksakt værdi for forsøgsfejls varians (støjen), som bevirker, at selv om man ikke har gentagelser, så kan man dog teste om der er vekselvirkning.

Det skal bemærkes, at testresultaterne er vanskelig at fortolke, så finder man der er vekselvirkning eller hovedvirkninger, så kan man sædvanligvis kun konkludere, at faktorerne har en virkning, men ikke komme nærmere ind på hvorledes denne virkning ytrer sig.

Eksempel 12.3 (E81) Variabeltransformation.

Fire forskellige metoder til anvendelse af et møldræbende middel på uldklæde ønskes sammenlignet ved et forsøg.

For hvert af fire forskellige fabrikater uldklæde udtoges 4 “ens” stykker klæde (20×20 cm), som blev behandlet med hver sin af de fire metoder.

På hvert af de 16 stykker uldklæde anbragtes 25 møllarver, hvorefter man observerede det møldræbende middels virkning på larverne i løbet af et givet tidsrum.

Resultaterne var (målt i antal døde larver):

		Klædefabrikat			
		1	2	3	4
Metode	1	19	18	20	21
	2	17	14	18	18
	3	19	19	20	22
	4	20	19	22	22

Det antages, at antallet af døde larver ved metode i anvendt på klædefabrikat j er binomialfordelt $b(25, p_{ij})$.

Foretag en statistisk analyse af om det møldræbende middels virkning afhænger af metoderne, og af klædefabrikatet.

Løsning:

Da antallet af dræbte larver anses for at være binomialfordelt, foretages den i oversigt 12.3 nævnte variabeltransformation.

De relative hyppigheder beregnes ved at alle tal i skemaet divideres med 25. Derefter beregnes $Y = \text{Arcsin}\sqrt{h_{ij}}$.

Eksempelvis for metode 1 klæde 1:

$$h_{11} = \frac{19}{25} = 0.76 \quad Y = \text{Arcsin}\sqrt{h_{11}} = \text{Arcsin}\sqrt{0.76} = 1.0588$$

		Klædefabrikat			
		1	2	3	4
Metode	1	1.0588	1.0132	1.1071	1.1593
	2	0.9695	0.8455	1.0132	1.0132
	3	1.0588	1.0588	1.1071	1.2171
	4	1.1071	1.0588	1.2171	1.2171

Vi kan nu foretage en sædvanlig tosidet variansanalyse.

Vi antager, at vi har fået beregnet SAK'erne enten som beskrevet i oversigt 11.2 eller ved benyttelse af et statistikprogram.

Varianstabellen kan derefter udfyldes, idet vi dog på forhånd kender variansen for “støjen”, fra

oversigt 12.3 til $\frac{1}{4n} = \frac{1}{4 \cdot 25} = 0.01$.

Variation	SAK	f	s ²	F
Metoder	0.08044448	3	0.0268149	2.68
Klæde	0.0576655	3	0.0192218	1.92
Vekselvirkning	0.00888146	9	0.000986829	0.009 < 1
Gentagelser		∞	0.01	

H_0 : AB = 0 (Model har ingen vekselvirkning) accepteres, da F - værdi er mindre end 1
I det følgende antages, at der ikke er vekselvirkning.

Vi pooler ikke, da vi har et eksakt værdi for "støjens varians".

H_0 : A = 0 (Klædefabrikat har ingen virkning) accepteres, da P -værdi = $P(F > 1.92)$
= $\text{FCdf}(1.92, \infty, 3, 1000) = 0.12 > 0.05$

Konklusion: Klædefabrikat har ingen virkning

H_0 : B = 0 (Metoder har ingen virkning) forkastes, da P - værdi = $P(F > 2.68)$
= $\text{FCdf}(2.68, \infty, 3, 1000) = 0.0457 < 0.05$

Konklusion: Metoder har en (svag) virkning

Skal vi finde ud af hvilken virkning der er størst, kan vi udregne konfidensintervaller for de transponerede tal.

		Klædefabrikat				
		1	2	3	4	Gennemsnit
Metode	1	1.0688	1.0132	1.1071	1.1593	1.0846
	2	0.9695	0.8455	1.0132	1.0131	0.9604
	3	1.0588	1.0588	1.1071	1.2171	1.1105
	4	1.1071	1.0588	1.2171	1.2171	1.1502

Radius i konfidensintervallet er

$$r_{kon} = t_{0.975}(\infty) \frac{s_{\infty}}{\sqrt{n \cdot q}} = 1.96 \frac{\sqrt{0.01}}{\sqrt{1 \cdot 4}} = 0.098$$

Konklusion: Metode 2 er ringere end metode 4, mens de øvrige ikke kan adskilles.



OVERSIGHT 12.1. Oversigt over test af parametre $p_1, p_2, \dots, p_k$ for binomialfordelte variable.

$X_1, X_2, \dots, X_k$ er **binomialfordelt** henholdsvis $b(n_1, p_1), b(n_2, p_2), \dots, b(n_k, p_k)$, hvor $n_1, n_2, \dots, n_k$ er kendte og $p_1, p_2, \dots, p_k$ ukendte.

Observerede stikprøveværdier $x_1, x_2, \dots, x_k$. Signifikansniveau er α . Y er en statistisk variabel, der er χ^2 -fordelt med $k - 1$ frihedsgrader.

Forudsætning:	Nulhypotese	Beregning	H_0 forkastes
Approximativ metode $n_1 \cdot \hat{p} \in [5; n_1 - 5], n_2 \cdot \hat{p} \in [5; n_2 - 5]$ $\dots n_k \cdot \hat{p} \in [5; n_k - 5]$ $\hat{p} = \frac{x_1 + x_2 + \dots + x_k}{n_1 + n_2 + \dots + n_k}$ hvor	$H_0: p_1 = p_2 = \dots = p_k$	$\chi^2 = \frac{1}{\hat{p}(1 - \hat{p})} \sum_{i=1}^k n_i (\hat{p}_i - \hat{p})^2$, hvor $\hat{p}_1 = \frac{x_1}{n_1}, \hat{p}_2 = \frac{x_2}{n_2}, \dots, \hat{p}_k = \frac{x_k}{n_k}$	P -værdi $< \alpha$, hvor P -værdi $= P(Y > \chi^2)$

OVERSIGHT 12.2. Oversigt over test af parametre $\mu_1, \mu_2, \dots, \mu_k$ for Poissonfordelt variable.

$X_1, X_2, \dots, X_k$ er Poissonfordelt henholdsvis $p(\mu_1), p(\mu_2), \dots, p(\mu_k)$, hvor $\mu_1, \mu_2, \dots, \mu_k$ er ukendte. Signifikansniveau er α .

Der foreligger for hver af de variable X_i en stikprøve af størrelsen n_i med gennemsnit $\bar{x}_i$. Y er en statistisk variabel, der er χ^2 -fordelt med $k - 1$ frihedsgrader..

Forudsætning	Nulhypotese	Beregning	H_0 forkastes
Approximativ metode. $n_i \cdot \bar{x} \geq 5$	$H_0: \mu_1 = \mu_2 = \dots = \mu_k$	$\chi^2 = \frac{\sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2}{\bar{x}}$ hvor $\bar{x} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2 + \dots + n_k \cdot \bar{x}_k}{n_1 + n_2 + \dots + n_k}$	P -værdi $< \alpha$ hvor P -værdi $= P(Y > \chi^2)$

OVERSIGHT 12.3. Oversigt over transformation af Binomialfordelte eller Poissonfordelt variable til tosidet variansanalyse.

Variabel	Transformation før tosidet variansanalyse	Forsøgsfejls varians s_0^2 .	Tilhørende frihedsgrad ∞
Relativ hyppighed H H binomialfordelt $b(n, p)$	$Y = \text{Arcsin} \sqrt{H}$	$\frac{1}{4n}$	
X X Poissonfordelt $p(\mu)$	$Y = \sqrt{X}$	$\frac{1}{4}$	

Appendix 12A Eksempler regnet på TI - 89.

1. Indledning.

Det forudsættes, at man kender de grundlæggende operationer på lommeregneren.

I Appendix "Vejledning i TI - 89" er beskrevet, hvorledes man beregner sandsynligheden for forskellige fordelinger, beregner gennemsnit og spredning, samt hvorledes man tester og beregner konfidensintervaller for funktion af 1 variabel. Dette forudsættes ligeledes bekendt.

Til eksemplerne 12.1 og 12.2 svarer intet program

2. Variabeltransformation.

Eksempel 12.3.

Fire forskellige metoder til anvendelse af et møldræbende middel på uldklæde ønskes sammenlignet ved et forsøg. På hvert af de 16 stykker uldklæde anbragtes 25 møllarver.

Resultaterne var (målt i antal døde larver):

		Klædefabrikat			
		1	2	3	4
Metode	1	19	18	20	21
	2	17	14	18	18
	3	19	19	20	22
	4	20	19	22	22

Foretag en statistisk analyse af om det møldræbende middels virkning afhænger af metoderne, og af klædefabrikatet.

Løsning:

Antallet af dræbte larver anses for at være binomialfordelt $b(25, p)$

For hver tal x foretages transformationen $Y = \text{Arcsin} \sqrt{\frac{x}{25}}$.

Klædefabrikat 1 gemmes i list1, klædefabrikat 2 gemmes i list 2 osv.

Der oprettes 4 lister med navnene k1, k2, k3 og k4, hvori de transformerede tal overføres.

APPS, STAT/LIST, I navnefeltet for listen efter list6 skrive k1, næste navnefelt k2 osv.

HOME, $\sin^{-1}(\sqrt{(list1 / 25.0)})$, STO, k1, ENTER osv.

F6, ANOVA2-Way, ENTER

DESIGN=Block, Levels of Col Factor =4, ENTER

Næste skema udfyldes med k1, k2, k3, og k4, ENTER

Resultatet kan umiddelbart aflæses: Nedenfor er de relevante resultater angivet i den sædvanlige variansanalysetabel

Variation	SAK=SS	df	MS= s^2	F	P-værdi
Factor: klædefabrikat	0.057665	3	0.019222		
Block: Metoder	0.080445	3	0.026815		
Error	0.008881	9	0.000987		

Her svarer "Error" en sum af "støj"+"vekselvirkning".

Da vi fra oversigt 12.3 kender den eksakte støj til $\frac{1}{4n} = \frac{1}{4 \cdot 25} = 0.01$ kan tabellen udbygges

Variation	SAK=SS	df	MS= s^2	F
Factor A: klædefabrikat	0.057665	3	0.019222	1.92
Block B: Metoder	0.080445	3	0.026815	2.68
Vekselvirkning AB	0.008881	9	0.000987	0.009
Error		∞	0.01	

H_0 : AB = 0 (Model har ingen vekselvirkning) accepteres, da F - værdi er mindre end 1
I det følgende antages, at der ikke er vekselvirkning.

Vi pooler ikke, da vi har et eksakt værdi for "støjens varians".

H_0 : A = 0 (Klædefabrikat har ingen virkning) accepteres, da P-værdi = $P(F > 1.92)$
= $FCdf(1.92, \infty, 3, 1000) = 0.12 > 0.05$

Konklusion: Klædefabrikat har ingen virkning

H_0 : B = 0 (Metoder har ingen virkning) forkastes, da P-værdi = $P(F > 2.68)$
= $FCdf(2.68, \infty, 3, 1000) = 0.0457 < 0.05$

Konklusion: Metoder har en (svag) virkning

Appendix 12B Eksempler regnet på Statgraphics.

1. Indledning

I “Grundlæggende begreber Appendix A” er beskrevet hvorledes man beregner sandsynligheden for forskellige fordelinger og beregner gennemsnit og spredning. Dette forudsættes bekendt. Til eksemplerne 12.1 og 12.2 svarer intet program

2. Variabeltransformation.

Eksempel 12.3.

Fire forskellige metoder til anvendelse af et møldræbende middel på uldklæde ønskes sammenlignet ved et forsøg. På hvert af de 16 stykker uldklæde anbragtes 25 møllarver. Resultaterne var (målt i antal døde larver):

		Klædefabrikat			
		1	2	3	4
Metode	1	19	18	20	21
	2	17	14	18	18
	3	19	19	20	22
	4	20	19	22	22

Foretag en statistisk analyse af om det møldræbende middels virkning afhænger af metoderne, og af klædefabrikatet.

Løsning:

Antallet af dræbte larver anses for at være binomialfordelt $b(25, p)$

For hver tal x foretages transformationen $Y = \text{Arcsin} \sqrt{\frac{x}{25}}$.

Data indtastes på sædvanlig måde:

```
larver metode klæde
19      M1      K1
18      M1      K2
20      M1      K3
```

osv.

Placer cursor på næste kolonne (Col 4), og omdøb på sædvanlig måde navnet til “transform”.

Placer cursor på “transform”, tryk som før på venstre musetast for at vælge kolonnen, og derefter på højre musetast for at få en lille menu frem.

Vælg (Generate data\ASINR(?))\erstat ? med sqrt(larver/25)\OK)

(ASINR (x) er Arcsin(x) hvor x er regnet i radianer)

Man ser nu, at søjlen “transform” er udfyldt med de transformerede tal.

```
larver metode klæde transform
19      M1      K1      1,05882363875
18      M1      K2      1,0131975001
20      M1      K3      1,10714871779
```

osv.

På de transformerede tal foretages så en tosidet variansanalyse.

Vælg (Compare\ Analysis of Variance\ Multifactor Anova\ok)

I den fremkomne menu: Vælg (“transform”\klik på pilen ved “Dependent variable”\Klik på “metode”\klik på pilen “Factors”\Klik på “klæde”\klik på pilen “Factors”\OK)

Der fremkommer en “Analysis Summary” med nogle statistiske betragtninger.

Vælg (gul ikon = Tabular Options\Anova Tables| OK) (fjern evt. krydset ved “Analysis summary”).

Der fremkommer følgende udskrift:

Analysis of Variance for transform - Type III Sums of Squares					
Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
MAIN EFFECTS					
A:metode	0,0804448	3	0,0268149		
B:klæde	0,0576655	3	0,0192218		
INTERACTIONS					
AB	0,00888146	9	0,000986829		
RESIDUAL	0,0	0			
TOTAL (CORRECTED)	0,146992	15			

All F-ratios are based on the residual mean square error.

Da vi fra oversigt 12.3 kender den eksakte støj til $\frac{1}{4n} = \frac{1}{4 \cdot 25} = 0.01$ kan tabellen udbygges

Variation	SAK	f	s ²	F
Metoder	0.08044448	3	0.0268149	2.68
Klæde	0.0576655	3	0.0192218	1.92
Vekselvirkning	0.00888146	9	0.000986829	0.009 < 1
Gentagelser		∞	0.01	

H_0 : AB = 0 (Model har ingen vekselvirkning) accepteres, da F -værdi er mindre end 1
I det følgende antages, at der ikke er vekselvirkning.

Vi pooler ikke, da vi har et eksakt værdi for “støjens varians”.

H_0 : A = 0 (Klædefabrikat har ingen virkning) accepteres, da P-værdi = $P(F > 1.92)$
= $FCdf(1.92, \infty, 3, 1000) = 0.12 > 0.05$

Konklusion: Klædefabrikat har ingen virkning

H_0 : B = 0 (Metoder har ingen virkning) forkastes, da P-værdi = $P(F > 2.68)$
= $FCdf(2.68, \infty, 3, 1000) = 0.0457 < 0.05$

Konklusion: Metoder har en (svag) virkning

For at få et overblik over hvilken metode, der kan anbefales, beregnes 95% konfidensintervaller (på basis af en model hvor vi har strøget vekselvirkning og klæde).

Vælg(Tabular options \Tables of Means \ OK),

Table of Least Squares Means for transform
with 95,0 Percent Confidence Intervals

Level	Count	Mean	Std. Error	Lower Limit	Upper Limit
GRAND MEAN	16	1,07637			
metode					
M1	4	1,08461	0,0372343	1,00349	1,16574
M2	4	0,960368	0,0372343	0,879241	1,04149
M3	4	1,11046	0,0372343	1,02934	1,19159
M4	4	1,15002	0,0372343	1,06889	1,23115

Heraf kan slutes, at metode 2 er dårligere end metode4.

Konfidensintervallerne er ikke korrekte, da de ikke er baseret på den “eksakte” spredning.

Mere præcist er radius i konfidensintervallet $r_{kon} = t_{0.975}(\infty) \frac{s_{\infty}}{\sqrt{n \cdot q}} = 1.96 \frac{\sqrt{0.01}}{\sqrt{1 \cdot 4}} = 0.098$ ◆

OPGAVER

Opgave 12.1.

I en virksomhed er på hvert af 3 skift arbejdsbetingelser og antal mennesker udsat for risiko tilsyneladende nogenlunde ens.

Ikke desto mindre synes følgende optælling at vise, at risikoen på skift 2 og 3 er større end på skift 1.

Skift	Antal arbejdsulykker
1	10
2	22
3	22

På grundlag af denne statistik finder man, at der bør gøres noget for at nedsætte risikoen i skift 2 og 3. Er dette statistisk velbegrunderet?

Opgave 12.2

5 typer vaccine mod en bestemt sygdom blev undersøgt ved, at 6 grupper på hver 200 forsøgsdyr (mus) blev udsat for smitte. De 5 af grupperne fik hver sin type vaccination, mens den sidste gruppe ikke blev vaccineret. Efter en passende tid undersøgte man hvor mange af de 200 dyr, der havde fået sygdommen. Følgende resultater fandtes:

Gruppe nr	1	2	3	4	5	6
Antal syge dyr	12	13	18	10	16	27

Vi ønsker at foretage en statistisk analyse af, om procenten af smittede dyr i de 6 grupper kan antages at være den samme.

Opgave 12.3.

Ved en tekstilfabrikation målt for to forskellige vævemetoder og 5 forskellige materialetyper antallet af garnbrud pr. 1000 m² klæde. Resultaterne var følgende:

		Materialetyper				
		1	2	3	4	5
Vævemetoder	1	4	12	23	6	9
	2	4	3	7	2	3

Foretag efter en passende variabeltransformation en analyse af, om og i bekræftende fald hvorledes middeltallet af garnbrud afhænger af vævemetoder og/eller materialetyper.

Opgave 12.4.

Ved en undersøgelse af, hvorledes virkningen af forskellige giftstoffer kunne bekæmpes, foretoges et fuldstændigt randomiseret forsøg, hvorved 2 giftstoffer og 4 vitaminbehandlinger inddrogtes i undersøgelsen, og overlevelsestiden (timer) af de benyttede forsøgsdyr målt.

Nedenfor er anført en skematisk oversigt over forsøgsresultaterne:

		Vitaminbehandlinger			
		1	2	3	4
Giftstoffer	1	3.1 4.5	8.2 11.0	4.3 4.5	4.5 7.1
	2	2.2 2.1	3.0 3.7	2.3 2.5	3.0 3.6

Teoretiske overvejelser i forbindelse med tidligere lignende forsøg har vist, at

variabeltransformationen $Y = \frac{1}{X}$ sikrer den for analysen nødvendige varianshomogenitet.

- 1) Analyser forsøgsresultaterne og drag konklusioner med hensyn til faktorenes virkemåde.
- 2) Opstil et 95%-konfidensinterval for den gennemsnitlige middelofterlevelsestid for hver enkelt vitaminbehandling og bestem den vitaminbehandling, som må antages at have bedst virkning.

Grundlæggende operationer på TI - 89.

1. Indledning.

Sædvanligvis vil man i statistik skulle analysere en eller flere kolonner af tal. I sådanne tilfælde benyttes

“Statistikmetoden”: Vælg APPS, Stats/List, indtast data i eksempelvis “list1”, og vælg en relevant “F- knap”.

Der fremkommer nu en menu, som er næsten selvforklarende.

Ønskes et resultat indsat på indtastningslinien: HOME, Var-Link.

I StatsVar mappen markeres den ønskede størrelse, ENTER .

Som eksempel se under afsnit 3: Beregning af gennemsnit m.m.

Skal man beregne sandsynligheder f.eks $P(X < 0.87)$, hvor X har en kendt sandsynlighedsfordeling, så benyttes

“Sandsynlighedsmetoden”: Vælg HOME, CATALOG,, F3, vælg den ønskede fordeling (benyt evt. ALPHA,+ forbogstav for hurtigt at komme til det ønskede navn).

ENTER bevirker at funktionen indsættes på indsætningslinien.

Man indsætter nu parametrene (nederst på skærmen kan man se rækkefølgen af parametrene), og ved et tryk på ENTER fås resultatet.

Oprette og vælge en mappe (“Folder”):

Det kan være praktisk at oprette én eller flere mapper til at have sine data gemt i.

VAR-Link, F1, 5: Create Folder, Skriv navn på folder.

Vælg F2: View, og under Folder vælg den mappe som du vil arbejde i

Vælg en mappe som den aktuelle mappe: MODE, Current Folder,

2. Sandsynlighedsfordelinger.

Normalfordeling $n(\mu, \sigma)$

a) Find $p = P(a \leq X \leq b)$, hvor a, b, μ, σ er givne konstanter.

HOME, Catalog, F3 ,vælg normcdf, ENTER

$\text{normcdf}(a, b, \mu, \sigma)$

b) Find fraktilen x_p : $P(X \leq x_p) = p$, hvor p, μ, σ er givne konstanter.

HOME, Catalog, F3

$x_p = \text{invNorm}(p, \mu, \sigma)$

t - fordeling.

Lad T være t - fordelt med frihedsgradstallet f .

a) Find $p = P(a \leq T \leq b)$, hvor a og b er givne konstanter.

$\text{tCdf}(a, b, f)$

b) Find fraktilen $t_\alpha(f)$: $P(T \leq t_\alpha(f)) = \alpha$ (α given konstant).

$\text{invT}(\alpha, f)$

F-fordeling.

Lad F være F -fordelt med tællerfrihedsgradstallet f_T og nævnerfrihedsgradstallet f_N .

a) Find $p = P(a \leq F \leq b)$, hvor a og b er givne konstanter. $FCdf(a, b, f_T, f_N)$

b) Find fraktilen $F_\alpha(f_T, f_N)$: $P(F \leq F_\alpha(f_T, f_N)) = \alpha$ $invF(\alpha, f_T, f_N)$

3 Beregning af gennemsnit, varians og spredning

APPS, Stats/List, Data indtastes i "list1", F4, 1: 1-Var Stats,

I menu sættes "List" til "List1" (Benyt evt. Var-Link til at finde List1)

Eksempel: (hentet fra "Statistiske grundbegreber" eksempel 5.1 side 53)

Forsøg nr	1	2	3	4	5	6	7	8	9	10	11	12
x	68.8	70.7	70.3	70.1	70.7	68.7	69.2	68.9	70.0	69.6	71	69.1

Udskriften består af en række statistiske størrelser, blandt hvilke man må finde det ønskede resultat. Man finder $\bar{x}=69.7583$, $s_x=0.8163$

Skal man regne videre med et resultat, f.eks finde variansen, skal resultatet indsættes på indtastningslinien:

Vælg Home og Var-Link. I StatsVar-mappen markerer man den ønskede størrelse. Tryk på ENTER indsætter så størrelsen på indsætningslinien.

Ønskes beregning af en enkelt størrelse f.eks gennemsnit = mean, spredning= stdDev eller varians = Variance

HOME, CATALOG,, F3,

Eksempel : Variance(list1)

Resultat: 0.666288

4. Hypotesetest og konfidensintervaller for 1 variabel.

APPS, STAT/LIST hvorefter eventuelle data indtastes i list1, list 2 osv.

4.1. Normalfordeling.

a1) Hypotesetest; σ kendt: F6, 1: Z-Test

I menu: Er data givet i en liste vælges Data.

Er kun gennemsnit (og σ) kendt vælges Stats.

Eksempel: Vælges ovennævnte data i list1, $\sigma=1$ og $H: \mu > 69.2$, fås P-værdi=2.48%.

a2) Konfidensinterval σ kendt: F7, 1: Z-Interval

Eksempel: Vælges ovennævnte data i list1 og er $\sigma=1$, fås $C\ Int=[69.2; 70.3]$

b1) σ ukendt: F6, 2: T-Test . Derefter som under punkt a)

b2) σ ukendt: F7, 2: T-Interval . Derefter som under punkt a)

4.2. Binomialfordeling.

a1) Hypotesetest: F6, 5: 1-Prop-ZTest

Menu selvforklarende. Bemærk: Kræver der kan approksimeres til normalfordeling

a2) Konfidensinterval: F7, 5: 1-Prop-ZInt

Bemærk: Kan kun benyttes, hvis kan approksimere til normalfordeling

Eksempel: Er af 100 forsøg de 85 en succes, så fås $C\ Int=[0.78; 0.93]$

Poissonfordeling: findes ikke, så her må formel for konfidensinterval benyttes

STATISTISKE TABELLER

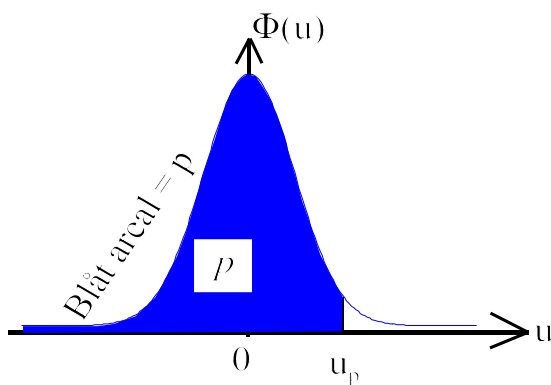
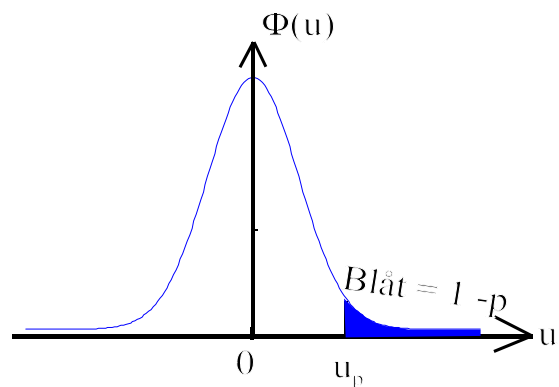
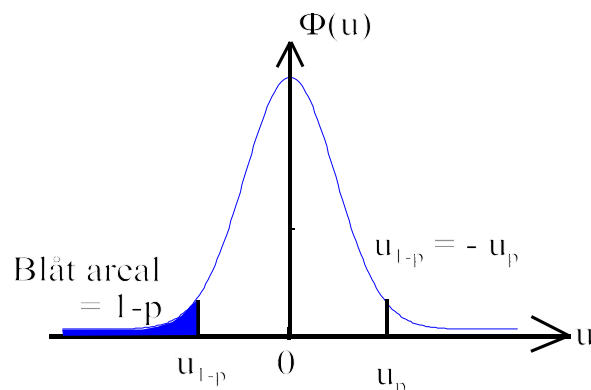
Tabel 1 Fraktiler u_p i U-fordelingen $n(0,1)$. $P(U \leq u_p) = p$.

Bemærk: $u_p = -u_{1-p}$

p	0.0005	0.001	0.005	0.01	0.025	0.05	0.10
u_p	-3.291	-3.090	-2.576	-2.326	-1.960	-1.645	-1.282

p	0.90	0.95	0.975	0.99	0.995	0.999	0.9995
u_p	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Eksempler: $u_{0.975} = 1.960$

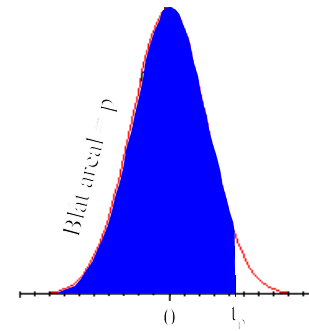


Tabel 2 Fraktiler t_p i t -fordelingen $t(f)$.

$$P(T \leq t_p) = p, \text{ hvor } T = \frac{(\bar{X} - \mu)\sqrt{n}}{s}$$

Eksempler: For $t(27)$ er $P(X \leq 3.42) = 0.999$.

$$t_{0.95}(1) = 6.31. \quad t_{0.05}(10) = -t_{0.95}(10) = -1.81.$$



f \ p	0.60	0.75	0.90	0.95	0.975	0.99	0.995	0.999	0.9995
1	0.33	1.00	3.08	6.31	12.7	31.8	63.7	318	637
2	0.29	0.82	1.89	2.92	4.30	6.97	9.93	22.3	31.6
3	0.28	0.74	1.64	2.35	3.18	4.54	5.84	10.2	12.9
4	0.27	0.74	1.53	2.13	2.78	3.75	4.60	7.17	8.61
5	0.27	0.72	1.48	2.02	2.57	3.37	4.03	5.89	6.86
6	0.27	0.72	1.44	1.94	2.45	3.14	3.71	5.21	5.96
7	0.26	0.71	1.42	1.90	2.37	3.00	3.50	4.79	5.41
8	0.26	0.71	1.40	1.86	2.31	2.90	3.36	4.50	5.04
9	0.26	0.70	1.38	1.83	2.26	2.82	3.25	4.30	4.78
10	0.26	0.70	1.37	1.81	2.23	2.76	3.17	4.14	4.59
11	0.26	0.70	1.36	1.80	2.20	2.72	3.11	4.03	4.44
12	0.26	0.70	1.36	1.78	2.18	2.68	3.06	3.93	4.32
13	0.26	0.69	1.35	1.77	2.16	2.65	3.01	3.85	4.22
14	0.26	0.69	1.35	1.76	2.15	2.62	2.98	3.79	4.14
15	0.26	0.69	1.34	1.75	2.13	2.60	2.95	3.73	4.07
16	0.26	0.69	1.34	1.75	2.12	2.58	2.92	3.69	4.02
17	0.26	0.69	1.33	1.74	2.11	2.57	2.90	3.65	3.97
18	0.26	0.69	1.33	1.73	2.10	2.55	2.88	3.61	3.92
19	0.26	0.69	1.33	1.73	2.09	2.54	2.86	3.58	3.88
20	0.26	0.69	1.33	1.73	2.09	2.53	2.85	3.55	3.85
21	0.26	0.69	1.32	1.72	2.08	2.52	2.83	3.53	3.82
22	0.26	0.69	1.32	1.72	2.07	2.51	2.82	3.51	3.79
23	0.26	0.69	1.32	1.71	2.07	2.50	2.81	3.49	3.77
24	0.26	0.69	1.32	1.71	2.06	2.49	2.80	3.47	3.75
25	0.26	0.68	1.32	1.71	2.06	2.49	2.79	3.45	3.73
26	0.26	0.68	1.32	1.71	2.06	2.48	2.78	3.44	3.71
27	0.26	0.68	1.31	1.70	2.05	2.47	2.77	3.42	3.69
28	0.26	0.68	1.31	1.70	2.05	2.47	2.76	3.41	3.67
29	0.26	0.68	1.31	1.70	2.05	2.46	2.76	3.40	3.66
30	0.26	0.68	1.31	1.70	2.04	2.46	2.75	3.39	3.65
40	0.26	0.68	1.30	1.68	2.02	2.42	2.70	3.31	3.55
50	0.26	0.68	1.30	1.68	2.01	2.40	2.68	3.26	3.50
60	0.25	0.68	1.30	1.67	2.00	2.39	2.66	3.23	3.46
80	0.25	0.68	1.29	1.66	1.99	2.37	2.64	3.20	3.42
100	0.25	0.68	1.29	1.66	1.98	2.37	2.63	3.17	3.39
120	0.25	0.68	1.29	1.66	1.98	2.36	2.62	3.16	3.37
200	0.25	0.68	1.29	1.65	1.97	2.35	2.60	3.13	3.34
500	0.25	0.68	1.28	1.65	1.97	2.33	2.59	3.11	3.31
∞	0.25	0.67	1.28	1.65	1.96	2.33	2.58	3.09	3.29

TABEL 3. Dimensioneringstabel.

Dimensionering med henblik på udførelse af **t-test vedrørende 2 statistiske variable** X_1 og X_2 som er normalfordelte $n(\mu_1, \sigma_1)$ og $n(\mu_2, \sigma_2)$, hvor μ_1 , μ_2 , σ_1 og σ_2 er ukendte.

Størrelse n af stikprøve på X_1 og X_2 :	Værdier af $\frac{\Delta}{\sigma}$ for $\alpha = 5\%$.			
	Ensided test $H_0: \mu_1 \leq \mu_2$ eller $H_0: \mu_1 \geq \mu_2$		Tosidet test $H_0: \mu_1 = \mu_2$	
n	$\beta = 5\%$	$\beta = 10\%$	$\beta = 5\%$	$\beta = 10\%$
3	3.32	2.94	4.02	3.59
4	2.65	2.36	3.09	2.77
5	2.29	2.03	2.62	2.35
6	2.05	1.82	2.32	2.08
7	1.87	1.66	2.10	1.89
8	1.73	1.54	1.94	1.75
9	1.62	1.44	1.82	1.64
10	1.53	1.36	1.71	1.53
11	1.45	1.29	1.62	1.46
12	1.39	1.23	1.54	1.39
13	1.33	1.18	1.48	1.33
14	1.28	1.14	1.42	1.27
15	1.23	1.10	1.37	1.23
16	1.19	1.06	1.32	1.19
17	1.15	1.03	1.28	1.15
18	1.12	1.00	1.24	1.11
19	1.09	0.97	1.20	1.08
20	1.06	0.94	1.17	1.05
22	1.01	0.90	1.11	1.00
24	0.96	0.86	1.06	0.96
26	0.93	0.82	1.02	0.92
28	0.89	0.79	0.98	0.88
30	0.86	0.76	0.95	0.85
40	0.74	0.66	0.82	0.73
60	0.60	0.54	0.66	0.60

FACITLISTE

KAPITEL 10

- 10.1** (1) 22 (2) 24 (3) - (4) -
10.2 30
10.3 (1) $u = 3.09$, P - værdi = 0.001 (2) [0.70 ; 3.12]
10.4 (1) $F = 7.42$ (2) $t = 0.837$, P - værdi = 0.2007
 (3a) [- 2.09 ; 4.84] (3b) [0.91 ; 6.40]
10.5 $F = 1.11$, P - værdi = 0.4557, $t = 1.393$, P - værdi = 0.0969
10.6 $F = 1.18$, P - værdi = 0.4117 $t = - 2.64$, P - værdi = 0.0102
10.7 $F = 3.68$, P - værdi = 0.1175, $t = 2.34$, P - værdi = 0.0237, [0.031 ; 1.87]
10.8 (1) 18 (2) $F = 2.171$ P - værdi = 0.113, $t = 4.74$ P - værdi = 0.00006, [11.9 ; 15.0]
10.9 $F = 6.00$, P - værdi = 0.0089, $t = 0.56$, P - værdi = 0.295
10.10 $F = 1.563$, P - værdi = 0.2125 $t = -1.497$, P - værdi = 0.0728
10.11. (1) $u = 1.98$, P - værdi = 0.0238
 (2) $F = 1.56$, P - værdi = 0.0627, $t = 2.75$, P - værdi = 0.00355
10.12 $u = - 0.729$, P - værdi = 0.233
10.13 (1) $u = 2.21$, P - værdi = 0.0135
 (2) $F = 7.29$, P - værdi = 0.0022, $t = 4.25$, P - værdi = 0.00012
10.14 (1) $u = 0.654$, P - værdi = 0.2566 (2) $u = 2.10$, P - værdi = 0.0179
10.15 $u = 3.45$, P - værdi = 0.00028

KAPITEL 11

- 11.1** ja, $F = 13.4$, P - værdi = 0.0004, B2 nok bedst, men et lille overlap med B3
11.2 $F = 3.51$, P - værdi = 0.0218
11.3 (1) $F = 4.94$, P - værdi = 0.0096 (2) S2, S4
11.4 (1) P - værdi = 0.838 (2) nej, P - værdi = 0.0108, (3a) A1 så K1, A2: frit valg, A3: frit valg.
 (3b) K1: frit valg, K2 så ikke A1, (3c) ikke A1 K2
11.5 (1) ja, P - værdi = 0.4575 (2) ja, P - værdi = 0.3178 (3) ja, ingen outliers,
 (4) glas 2 og enten fosfor 1 eller 3.
11.6 (1) - (2) limtyper (3) II: [23.76 ; 28.10], III: [22.69 ; 27.04]
11.7 (1) Kun syrer har virkning, Svovlsyre størst. (2) 26.45, [25.17 ; 27.73]
11.8 (1) Nej, P - værdi = 0.0463 (2) ja, P - værdi = 0.0777 (3) -
11.9 (1) - (2) -
11.10 ja, R forskellig fra de øvrige, ja
11.11 (1) T evt. Q må foretrækkes. (2) -

KAPITEL 12

- 12.1** $\chi^2 = 5.33$, P - værdi = 0.0696
12.2 $\chi^2 = 12.64$, P - værdi = 0.027
12.3 Begge har signifikant virkning, vælg vævemetode 2, og ikke matrialtype 3.
12.4 Begge, Vitaminbehandling 1 eller 3.

STIKORDSREGISTER

A

additiv model 39 , 42
 appendix
 10A Eksempler regnet på TI-89 19
 10B Eksempler regnet på Statgraphics 22
 11A Eksempler regnet på TI-89 53
 11B Eksempler regnet på Statgraphics 59
 12A Eksempler regnet på TI-89 79
 12B Eksempler regnet på Statgraphics 81

B

bagatelgrænse 3
 Bartlett's test 33, 48
 behandlinger 1
 binomialfordeling
 test
 to variable 13, 16
 mere end 2 variable 74, 78
 Statgraphics 24
 TI - 89 21
 blokforsøg, randomiseret 44

C

D

delforsøg 3
 dimensionering 3
 Maple program 4, 18
 tabel 89

E

én faktor ad gangen 37
 ensidet variansanalyse 32
 beregninger-formler 32, 35, 47
 med Statgraphics 59
 med TI-89 53

F

facitliste 90
 fejl af type I 3
 fejl af type II 3
 F - fordeling
 test 5, 34, 48
 fordeling, tabel over
 fraktiler for normalfordeling 87

fraktiler for t - fordeling 88
 forudsætninger
 variansanalyse 33, 39
 fraktiltabel for
 normalfordeling 87
 t - fordeling 88
 fuldstændig faktorstruktur 38
 fuldstændig randomiseret blokforsøg 44

G

H

hypotesetest 1
 hyppighed, relativ 76

I,J

interaction plot 62

K

konfidensinterval
 differens, 2 normalfordelte variable 7,12,
 13, 14
 differens, 2 binomialfordelte variable 10,
 16
 differens, 2 Poissonfordelte variable 11, 17
 i ensidet variansanalyse 34, 36, 47
 i tosidet variansanalyse 41, 43, 52
 LSD 34, 36, 47
 kvalitativ faktor 31, 37
 kvantitativ faktor 31

L

Levines test 33, 49
 LSD (Least Signifikant Difference) 34, 36

M

Maple program til dimensionering 4, 18
 median 49

N

niveau for faktor 37
 normalfordeling
 plot 33, 49
 tabel 87
 test, 2 variable

Stikord

middelværdi 4, 5, 7, 8, 9, 12, 13, 14
varianser 5, 15

O

opgaver

kapitel 10 26
kapitel 11 69
kapitel 12 84

oversigt

kapitel 10 12
kapitel 11 47
kapitel 12 78

P

planlægning af forsøg 1, 37

Poissonfordeling

test

2 variable 11, 17
mere end 2 variable 75, 78

poolet estimat for varians 33

Q

R

randomisering 2

randomiseret forsøg 2

randomiseret blokforsøg 44

S

SAK 41, 50

Satterthwaites test 7, 13

Statgraphics

løsning

blokforsøg 67
to binomialfordelte variable 24
to normalfordelte variable 22
to Poissonfordelte variable 25
variansanalyse, ensidet 59
variansanalyse, tosidet 61

statistisk uafhængige 33

styrke af test 7

støj 1

T

t - tabel over fraktiler 88

TI-89

løsning

blokforsøg 58
to binomialfordelte variable 21

to normalfordelte variable 19

variansanalyse, ensidet 53

variansanalyse, tosidet 54

tosidet variansanalyse 50

transformation 75, 78

U

uafhængige statistiske variable 33

V

variabeltransformation 75, 78

variansanalyse

ensidet 32

beregninger, formler 32, 35, 47

med Statgraphics 59

med TI-89 53

tosidet 37

beregninger, formler 39, 40, 50

med Statgraphics 61

med TI-89 54

vekselvirkning 39, 40

varianshomogenitet 33, 48