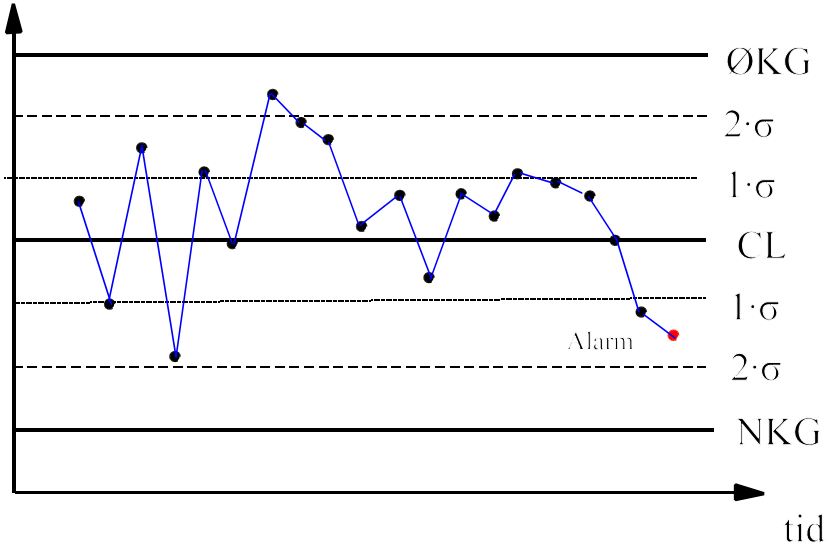


VIDEREGÅENDE STATISTIK

herunder kvalitetskontrol



FORORD

Denne lærebog kan læses på baggrund af en statistisk viden svarende til lærebogen "M. Oddershede Larsen : Statistiske grundbegreber".

Læsning:

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

I slutningen af nogle af kapitlerne er givet en oversigt over centrale formler eller fremgangsmåder.

I et appendix sidst i bogen er givet en mere dybtgående statistisk forklaring på formlerne.

Efter hvert kapitel er der nogle opgaver og en facitliste til opgaverne findes bagerst i bogen.

Hypotesetest for 1 variabel er en repetition

De tre hovedemner

"Faktorer på 2 eller flere niveauer": kapitel 5, 6 og 7

Regressionsanalyse" kapitel 8 og

"Kontrolteori": kapitel 2 og 3

kan læses stort set uafhængigt af hinanden.

Regnemidler

I eksemplerne er beregningerne er i videst mulig omfang foretaget ved anvendelse af PC-programmet TI-Nspire idet dog der også er angivet ordrene i lommeregnerne TI-nspire og TI-89.

Man kan på min hjemmeside www.larsen-net.dk under "statistik 3" finde bøger hvor der er anvendt statistikprogrammerne SAS.JMP og STATGRAPHICS og regnearket EXCEL.

Ønskes bøger indenfor grundlæggende statistik findes de også på ovenstående adresse under "statistik 2"

14. juli 2016

Mogens Oddershede Larsen.

INDHOLD

1 REPETITION AF HYPOTESETEST FOR 1 VARIABEL

1.1	Indledning	1
1.2	Normalfordelt variabel	1
1.3	Binomialfordelt variabel	3
1.4	Poissonfordelt variabel	4
	Opgaver	5

2 STATISTISK PROCESKONTROL

2.1	Indledning	7
2.2	Proces i statistisk kontrol	8
2.3	Opbygning og alarmkriterier for kontrollkort	9
2.4	Kontrollkortanalyse	9
2.5	Procesvariablen er normalfordelt	10
	2.5.1 Stikprøvestørrelse større end 1	10
	2.5.2 Stikprøvestørrelse $n = 1$	14
2.6	Løbende kontrol	15
2.6	Tolerancegrænser og kapabilitet	18
2.7	Procesvariablen X er diskret	20
	2.7.1 X er binomialfordelt	20
	2.7.2 X er Poissonfordelt	22
	Opgaver	23

3 STATISTISK GODKENDELSESKONTROL

3.1	Indledning	27
3.2	Enkelt stikprøveplan	28
3.3	Rektificerende kontrol	31
3.4	Dobbelt stikprøveplan	32
	Opgaver	38

4 PLANLÆGNING AF FORSØG

4.1	Indledning	40
4.2	Nomenklatur	40
4.3	Krav til statistisk gyldigt forsøg	41

5 1 FAKTOR PÅ 2 NIVEAUER

Indhold

5.1	Normalfordelte variable	45
5.1.1	Test af differens mellem middelværdier	45
5.1.2	Parvise observationer (blokforsøg)	49
5.2	Binomialfordelte variable	51
5.3	Poissonfordelte variable	52
5.4	Fordeling ukendt (rangtest)	53
5.5	Oversigt over centrale formler i kapitel 3	54
5.5.1	Test af differens mellem middelværdier for 2 normal-fordelte variable	54
5.5.2	Test af differens mellem varians for 2 normalfordelte variable	56
5.5.3	Test og konfidensintervaller af $p_1 - p_2$ for 2 binomialfordelte variable	57
5.5.4	Test og konfidensintervaller af $\mu_1 - \mu_2$ for 2 Poissonfordelte variable	57
	Opgaver	58
6	1 FAKTOR PÅ MERE END 2 NIVEAUER	
6.1	Indledning	62
6.2	Ensidet variansanalyse (normalfordelte variable)	63
6.2.1	Indledning	63
6.2.2	Forklaring af metoder og formler	64
6.2.3	Beregning af ensidet variansanalyse	68
6.3	Fuldstændigt randomiseret blokforsøg	71
6.4	Binomialfordelte variable	73
6.5	Poissonfordelte variable	74
6.6	Oversigt over centrale formler i kapitel 6	75
6.6.1	Oversigt over fremgangsmåde ved ensidet variansanalyse	75
6.6.2	Test af parametre p_1, p_2, \dots, p_k for binomialfordelte variable	76
6.6.3	Test af parametre $\mu_1, \mu_2, \dots, \mu_k$ for Poissonfordelte variable	76
	Opgaver	77
7	2 FAKTORER PÅ 2 ELLER FLERE NIVEAUER. TOSIDET VARIANSANALYSE	
7.1	Indledning	80
7.2	Planlægning af forsøg	80
7.2.1	Een faktor ad gangen	80
7.2.2	Fuldstændig faktorstruktur	81
7.3	Formler og metode	82
7.4	Beregning af tosidet variansanalyse	84
7.5	Fuldstændigt randomiseret blokforsøg	89

7.6	To binomialfordelte eller Poissonfordelte faktorer i et fuldstændigt faktorforsøg. .	91
7.7	Oversigt over fremgangsmåde ved tosidet variansanalyse.	92
	Opgaver	95
8	FLERE END TO FAKTORER, SCREENINGSFORSØG	
8.1	Beskrivelse af forsøg	98
8.2	Statistikprogrammer, der kan foretage beregningerne	99
9	ENKELT REGRESSIONSANALYSE	
9.1	Indledning	100
9.2	Bestemmelse af regressionsligning	101
9.3	Vurdering af om regressionsligning beskriver data godt.	103
9.4	Test og konfidensintervaller.	109
9.5	Transformation af data.	111
9.6	Enkelt regressionsanalyse med flere y - observationer for hver x - værdi	114
9.7	Multipel regressionsanalyse.	117
	9.7.1 Indledning	117
	9.7.2 Analyse med én y - observation for hver x - værdi	117
9.8	Polynomial regressionsanalyse	120
	9.8.1 Indledning	120
	9.8.2 Beregning af polynomial regressionsanalyse	122
9.9	Oversigt over fremgangsmåde ved regressionsanalyse	126
	Opgaver	128
10	RANGTEST (FORDELING UKENDT)	
10.1	Indledning.	135
10.2	Wilcoxons rangtest for 1 stikprøve.	136
10.3	Wilcoxons rangtest for 2 uafhængig stikprøver	138
10.4	Kruskal Wallis test	140
	Opgaver	142
	APPENDIX	
2.1	Begrundelse for grænserne for kontrollkort	144
6.1	Formler til beregning af ensidet variansanalyse	145
7.1	Formler til beregning af tosidet variansanalyse	148
7.2.	Transformation af binomial-og Poissonfordelte variable til tosidet variansanalyse	151
9.1.	Formler til beregning af enkelt regressionsanalyse uden gentagelser	152

Indhold

9.2	Formler til beregning af enkelt regressionsanalyse med lige mange gentagelser .	155
9.3	Transformation til lineær model	156
9.4	Formler til beregning af multipel regressionsanalyse.....	157
10.1	Kruskal-Wallis rangtest for to eller flere variable	160
	GRUNDLÆGGENDE OPERATIONER PÅ TI89 og TI-Nspire	161
	FACITLISTE	165
	STIKORD	168

1 REPETITION AF HYPOTESETEST FOR 1 VARIABEL

1.1. INDLEDNING

De grundlæggende begreber vedrørende hypotesetest, konfidensintervaller og dimensionering af forsøg blev i "Statistiske Grundbegreber" grundigt beskrevet når vi havde én stikprøve. Beregningerne blev der udført ved anvendelse af pc-programmet TI-Nspire og regnearket Excel.

1.2. NORMALFORDELT VARIABEL

Eksempel 1.1 Hypotesetest. Normalfordelt variabel .

En fabrik der fremstiller plastikprodukter ønsker at evaluere hårheden af rektangulære støbte plastik blokke som anvendes i møbelfabrikationen.

Der udtages tilfældigt 50 blokke, og deres hårhed måles (i Brinell enheder) .

Resultaterne var følgende

283.5	273.3	278.8	238.7	334.9	302.6	239.9	254.6	281.9	270.4	269.1	250.1	301.6	289.2
240.8	267.5	279.3	228.4	265.2	285.9	279.3	252.3	271.7	235.0	313.2	277.8	243.8	295.5
249.3	228.7	255.3	267.2	253.3	281.0	302.1	256.3	233.0	194.4	219.9	263.7	273.6	267.7
283.1	260.9	274.8	277.4	276.9	259.5	262.0	263.5						

- Undersøg om tallene er rimeligt normalfordelt ved at tegne et histogram, et boxplot og et normalfordelingsplot.
Angiv endvidere gennemsnit og spredning.
- Hårheden bør være over 260 (brinell enheder).
Test på et signifikansniveau på $\alpha = 5\%$ om dette er tilfældet.
- Forudsat hårheden er signifikant over 260 brinell, skal angives et estimat for hårheden, samt et 95% konfidensinterval for denne.

Løsning:

- a) Lister og regneark ► kald liste eksempelvis for x , og indtast data.

Da vi ønsker at tegne et histogram, med lige store intervaller, og med ca. $\sqrt{50} \approx 7$ intervaller , vælges

Indsæt ► noter ► matematikfelt ► beregninger ► statistik ► listematematik ► maksimum osv.

$$\min(\mathbf{x}) \rightarrow 194.4 \quad \max(\mathbf{x}) \rightarrow 334.9 \quad \frac{334.9-194.4}{7} \rightarrow 20.0714$$

A _x
283.5
273.3
278.8
238.7
334.9
302.6

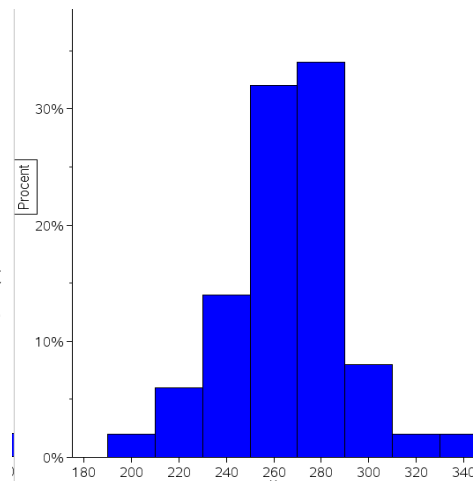
1. Repetition af centrale begreber

Histogram:

Marker listen ved at trykke på listebogstav (A) , højre musetast ► vælg Data ► hurtiggraf ► højre musetast på graf ► vælg histogram

Der fremkommer nu et histogram

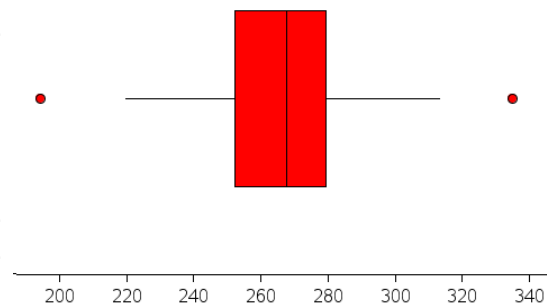
Højre musetast på figur vælg søjleindstillinger ► lige store intervaller ► vælg bredde og start ► vælg bredde 20 ► søjlestart 190 ► skala procent.



Histogrammet ser nogenlunde symmetrisk ud med et enkelt maksimum, så det tyder på at data er normalfordelt

Boxplot

Gentag , men vælg nu boxplot fremfor histogram.



Den midterste streg angiver medianen og kassens grænser angiver henholdsvis 1. og 3. kvartil.

Det betyder, at hvis man opstillede de 50 tal i rækkefølge efter størrelse, så er tal nr $50/2 = 25$ medianen. Peger man på linien fremgår det, at median er 267.6. 1. kvartil er 225.175 osv.

Da boxplottet er nogenlunde symmetrisk om medianen, så kan man igen antage at data er rimelig normalfordelt.

De isolerede prikker yderst viser, at der er et par værdier, som afviger kraftigt fra de øvrige, og muligvis er fejlmålinger (kaldes outliers).

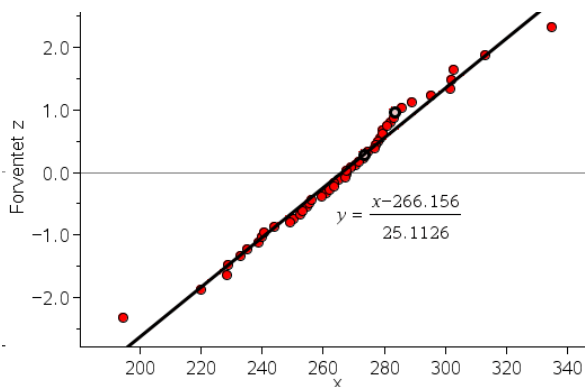
Normalfordelingsplot

Gentag, men vælg nu normalfordelingsplot.

Her har man ud af y - akse sørget for at skalaen er sådan, at punkterne burde ligge på den rette linie, hvis de fuldstændigt eksakt var normalfordelt.

Den rette linie går gennem (mean, 0)

Som det ses, ligger punkterne tæt på linien for de midterste 75% af tallene. De yderste punkter kan man ikke forvente ligger på linien



Man må derfor igen antage, at data er tilnærmelsesvis normalfordelt.

I noter vælges som da man fandt maksimum $\text{mean}(\mathbf{x}) \rightarrow 266.156$ $\text{stDevSamp}(\mathbf{x}) \rightarrow 25.1126$

Gennemsnittet $\bar{x} = \underline{266.16}$

Spredningen $s = \underline{25.11}$

b) $X =$ holdbarheden af plastblokke

X antages normalfordelt med ukendt middelværdi μ og σ .

$H_0: \mu = 260$ $H: \mu > 260$

Da spredningen ikke er kendt eksakt anvendes en t-test.

TI-Nspire: Beregner ► Statistik ► Statistiske test ► t-test for 1 middelværdi ► statistik ► data ► menu udfyldes ► ENTER

Vi får: P-værdi = 0.0429

P-værdien = sandsynligheden for at begå en "type 1 fejl", dvs. påstå at $\mu > \mu_0 = 260$ selv om det ikke er tilfældet.

Da P-værdi = 4.29 % < 5%, forkastes H_0 (svagt).

Konklusion: Vi har bevist, at holdbarheden i middel er over 260 brinell.

c) **TI-Nspire:** Beregner ► Statistik ► konfidensintervaller ► t-interval for 1 middelværdi ► statistik ► menu udfyldes ► ENTER

Et estimat for holdbarheden er 266.2 brinell Et 95% konfidensinterval [259.1 ; 273.3] ◆

1.3. BINOMIALFORDELT VARIABEL

Eksempel 1.2. Binomialtest

En fabrikant af chip til computere reklamerer med, at højst 2% af en bestemt type chip, som fabrikken sender ud på markedet er defekte.

Et stort computerfirma, vil købe et meget stort parti af disse chip, hvis påstanden er rigtigt. For at teste påstanden købes 1000 af dem. Det viser sig, at 33 ud af de 1000 er defekte.

a) Kan fabrikantens påstand på denne baggrund forkastes på signifikansniveau 5% ?

b) Forudsat påstanden forkastes, skal angives et estimat for % defekte, samt et 95% konfidensinterval for denne.

Løsning:

$X =$ antal defekte chips af 1000

X er binomialfordelt $b(1000, p)$.

Nulhypotese: $H: p = 0.02$ Alternativ hypotese $H: p > 0.02$

TI-Nspire

a) P -værdi = $P(X \geq 33) = \text{binomCdf}(1000, 0.02, 33, 1000) = 0.00433$

Da P-værdi < 0.05 forkastes H_0 , dvs. fabrikantens påstand om færre end 2% defekte forkastes.

b) Da $x = 33 > 5$ og $33 < 1000 - 5$ kan approksimeres med normalfordelingen

TI-Nspire: Beregner ► Statistik ► konfidensintervaller ► z-interval for 1 andel ► menu udfyldes ► ENTER

TI89: APPS, STAT/LIST ► F7 ► 5: 1-Prop-Z-interval ► udfyld menu

Estimat for p: 3.3% 95% konfidensinterval : [2.19% ; 4.41%] ◆

1.4 Poissonfordelt variabel

Eksempel 1.5: Poissontest

På en fabrik fremstilles kobberkabler af en bestemt tykkelse. Mikroskopiske revner forekommer tilfældigt langs disse kabler. Man har erfaring for, at der i gennemsnit er 12.3 af den type revner p. 10 meter kabel.

Fabrikken går nu over til en anden og billigere produktionsmetode. For at få et estimat for middelværdien ved den nye metode målt antallet af revner på 12 kabelstykker på hver 10 meter. Man fandt i alt 94 revner i 120 m kabel.

Test, om den nye metode giver færre revner end den gamle metode.

Løsning

X = antal revner i 120 m kabel ved ny metode

X er Poissonfordelt

Ved gammel metode er antal revner i 120 m kabel i middel $\mu_0 = 12.3 \cdot 12 = 147.6$

Nulhypotese $H_0: \mu = 147.6$ mod den alternative hypotese $H: \mu < 147.6$.

$$P\text{-værdi} = P(X \leq 94) = \text{poissCdf}(147.6, 0, 94) \approx 0.000002$$

Da P - værdi < 0.05 forkastes nulhypotesen (stærkt), dvs. vi er sikre på, at middelantallet af revner er blevet formindsket ved at anvende den nye metode



OPGAVER

Opgave 1.1

Færdselspolitiet overvejede, om der burde indføres en fartgrænse på 70 km/h på en bestemt landevejsstrækning, hvor der hidtil havde været fartgrænsen 80 km/h.

Som et led i analysen af hensigtsmæssigheden af den overvejede ændring observeredes inden for et bestemt tidsrum ved hjælp af radarkontrol de forbipasserende bilers fart. Resultatet af målingerne var:

50 observationer									
64	72	82	52	60	95	86	70	63	48
50	63	35	60	77	41	47	88	62	66
59	49	55	99	65	76	76	68	51	80
75	74	64	74	62	70	85	73	93	65
98	55	85	80	78	53	96	71	84	103

Undersøg om tallene er rimeligt normalfordelt ved i TI-Nspire, at tegne et histogram, et boxplot og et normalfordelingsplot.

Angiv endvidere gennemsnit og spredning.

Opgave 1.2

Under produktionen forekommer blandt en fabriks affaldsprodukter 1.5 mg/l af et stof A., som i større mængder kan være kræftfremkaldende. Man håber ved en ny og mere kostbar metode, at formindske indholdet af det pågældende stof.

a) Ved en række kontrolmålinger efter tilsætning af additivet fandtes følgende resultater (i mg/l)

1.12	1.47	1.35	1.27	1.17	1.26	1.83	1.10	1.39	1.25	1.44	1.14
------	------	------	------	------	------	------	------	------	------	------	------

Test på 5% niveau, om målingerne beviser, at der er sket en formindskelse af middelindholdet af stoffet A.

b) Forudsat middelindholdet er signifikant under 1.5 mg/l, skal angives et estimat for det nye middelindhold, samt et 95% konfidensinterval for denne.

Opgave 1.3

Det forventes, at lovgivningen bliver strammet omkring mængden af skadelige partikler i bilers udstødningssgas. En person mener, at mere end 20% af forsvarets biler ikke vil opfylde de forventede nye krav. Ved en undersøgelse af 40 af forsvarets biler tilfældigt udvalgt, fandt man, at 13 af disse ikke kunne opfylde de nye krav.

- 1) Test om dette på et signifikansniveau på 5 % er et bevis for, at mere end 20% af forsvarets biler udsender flere skadelige partikler end ønskeligt.
- 2) Under forudsætning af at det er signifikant, at 20% af bilerne ikke opfylder kravet, skal man angive et estimat for hvor mange procent af bilerne, der ikke opfylder de nye krav, samt angive et 95% konfidensinterval herfor.

1. *Repetition af centrale begreber*

Opgave 1.4

Indenfor en stor virksomhed, der producerer udstyr til forsvaret, er der i middel 20 driftsuheld pr. måned. Da antallet efter indførelsen af nye arbejdsrutiner synes at være vokset målte man i 5 på hinanden følgende måneder antallet af driftsuheld. Resultaterne var

måned nr.	1	2	3	4	5
antal/måned	23	19	23	27	24

Test, om disse data giver et eksperimentelt bevis for, at middelværdien μ er større end 20 driftsuheld/måned?

2 STATISTISK PROCESKONTROL

2.1. INDLEDNING

Formålet med det meget omfattende begreb kvalitetsstyring (også kaldet kvalitetskontrol¹) er at kontrollere, styre og forbedre kvaliteten af et produkt. Endvidere at nedsætte virksomhedens samlede styrings- og fejlomkostninger. Forskellige undersøgelser antyder, at mange virksomheder ved overgang til kvalitetsstyring kan nedsætte deres samlede kvalitetsomkostninger fra størrelsesordenen 10-15% af omsætningen til 3-5% af denne.

Der er på international basis opstillet krav, som skal opfyldes for opnåelse af en "kvalitetscertificering" af virksomheders kvalitetsstyring, bl.a. anført i serien af "ISO 9000-standarder". Virksomheder, som opfylder de pågældende krav, kan opnå et såkaldt "ISO 9000 - certifikat".

Vi begrænser os i nærværende kapitel til omtale af **statistisk proceskontrol** (SPC = Statistical Process Control) i egentligt materialeproducerende virksomheder.

Formålet med denne er at styre produktionsprocesserne således, at fejlproduktion forebygges.

En industriel produktionsproces kan formelt betragtes som en talfrembringende proces.

Eksempler herpå er angivet i det følgende skema:

nr	Produktion af	Frembragte tal
1	Kunstgødning	Kvælstofindhold
2	Jernbjælker	Trykstyrke , Siliciumindhold
3	TV-apparater	Antal loddefejl pr. apparat
4	Patroner	Antal defekte patroner i en stikprøve
5	Leverpostej	0: Dårlig smag , 1: Mindre god smag, 2: God smag , 3: Særdeles god smag
6	Vin	Alkoholprocent
7	Flasker	Rumindhold
8	Film	0: defekt, 1: ikke defekt
9	Tekstil	Antal fejl pr. m ²
10	Aksler	Diameter

¹ På engelsk kaldet "Quality Control" eller "Quality Management".

2.2 PROCES I STATISTISK KONTROL.

Ved enhver proces vil de frembragte tal udvise en "naturlig" variation, uanset hvor godt processen er planlagt og hvor omhyggeligt den bringes til udførelse og vedligeholdes. Denne "naturlige" variation der i praksis ikke kan kontrolleres, er et resultat af talrige små påvirkninger/variationsårsager ("common causes").

Hvis en proces **kun** er påvirket af tilfældige variationsårsager, siges den at være i **statistisk kontrol**.

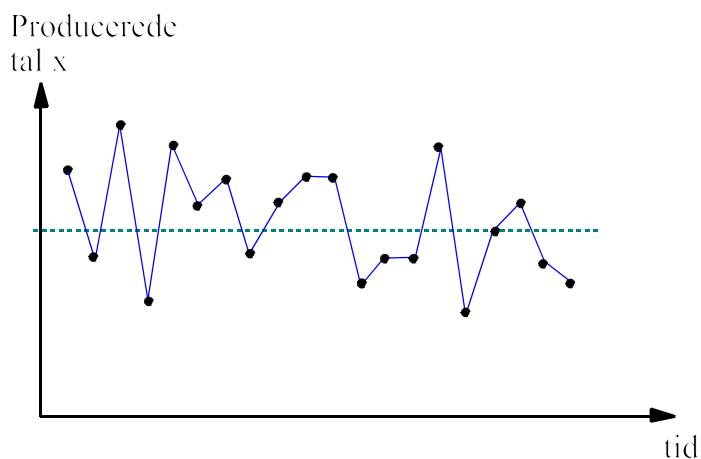


Fig 2.1. *Proces i statistisk kontrol*

De producerede tal er uafhængige observationer fra en population med en bestemt sandsynlighedsfordeling. Den dertil svarende statistiske variabel kaldes procesvariablen. Er en proces i statistisk kontrol kan man derfor på basis af en stikprøve estimere dens parametre, og kan eksempelvis beregne størrelsen af en kommende fejlproduktion.

Proces ude af statistisk kontrol. Udover den tilfældige variation kan der forekomme en variation, som kan tilskrives bestemte årsager, f.eks. maskiner, operatører eller råvarer; maskiner kan være indstillet forkert, en operatør kan være træt eller uopmærksom, et råvareparti kan være af dårlig kvalitet. Man taler i disse situationer om væsentlige, specielle eller påviselige variationsårsager ("assignable causes")

Hvis en proces i løbet af et givet tidsrum går fra at være i statistisk kontrol i en procestilstand til at være i statistisk kontrol i en anden procestilstand, har processen været **ude af statistisk kontrol** indenfor det pågældende tidsrum.

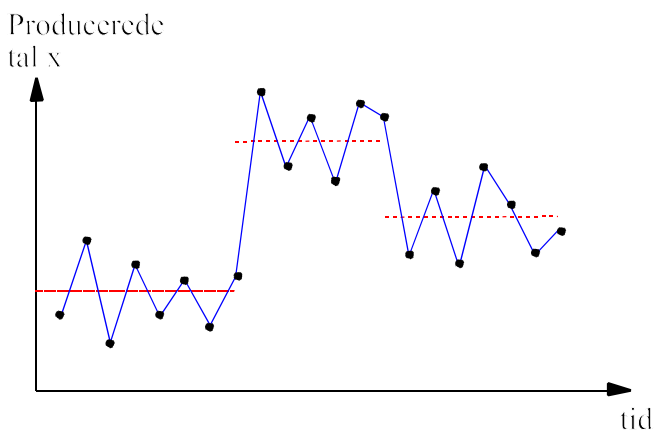


Fig.2.2. *Proces ude af statistisk kontrol*

2.3 OPBYGNING OG ALARMGRÆNSER FOR KONTROLKORT .

Kontrolkort (også kaldet Shewart kontrolkort efter opfinderen) er baseret på, at man af den løbende produktion med regelmæssige mellemrum udtager stikprøver af størrelsen n .

Figur 2.3 viser en normalfordeling sammen med et enkelt kontrolkort.

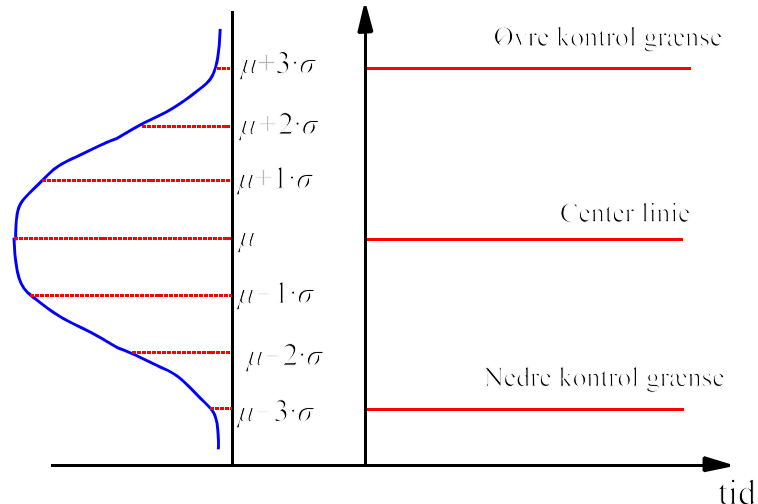


Fig 2.3. Kontrolkort med øvre og nedre kontrolgrænser

Hvis en på basis af stikprøven beregnet parameter eksempelvis gennemsnittet \bar{x} holder sig indenfor nogle kontrolgrænser, antager man at processen er i kontrol med hensyn til denne parameter. Hvis den beregnede størrelse ligger udenfor kontrolgrænserne vil man “slå alarm” fordi så kan det tænkes at processen er kommet ud af kontrol.

Stikprøvernes størrelse n .

Generelt gælder det, at det er bedre at tage små stikprøver på 4-5 emner ud med korte mellemrum end at tage store stikprøver på 20-25 emner ud med lange mellemrum. Formålet er jo, at man ønsker hurtigt at opdage, hvis processen er ved at komme ud af kontrol (fordi eksempelvis en maskine er ved at gå i stykker). Bruger man store stikprøver vil man kunne opdage selv forholdsvis små forskydninger i niveauet, men på grund af at der går temmelig lang tid mellem man tager stikprøverne, kan en stor katastrofal forskydning blive opdaget for sent.

2.4. KONTROLKORTANALYSE.

Anvendelsen af kontrolkort kræver, at processen fra starten er i statistisk kontrol. Indførelsen af proceskontrol i en “ny” produktion kræver derfor, at man ved en såkaldt kontrolkortanalyse får undersøgt og om nødvendigt justeret processen således, at den kommer i statistisk kontrol. Først når dette er tilfældet, kan man estimere de ukendte parametre og konstruere kontrolkortet.

Ved en kontrolkortanalyse foretages en række målinger af procesvariablen. Samtidig med en måling registreres en række oplysninger, f.eks. observationstidspunkt, temperatur, råmaterialeparti, arbejdshold m.v., samt hvilke personer der har foretaget målingerne. På grundlag heraf inddeles måltallene i “rationelle undergrupper”, inden for hvilke de formodes at være produceret under samme væsentlige betingelser. Endvidere tilstræber man, at antal målinger er det samme i hver undergruppe.

Indsamling og opdeling af et observationsmateriale i rationelle undergrupper kræver ofte en betydelig teknisk indsigt i den betragtede produktionsproces. Alle de i praksis uundgåelige

variationsårsager skal bidrage til den faktiske variation indenfor undergrupper, hvorimod alle de variationsårsager, som man mener kan være væsentlige variationsårsager, ikke må bidrage til variationen indenfor undergrupper.

De variationsårsager, som måske er væsentlige, må altså kun indvirke fra undergruppe til undergruppe. Mener man således, at tal produceret af forskellige maskiner kan være en væsentlig variationsårsag, så må disse tal ikke være placeret i samme undergruppe.

På basis af undergrupperne konstrueres så kontrolkort som beskrevet i eksempel 2.1. Falder en undergruppe udenfor de beregnede kontrolgrænser, så må man undersøge nærmere hvad årsagen kan være. Er det eksempelvis en gruppe hvis tal er produceret af en bestemt maskine, så må man ofre en hovedreparation på maskinen eller kassere den. På den måde får man ikke alene processen i kontrol, men man øger også kvaliteten af den igangværende proces.

Der udarbejdes forskellige typer kontrolkort afhængig af sandsynlighedsfordelingen for procesvariablen X .

I afsnit 2.6 behandles tilfældet hvor X approksimativt er en kontinuert normalfordelt variabel (jævnfør tilfældene 1, 2, 6, 7 og 10 i skemaet på side 2), og størrelsen af stikprøverne er større end 1. I afsnit 2.7 ses på den situation hvor stikprøvestørrelsen er 1, hvilket eksempelvis kan ske hvis det tager så lang tid at analysere et enkelt element. I afsnit 2.8 ser på de tilfælde, hvor X approksimativt er binomialfordelt eller Poissonfordelt.

Det er i alle tilfælde væsentligt, at observationerne er uafhængige, mens mindre afvigelser fra den forventede sandsynlighedsfordeling ikke vil give nogen væsentlig forøgelse i falske alarmer.

2.5 PROCESVARIABLEN ER NORMALFORDELT

2.5.1 Stikprøvestørrelse større end 1

Lad os antage at processen er i kontrol med en middelværdi på μ og en spredning på σ . Vi kender ikke de eksakte værdier, men ønsker at beregne estimater herfor.

For at kunne beregne kontrolgrænserne for \bar{x} -kortet: $\bar{x} \pm 3 \frac{\tilde{\sigma}}{\sqrt{n}}$ må man kende et estimat for spredningen σ . Man starter derfor altid med at lave et kontrolkort for spredningen. Man udtager i produktionen løbende ("af folk på gulvet") stikprøver. Da det tidligere var besværligt at beregne estimatet s , foretrak man ofte, at benytte variationsbredderne R_i som et mål for spredningen. Derfor vil man stadig møde mange "R-kontrolkort" selv om "s-kontrolkort" er blevet mere almindelige.

Sædvanligvis vil man benytte et statistikprogram som SAS.JMP til beregning af kontrolgrænser m.m., men som eksempel 2.1 viser, kan man godt beregne dem ved "håndkraft" ved hjælp af de formler og konstanter, der findes i nedenstående 2 tabeller.

En nærmere forklaring på formlerne findes i appendix 2.1

Tabel 2.1: Kontrollkort.

$\bar{x} - R$ - kontrolkort. Procesvariablen X er normalfordelt $n(\mu, \sigma)$.					
Procestilstand		Centerlinie	Nedre kontrolgrænse	Øvre kontrolgrænse	Estimer
Kendt:	\bar{x} - kort R - kort	μ $d_2 \cdot \sigma$	$\mu - A_1 \cdot \sigma$ $D_1 \cdot \sigma$	$\mu + A_1 \cdot \sigma$ $D_2 \cdot \sigma$	
Ukendt:	\bar{x} - kort R - kort	$\bar{\bar{x}}$ \bar{R}	$\bar{\bar{x}} - A_2 \cdot \bar{R}$ $D_3 \cdot \bar{R}$	$\bar{\bar{x}} + A_2 \cdot \bar{R}$ $D_4 \cdot \bar{R}$	$\hat{\sigma} \approx \frac{\bar{R}}{d_2}$
$\bar{x} - s$ - kontrolkort. Procesvariablen X er normalfordelt $n(\mu, \sigma)$.					
Kendt	\bar{x} - kort s - kort	μ $c_4 \cdot \sigma$	$\mu - A_1 \cdot \sigma$ $B_5 \cdot \sigma$	$\mu + A_1 \cdot \sigma$ $B_6 \cdot \sigma$	
Ukendt	\bar{x} - kort s - kort	$\bar{\bar{x}}$ \bar{s}	$\bar{\bar{x}} - A_3 \cdot \bar{s}$ $B_3 \cdot \bar{s}$	$\bar{\bar{x}} + A_3 \cdot \bar{s}$ $B_4 \cdot \bar{s}$	$\hat{\sigma} \approx \frac{\bar{s}}{c_4}$

Tabel 2.2. Konstanter til fastlæggelse af kontrolgrænser ved hjælp af \bar{x} - kort, R - kort og s -kort.

n	\bar{x} - kort			R - kort						s - kort				
	A_1	A_2	A_3	d_2	d_3	D_1	D_2	D_3	D_4	c_4	B_3	B_4	B_5	B_6
2	2.121	1.880	2.659	1.128	0.853	0	3.686	0	3.267	0.798	0	3.267	0	2.606
3	1.732	1.023	1.954	1.693	0.888	0	4.358	0	2.575	0.886	0	2.568	0	2.276
4	1.500	0.729	1.628	2.059	0.880	0	4.698	0	2.282	0.921	0	2.266	0	2.088
5	1.342	0.577	1.427	2.326	0.864	0	4.918	0	2.115	0.940	0	2.089	0	1.964
6	1.225	0.483	1.287	2.534	0.848	0	5.078	0	2.004	0.952	0.030	1.970	0.029	1.874
7	1.134	0.419	1.182	2.704	0.833	0.205	5.203	0.076	1.924	0.959	0.118	1.882	0.113	1.806
8	1.061	0.373	1.099	2.847	0.820	0.387	5.307	0.136	1.864	0.965	0.185	1.815	0.179	1.751
9	1.000	0.337	1.032	2.970	0.808	0.546	5.394	0.184	1.816	0.969	0.239	1.761	0.232	1.707
10	0.949	0.308	0.975	3.078	0.797	0.687	5.469	0.223	1.777	0.973	0.284	1.716	0.276	1.669
11	0.905	0.285	0.927	3.173	0.787	0.812	5.534	0.256	1.744	0.975	0.321	1.679	0.313	1.637
12	0.866	0.266	0.886	3.258	0.778	0.924	5.592	0.284	1.716	0.978	0.354	1.646	0.346	1.610
13	0.832	0.249	0.850	3.336	0.770	1.026	5.646	0.308	1.692	0.979	0.382	1.618	0.374	1.585
14	0.802	0.235	0.817	3.407	0.762	1.121	5.693	0.329	1.671	0.981	0.406	1.594	0.399	1.563
15	0.775	0.223	0.789	3.472	0.755	1.207	5.737	0.348	1.652	0.982	0.428	1.572	0.421	1.544
16	0.750	0.212	0.763	3.532	0.749	1.285	5.779	0.364	1.636	0.983	0.448	1.552	0.440	1.526
17	0.728	0.203	0.739	3.588	0.743	1.359	5.817	0.379	1.621	0.984	0.466	1.534	0.458	1.511
18	0.707	0.194	0.718	3.640	0.738	1.426	5.854	0.392	1.608	0.985	0.482	1.518	0.475	1.496
19	0.688	0.187	0.698	3.689	0.733	1.490	5.888	0.404	1.596	0.986	0.497	1.503	0.490	1.483
20	0.671	0.180	0.680	3.735	0.729	1.548	5.922	0.414	1.586	0.987	0.510	1.489	0.504	1.470

Eksempel 2.1. Kontrol af stof i levnedsmiddelprodukt.

En levnedsmiddelvirkomhed har problemer med at holde koncentrationen af et bestemt stof A i et konservesprodukt mellem nogle tolerancegrænser. Man vælger derfor at få foretaget en kontrollkortanalyse. På basis af tidligere erfaringer inddeles målingerne i 30 undergrupper, som hver har deres karakteristika:(råvarecharge, apparatur, tidspunkt på dagen osv.). For hver undergruppe (som er på 5 målinger) er der af hensyn til de følgende beregninger også beregnet gennemsnit \bar{x}_i , variationsbredde R_i og spredning s_i .

Gruppe	Målinger	\bar{x}_i	R_i	s_i	Gruppe	Målinger	\bar{x}_i	R_i	s_i
1	13 8 2 5 8	7.2	11	4.0866	16	16 11 14 8 17	13.2	9	3.7014
2	0 6 1 9 15	6.2	15	6.1400	17	9 4 4 8 9	6.8	5	2.5884
3	4 2 4 3 4	3.4	2	0.8944	18	6 1 1 3 13	4.8	12	5.0200
4	3 15 8 3 5	6.8	12	5.0200	19	7 0 5 7 2	4.2	7	3.1145
5	5 10 5 4 0	4.8	10	3.5637	20	10 0 10 12 7	7.8	12	4.7117
6	9 5 13 7 7	8.2	8	3.0332	21	3 7 5 10 12	7.4	9	3.6469
7	0 4 4 3 9	4.0	9	3.2404	22	3 0 10 5 4	4.4	10	3.6469
8	9 3 0 6 0	3.6	9	3.9115	23	3 3 0 6 9	4.2	9	3.4205
9	14 0 0 5 3	4.4	14	5.7706	24	0 2 3 6 7	3.6	7	2.8810
10	3 9 5 0 2	3.8	9	3.4205	25	2 3 5 4 10	4.8	8	3.1145
11	5 8 0 7 8	5.6	8	3.3616	26	3 1 4 2 4	2.8	3	1.3038
12	3 2 2 7 4	3.6	5	2.0736	27	2 4 5 13 4	5.6	11	4.2778
13	5 11 14 8 3	8.2	11	4.4385	28	0 22 7 2 11	8.4	22	8.7350
14	13 5 5 12 7	8.4	8	3.8471	29	3 5 9 8 6	6.2	6	2.3875
15	7 0 1 0 6	2.8	7	3.4205	30	9 7 10 13 0	7.8	13	4.8683
SUM							173.0	281	113.64

Foretag ved hjælp af \bar{x} - og R - kort en kontrollkortanalyse og opstil kontrollkort, der kan benyttes til en løbende kontrol af indholdet af det undersøgte stof.

Løsning .

1) Først foretages en R - kort analyse.

Idet $\bar{R} = \frac{281}{30} = 9.366$ er grænserne for R - kortet ifølge tabel 5.1 (og tabel 5.2):

$$NKG_R = D_3 \cdot \bar{R} = 0 \cdot 9.366 = 0 \quad \text{og} \quad \text{ØKG}_R = D_4 \cdot \bar{R} = 2.115 \cdot 9.366 = 19.809.$$

Det ses, at alle grupper på nær gruppe 28 falder indenfor grænserne. Man foretager nu en nærmere undersøgelse af hvilke forhold i gruppe 28 der kan tænkes at bevirke dette. Hvis det eksempelvis skyldes et bestemt apparat, kan man kassere dette eller reparere det. Vi vil i det følgende udskyde gruppen, revidere grænserne og se om de resterende grupper nu falder indenfor de nye grænser.

Vi får $\bar{R}^* = \frac{281 - 22}{29} = 8.931$ og dermed

$$NKG_R = D_3 \cdot \bar{R}^* = 0 \cdot 8.931 = 0 \quad \text{og} \quad \text{ØKG}_R = D_4 \cdot \bar{R}^* = 2.115 \cdot 8.931 = 18.889.$$

Nu falder alle grupper indenfor kontrolgrænserne, og vi konkluderer derfor at processen er i kontrol med hensyn til spredningen, og at denne estimeres ved $\tilde{\sigma} = \frac{\bar{R}^*}{d_2} = \frac{8.931}{2.326} = 3.840$.

Dernæst foretages en \bar{x} - kort analyse.

$\bar{x}^* = \frac{173 - 8.4}{29} = 5.676$, og dermed

$$NKG_{\bar{x}} = \bar{x}^* - A_2 \cdot \bar{R}^* = 5.676 - 0.577 \cdot 8.931 = 0.5228$$

$$\emptyset KG_{\bar{x}} = \bar{x}^* + A_2 \cdot \bar{R}^* = 5.676 + 0.577 \cdot 8.931 = 10.829$$

Det ses, at alle grupper på nær gruppe 16 falder indenfor grænserne.

Man foretager nu en nærmere undersøgelse af hvilke forhold i gruppe 16 der kan tænkes at bevirke dette. Vi vil i det følgende udskyde gruppen, revidere grænserne og se om de resterende grupper nu falder indenfor de nye grænser.

$$\bar{x}^{**} = \frac{173 - 8.4 - 13.2}{28} = 5.407.$$

Strengt taget burde vi også revidere R - kortet, men da en udskydelse af punkt 16 kun vil give en ubetydelig ændring af R - grænserne beholdes disse.

$$NKG_{\bar{x}} = \bar{x}^{**} - A_2 \cdot \bar{R}^* = 5.407 - 0.577 \cdot 8.931 = 0.2538$$

$$\emptyset KG_{\bar{x}} = \bar{x}^{**} + A_2 \cdot \bar{R}^* = 5.407 + 0.577 \cdot 8.931 = 10.560$$

Vi antager herefter at processen også er i kontrol med hensyn til middelværdien.

De to kort med de reviderede grænser kan nu benyttes til den løbende proceskontrol. ◆

Eksempel 2.2. Kontrol af stof i levnedsmiddelproduktion.

Samme spørgsmål som i eksempel 2.1, men udarbejd i stedet for R - kortet et s - kort.

Løsning:

Først foretages en s - kort analyse.

Idet $\bar{s} = \frac{113.64}{30} = 3.7880$ er grænserne for s - kortet:

$$NKG_s = B_3 \cdot \bar{s} = 0 \cdot 3.7880 = 0 \text{ og } \emptyset KG_s = B_4 \cdot \bar{s} = 2.089 \cdot 3.7880 = 7.9132.$$

Det ses, at alle grupper på nær gruppe 28 falder indenfor grænserne. Denne gruppe udskydes og grænserne revideres. Vi får $\bar{s}^* = \frac{113.64 - 8.735}{29} = 3.6174$ og dermed

$$NKG_s = B_3 \cdot \bar{s}^* = 0 \cdot 3.6174 = 0 \text{ og } \emptyset KG_s = B_4 \cdot \bar{s}^* = 2.089 \cdot 3.6174 = 7.5568.$$

Nu falder alle grupper indenfor kontrolgrænserne, og vi konkluderer derfor at processen er i kontrol med hensyn til spredningen, og at denne estimeres ved $\tilde{\sigma} = \frac{\bar{s}^*}{c_4} = \frac{3.6174}{0.940} = 3.8483$.

Dernæst foretages en \bar{x} - kort analyse. $\bar{x}^* = \frac{173 - 8.4}{29} = 5.676$, og dermed

$$NKG_{\bar{x}} = \bar{x}^* - A_3 \cdot \bar{s}^* = 5.676 - 1.427 \cdot 3.6174 = 0.5140$$

$$\emptyset KG_{\bar{x}} = \bar{x}^* + A_3 \cdot \bar{s}^* = 5.676 + 1.427 \cdot 3.6174 = 10.838.$$

Det ses, at alle grupper på nær gruppe 16 falder indenfor grænserne.

Denne gruppe udskydes og grænserne revideres: $\bar{x}^{**} = \frac{173 - 8.4 - 13.2}{28} = 5.407$.

Strengt taget burde vi også revidere s - kortet, men da en udskydelse af punkt 16 kun vil give en ubetydelig ændring af R - grænserne beholdes disse.

$$NKG_{\bar{x}} = \bar{x}^{**} - A_3 \cdot \bar{s}^* = 5.407 - 1.427 \cdot 3.6174 = 0.2450$$

$$\emptyset KG_{\bar{x}} = \bar{x}^{**} + A_3 \cdot \bar{s}^* = 5.407 + 1.427 \cdot 3.6174 = 10.569$$

Vi antager herefter at processen også er i kontrol med hensyn til middelværdien.

De to kort med de reviderede grænser kan nu benyttes til den løbende proceskontrol. ◆

2.5.2 Stikprøvestørrelse $n = 1$

I visse situationer er det praktisk umuligt at udtage mere end et element i stikprøven. Det kan skyldes, at det tager så lang tid at få de nødvendige data, at det er urealistisk at undersøge mere end 1 element.

For at få et skøn for spredningen bruger man R-værdierne for 2 på hinanden følgende observationer (moving range) $MR_x = |x_i - x_{i-1}|$

Derfor kaldes kontrollkortet for spredningen for et "Moving range control chart"

Kontrollkortet er bestemt ved de formler der fremgår af følgende eksempel.

Imidlertid er det naturligvis vigtigt, at fordelingen er rimelig normalfordelt, så man vil sædvanligvis først tegne et normalfordelingsplot.

Eksempel 2.3 Kontrollkort for enkeltobservationer

En olieraffineringsproces forløber så langsomt, at man må benytte en stikprøvestørrelse på $n = 1$.

Følgende observationer af oktantallet x for et bestemt produkt fandtes.

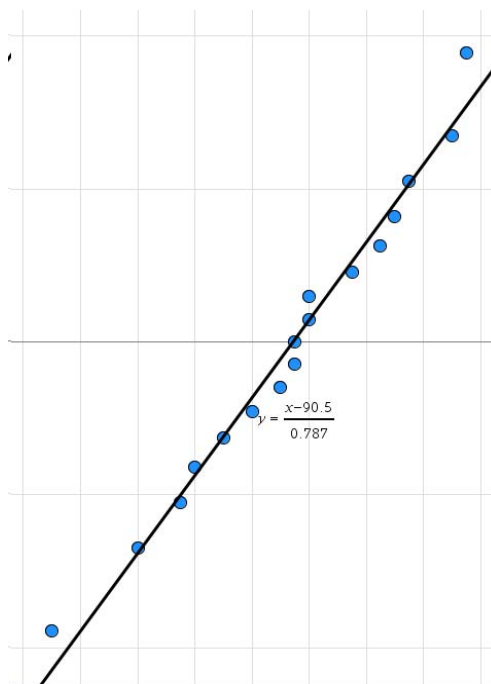
x	90.2	90.6	89.4	91.3	91.2	89.7	90.6	88.8	90.5	90.9	90.0	91.1	89.8	90.5	90.5	91.6	90.4	91.7
---	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

- 1) Vurder først om tallene er rimelig normalfordelt ved at tegne et normalfordelingsplot.
- 2) I det følgende antages, at fordelingen er rimelig normalfordelt.

Lav et "Moving Range kontrollkort" for processen.

Løsning.

- 1) Plottet er tegnet i TI-Nspire (se evt eksempel 1.1 hvordan)
Da punkterne ligger nogenlunde på den rette linie antages tallene at være rimelig normalfordelt.
- 2) Vi udregner nu MR-værdierne



x	90.2	90.6	89.4	91.3	91.2	89.7	90.6	88.8	90.5	90.9	90.0	91.1	89.8	90.5	90.5	91.6	90.4	91.7
MR	0.4	1.2	1.9	0.1	1.5	0.9	1.8	1.7	0.4	0.9	1.1	1.3	0.7	0.0	1.1	1.2	1.3	

Vi finder $\bar{x} = 90.49$ og $\overline{MR} = 1.029$

Der gælder nu, at et estimat for spredningen er $\tilde{\sigma} = \frac{\overline{MR}}{d_2} = \frac{1.029}{1.128} = 0.912$

d_2 er tallet i tabel 2.2 svarende til $n = 2$.

Vi får derfor for R-kortet

$$NKG = D_3 \cdot \overline{MR} = 0 \quad \text{ØKG} = D_4 \cdot \overline{MR} = 3.267 \cdot 1.029 = 3.356$$

Det ses, at alle R-værdier falder indenfor kontrolgrænserne.

$$\text{For } \bar{x} \text{ kort gælder } NKG = \bar{x} - 3 \cdot \frac{MR}{d_2} = 90.49 - 3 \cdot 0.912 = 87.75$$

$$\text{ØKG} = \bar{x} + 3 \cdot \frac{MR}{d_2} = 90.49 + 3 \cdot 0.912 = 93.23$$

Vi ser, at alle \bar{x} -værdier falder indenfor kontrolgrænserne. ◆

Behandlingen af disse kort svarer til dem vi anvendte ved de øvrige kontrolkort.

Det skal dog nævnes, at disse kort har svært ved at opdage små ændringer i processens middelværdi. Den er også meget følsom overfor selv mindre afvigelser fra normalfordelingen.

Derfor anvendes ofte i stedet enten såkaldte "Cusum" kort eller "EWMA" kort.

2.6 Løbende kontrol

Når man har lavet kontrolkortet skal det jo anvendes til at holde kontrol med at produktionen ikke af en eller anden grund begynder at ændre sig.

"Alarmkriterier".

Man finder der er grund til at tro at processen er ude af kontrol (at slå alarm), hvis der sker et af følgende (se også de følgende figurer).

- 1) 1 punkt (mindst) udenfor $3 \cdot \sigma$ - kontrolgrænserne
- 2) 9 på hinanden følgende punkter ligger over /under en $1 \cdot \sigma$ - grænse.
- 3) 6 på hinanden følgende punkter er konstant stigende/ faldende.
- 4) 14 på hinanden følgende punkter skiftevis stiger og falder i værdi.
- 5) 2 af 3 på hinanden følgende punkter falder udenfor en $2 \cdot \sigma$ - grænse
- 6) 4 ud af 5 på hinanden følgende punkter falder udenfor en $1 \cdot \sigma$ - grænse.
- 7) 15 på hinanden følgende punkter ligger indenfor $1 \cdot \sigma$ - grænse
- 8) 8 på hinanden følgende punkter ligger på begge sider af centerlinien, men ingen indenfor $1 \cdot \sigma$ - grænse.

De følgende figurer illustrerer nogle af disse grænser

Fig 2.4. Mindst 1 punkt udenfor $3 \cdot \sigma$ - kontrolgrænser

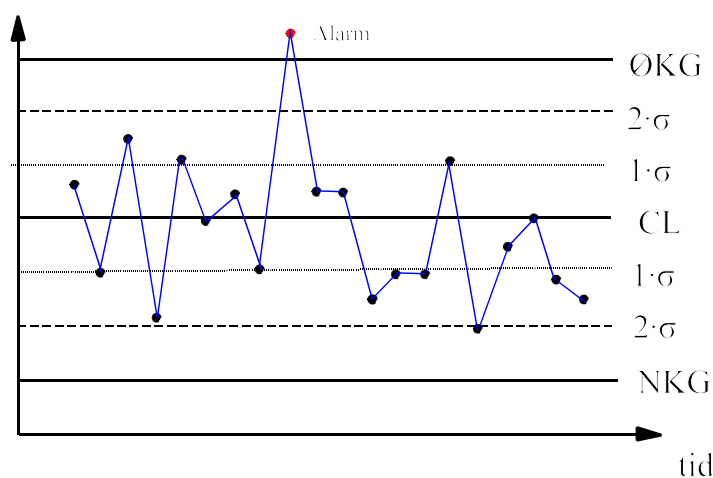
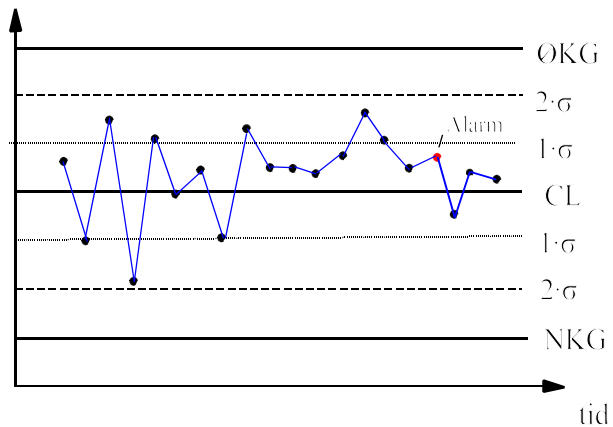


Fig 2.5.

9 på hinanden følgende punkter over centerlinien eller 9 på hinanden følgende punkter under centerlinien



Figur 2.6

6 på hinanden følgende punkter stiger i værdi eller falder i værdi.

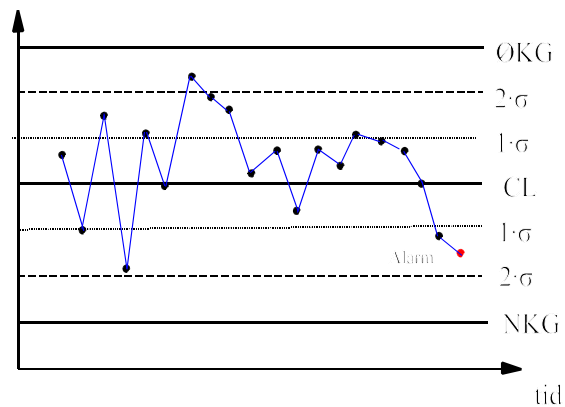


Fig 2.7. Mindst 2 punkter ud af 3 på hinanden følgende punkter udenfor en $2\cdot\sigma$ - grænse

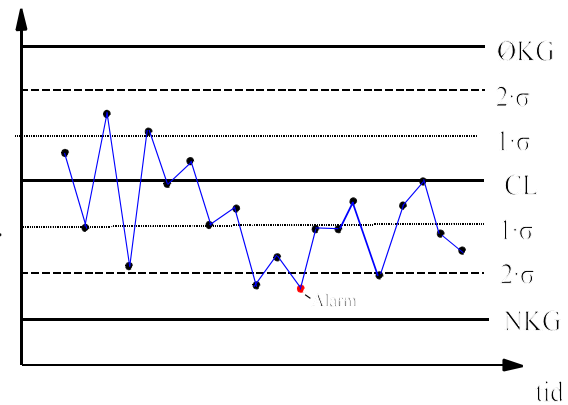
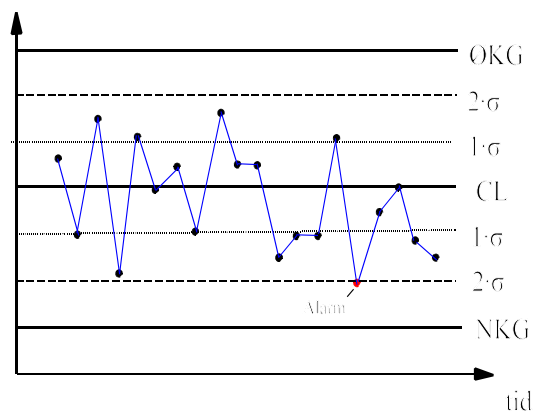


Fig 2.8. Mindst 4 punkter ud af 5 på hinanden følgende punkter falder udenfor en $1\cdot\sigma$ - grænse.

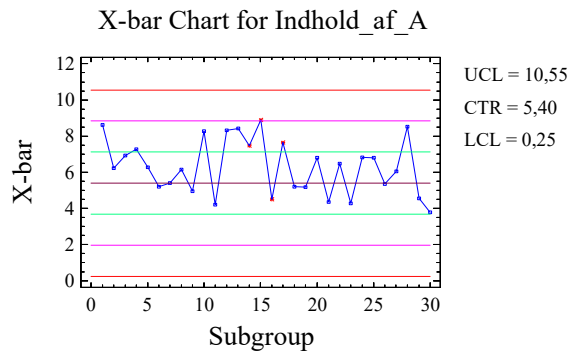
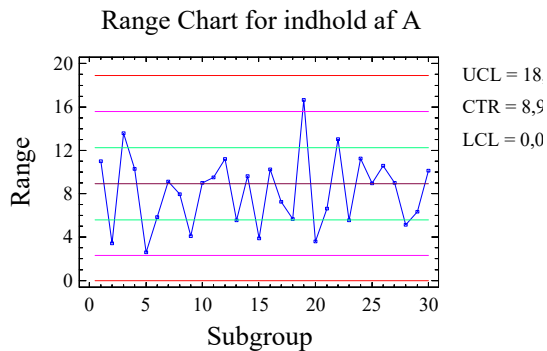


Det følgende eksempel viser hvordan.

Eksempel 2.4. Løbende kontrol.

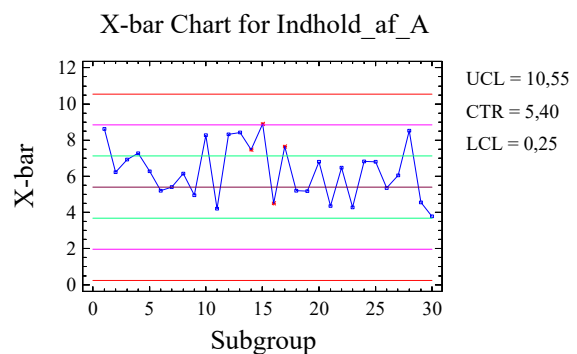
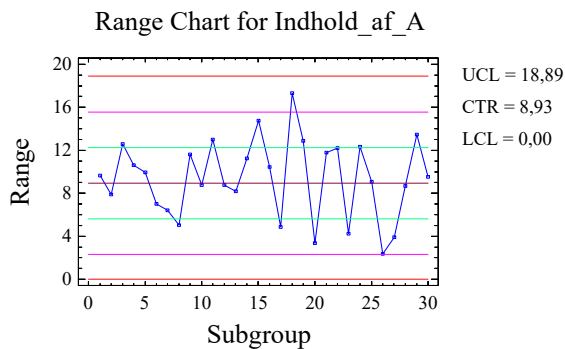
De i eksempel 2.1 fundne kontrollkort er i de følgende dage blevet benyttet til løbende kontrol af processen.

1) I de første 30 dage gav det følgende resultat:



Er processen stadig i kontrol?

2) Man har nu en mistanke om, at koncentrationerne af A har ændret sig. Kan dette bekræftes af kontrollkortene for de følgende 30 dage?



Løsning

1) Det ses, at processen er i kontrol med hensyn til R - kortet, da punkterne fordeler sig tilfældigt omkring centerlinien, og ingen af "alarmkriterierne" er overtrådt.

For \bar{x} - kortets vedkommende er alarmregel 2 (9 punkter i træk over centerlinie) tæt ved at være opfyldt, men da dag 17 lige er placeret på centerlinien, så anses processen også her at være i kontrol

2) Det ses, at for R - kortets vedkommende er alarmregel 3 opfyldt (6 på hinanden følgende punkter nemlig dagene 3,4,5,6,7,8 falder i værdi

For \bar{x} - kortets vedkommende er alarmregel 6 (mindst 4 punkter ud af 5 falder udenfor en 1σ grænse nemlig dagene 12,13,14,15,16,17).

Vi må derfor konkludere, at der er grund til at formode, at en nøjere undersøgelse er påkrævet.

2.7 TOLERANCEGRÆNSER OG KAPABILITET

I forbindelse med en fabrikation er der ofte fastsat "tolerancegrænser" eller "specifikationsgrænser (jævnfør eksempel 2.1).

Det kan enten være en nedre tolerancegrænse NTG_x og/eller en øvre tolerancegrænse $ØTG_x$, idet man ved fabrikationen ønsker/forlanger, for procesvariablen X , at $X \geq NTG_x$ henholdsvis $X \leq ØTG_x$.

Produktion, for hvilken X falder udenfor tolerancegrænsen/tolerancegrænserne, betragtes altså som fejlproduktion.

Bemærk, at der her ikke er tale om stikprøvens gennemsnit, men om den enkelte værdi af procesvariablen.

I det følgende forudsætter vi, at der er fastsat såvel en nedre som en øvre tolerancegrænse.

Ved "**kapabiliteten**" af en proces forstås processens evne til at producere inden for et specificeret toleranceinterval.

Som omtalt, er "det naturlige variationsområde" for en proces en variation på 6σ , nemlig fra -3σ til $+3\sigma$.

Når der ved en given fabrikation er fastsat et toleranceinterval $[NTG_x; ØTG_x]$ er det derfor nærliggende at sammenligne dette med intervallet for den naturlige variation $[-3\sigma; 3\sigma]$.

Hertil udregnes et **kapabilitetsindeks** $C_p = \frac{ØTG_x - NTG_x}{6\sigma}$

(PCR = proces capability ratio = C_p)

C_p er et mål for processens evne til produktion indenfor toleranceintervallet.

Der foreligger herved i en kontrolsituation en af flere muligheder:

Hvis $C_p \approx 1$ dvs $ØTG_x - NTG_x \approx 6\sigma$ skulle man tro, at dette ville være en proces med lille fejlproduktion, da højst 0.27% af produktionen vil falde udenfor.

Imidlertid er det med en stor løbende produktion svært hele tiden at holde middelværdien midt mellem de to tolerancegrænser.

Endvidere er fordelingen muligvis ikke helt normalfordelt, og endvidere kendes spredningen jo ikke eksakt men man udregner på basis af nogle målinger et estimat s .

Man vil derfor sædvanligvis først sige, at man har en produktion med en lille fejlproduktion, hvis $C_p > 1.33$. Det afhænger naturligvis af produktionen og hvilke krav man stiller. hvor man sætter grænsen, og i den såkaldte "6 sigma proces" kræves et $C_p \geq 2 \Leftrightarrow ØTG_x - NTG_x \geq 12\sigma$

1) C_p stor (større end eksempelvis 1.33 eller 2)

I dette tilfælde kan fejlproduktion (næsten) helt undgås; en processtyring er fortsat anbefalelsesværdig, men kontrollen kan foretages relativt afslappet

2) C_p lille (under 1.33 eller 2)

I dette tilfælde udøves sædvanlig processtyring, f.eks. med anvendelse af \bar{x} - kort og R - kort. En (for stor) fejlproduktion kan ikke undgås, men den kan søges minimeret ved stram styring af processen, hvorved procesniveauet søges fastholdt på midterværdien $\frac{ØTG_x - NTG_x}{2}$.

Eventuelt foretages en totalinspektion; eventuelt gennemføres en produktionsforbedring, hvorved processens spredning formindskes; eventuelt aftales nye tolerancegrænser og/eller

en ny pris for det fremstillede produkt osv.

Tallet $P = \frac{100}{C_p}$ angiver hvor stor en procentdel af toleranceområdet der bruges af processen.

Er eksempelvis $C_p = 1$ bruger processen 100% af toleranceintervallet fra NTG til ØTG.

Som et mål for den aktuelle kapabilitet indføres et centreringsindeks (performanceindeks)

$$C_{pk} = \min\left(\frac{\text{ØTG}_x - \bar{x}}{3\sigma}, \frac{\bar{x} - \text{NTG}_x}{3\sigma}\right)$$

Hvis $C_p \approx C_{pk}$ er processen meget godt centreret, mens $C_{pk} < C_p$ viser, at dette ikke er tilfældet.

Det kan ofte være nyttigt at beregne et konfidensinterval for C_p

$$\text{Konfidensinterval for } C_p: \quad \tilde{C}_p \cdot \sqrt{\frac{\chi_{\alpha/2, n-1}^2}{n-1}} \leq C_p \leq \tilde{C}_p \cdot \sqrt{\frac{\chi_{1-\alpha/2, n-1}^2}{n-1}}$$

Der er andre forslag til kapabilitetsindeks, som er mere følsom over for den situation, at middelværdien μ for processen afviger fra midtpunktet T af tolerancegrænserne.

$$\text{En af disse er } C_{pkm} = \frac{C_{pk}}{\sqrt{1 + \left(\frac{\mu - T}{\sigma}\right)^2}}$$

Ensidede kapabilitetsindeks

Hvis der kun er en øvre eller en nedre tolerancegrænse anvendes i stedet

$$C_{pu} = \frac{\text{ØTG} - \mu}{3 \cdot \sigma} \text{ hvis der kun er angivet en øvre tolerancegrænse.}$$

$$C_{pl} = \frac{\mu - \text{NTG}}{3 \cdot \sigma} \text{ hvis der kun er angivet en nedre tolerancegrænse.}$$

Eksempel 2.5. Beregning af kapabilitetsindeks

En proces var i statistisk kontrol med en estimeret processtilstand med middelværdi $\tilde{\mu} = 7.407$ og spredningen $\tilde{\sigma} = 2.840$

De udarbejdede kontrolkort kan derfor benyttes til den løbende kontrol.

Der bliver nu fastlagt følgende tolerancegrænser $\text{ØTG} = 24$, $\text{NTG} = 2$

1) Beregn kapabilitetsindeks C_p

Forklar hvorfor dette tal ikke giver et sandt billede af processens evne til holde sig indenfor tolerancegrænserne.

2) Hvilken fejlprocent vil fabrikationen have?

3) Hvilket middelniveau bør processen ideelt søges indstillet og fastholdt på.

Løsning:

$$1) C_p = \frac{24 - 2}{6 \cdot 2.84} = 1.29$$

Selv om værdien er over 1 er der uden tvivl en ret stor fejlprocent, da estimatet for middelværdien 7.407 ligger ret langt gennemsnittet $T = \frac{24 + 2}{2} = 13$.

2. Statistisk Proceskontrol

2) X = koncentrationen af stoffet ved en enkelt måling

X antages normalfordelt med $\tilde{\mu} = 7.407$ og spredning $\tilde{\sigma} = 2.840$

$$P(X > 24) + P(X < 2) = 1 - P(2 < X < 24) = 1 - \text{normCdf}(2, 24, 7.407, 2.84) = 2.85\%$$

3) Bør indstilles på en middelværdi på $T = 8$.

Eksempel 2.6. Beregning af konfidensinterval for C_p

I forbindelse med en analyse af en proces fandt man, på basis af 20 målinger, at processens midtpunkt lå nogenlunde midt mellem øvre tolerancegrænse på 62 og nedre tolerancegrænse på 38. Endvidere skønnede man, at spredningen var ca 1.75.

Find et 95% konfidensinterval for C_p

Løsning

$$c_p = \frac{62 - 38}{6 \cdot 1.75} \approx 2.2857$$

$$c_p \cdot \sqrt{\frac{\text{inv}\chi^2(0.025, 19)}{19}} \approx 1.5649 \quad c_p \cdot \sqrt{\frac{\text{inv}\chi^2(0.975, 19)}{19}} \approx 3.0056$$

95% konfidensinterval for C_p $1.56 \leq C_p \leq 3.01$

2.8 PROCESVARIABLEN X ER DISKRET.

Vi vil her behandle den situation, hvor X enten er binomialfordelt eller Poissonfordelt.

2.8.1: X er binomialfordelt

Procesvariablen X er bestemt ved :

X = antal enheder med fejl ud af en stikprøve på n .

X er binomialfordelt $b(n, p)$.

Sædvanligvis benyttes ved kontrol af fejlprocent ikke p , men middelværdien for binomialfordelingen $\mu = n \cdot p$ som parameter, og man siger man laver et np - kontrolkort.

Som beskrevet i forrige afsnit foretages en kontrolkortanalyse, ved at man opdeler i k undergrupper, som hver er karakteriseret ved en bestemt egenskab.

For hver undergruppe i på n enheder findes antallet af fejlenheder y_i .

Beregnes nu gennemsnittet $\bar{y} = \frac{1}{k} \sum_{i=1}^k y_i$, fås hermed et estimat for $\mu = n \cdot p$.

Et estimat for p er derfor $\hat{p} = \frac{\bar{y}}{n}$ og da spredningen for en binomialfordeling er $\sqrt{n \cdot p \cdot (1 - p)}$

forklarer dette formlerne i oversigt 2.3

Oversigt 2.3

<i>np</i> - kontrolkort. Procesvariablen X er binomialfordelt $b(n,p)$				
Forudsætning	Centerlinie	Nedre kontrolgrænse	Øvre kontrolgrænse	Estimerer
Der udtages k undergrupper hver på n enheder. For hver undergruppe i findes antallet af fejlenheder y_i .	$\bar{y} = \frac{\sum_{i=1}^k y_i}{k}$	$\bar{y} - 3 \cdot \sqrt{\bar{y} \cdot \left(1 - \frac{\bar{y}}{n}\right)}$ (dog altid mindst 0)	$\bar{y} + 3 \cdot \sqrt{\bar{y} \cdot \left(1 - \frac{\bar{y}}{n}\right)}$	$\hat{p} \approx \frac{\bar{y}}{n}$

Eksempel 2.7 (*np* - kort)

En fabrikant af nogle specielle typer keramikfliser som er beregnet til at kunne klare høje temperaturer ønsker udarbejdet et kontrolkort. Ved en løbende produktion af fliser udtoges 40 gange en stikprøve på 100 fliser. De blev undersøgt om de levede op til de forventede kvalitetsmål. Fliser der ikke opfyldte disse krav blev klassificeret som defekte

Resultatet var følgende:

Gruppe	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Antal defekte	8	6	4	4	3	7	3	6	9	5	7	2	6	11	4	6	7	4	9	6

Gruppe	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Antal defekte	6	2	5	7	6	4	6	10	5	5	7	9	3	8	5	3	14	6	4	5

Løsning

Procesvariablen X er bestemt ved :

X = antal enheder uden fejl af en produktion på 100 fliser

X er binomialfordelt $b(100, p)$

Der er i alt 237 defekte fliser fordelt på 40 stikprøver, dvs. i gennemsnit $\bar{y} = \frac{237}{40} = 5.925$ pr. 100 fliser.

Af oversigt 2.3 fås derfor kontrolgrænserne

$$\text{ØKG} = \bar{y} + 3 \cdot \sqrt{\bar{y} \cdot \left(1 - \frac{\bar{y}}{n}\right)} = 5.925 + 3 \cdot \sqrt{5.925 \cdot \left(1 - \frac{5.925}{100}\right)} = 5.925 + 7.0828 = 13.008$$

$$\text{NKG} = 5.925 - 3 \cdot \sqrt{5.925 \cdot \left(1 - \frac{5.925}{100}\right)} = 5.925 - 7.0828 = -1.1578 \text{ dvs. } \text{NKG} = 0$$

Da punkt 37, falder udenfor kontrolgrænserne, foretoges en nærmere undersøgelse af produktionsforholdene på det pågældende tidspunkt, men der blev herved ikke afsløret nogen tegn på væsentlige variationsårsager, jævnfør også, at punkt 37 ikke ligger meget over ØKG.

Man sluttede nu, at processen indtil videre kunne betragtes som værende i kontrol med en procestilstand på $\hat{p} = \frac{\bar{y}}{100} = 0.0593$. ◆

2.8.2: X er Poissonfordelt

Procesvariablen X er bestemt ved :
 X = antal fejl i en stikprøve på n enheder.
 X antages Poissonfordelt $p(\mu)$

Middelværdien af Poissonfordelingen er $c = n \cdot \mu$

Som beskrevet i forrige afsnit foretages en kontrolkortanalyse, ved at man opdeler i k undergrupper, som hver er karakteriseret ved en bestemt egenskab.

For hver undergruppe i på n enheder findes antallet af fejlenheder c_i .

Beregnes nu gennemsnittet $\bar{c} = \frac{1}{k} \sum_{i=1}^k c_i$, fås hermed et estimat for middelværdien $c = n \cdot \mu$.

Vi har derfor, at et estimat for μ er $\tilde{\mu} = \frac{\bar{c}}{n}$, og da spredningen for en Poissonfordeling er $\sqrt{\mu}$ forklarer dette formlerne i oversigt 2.4

Oversigt 2.4

c - kontrolkort. Procesvariablen X er Poissonfordelt $P(\mu)$				
Forudsætning	Centerlinie	Nedre kontrolgrænse	Øvre kontrolgrænse	Estimater
Der udtages k undergrupper hver på n enheder. For hver undergruppe i findes samlet antal fejl c_i .	$\bar{c} = \frac{\sum_{i=1}^k c_i}{k}$	$\bar{c} - 3 \cdot \sqrt{\bar{c}}$ (dog altid mindst 0)	$\bar{c} + 3 \cdot \sqrt{\bar{c}}$	$\hat{\mu} = \frac{\bar{c}}{n}$

Eksempel 2.5.(c - kort)

Ved en tekstilproduktion taltes antal fejl pr. 100 m^2 klæde. Følgende resultater fandtes (tidsmæssig rækkefølge for produktionen) :

nr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
antal fejl	3	3	6	3	0	1	3	5	8	7	4	10	5	5	5	4	2	4	5	1	2	0	1	1	4

Med henblik på en kontrolkortanalyse skal konstrueres et c-kort for processen

Løsning

$$\bar{c} = \frac{\sum_{i=1}^k c_i}{k} = \frac{92}{25} = 3.68$$

$$NKG = 3.68 - 3 \cdot \sqrt{3.68} < 0, \text{ dvs. } NKG = 0$$

$$ØKG = 3.68 + 3 \cdot \sqrt{3.68} = 9.435$$

Punktet (12,10) falder uden for kontrolgrænserne, dvs. processen formodes at have været ude af statistisk kontrol på det pågældende tidspunkt. Fjernes den pågældende undergruppe, fås følgende reviderede kontrolgrænser ud fra den reviderede estimerede procestilstand:

$$\bar{c}_1 = \frac{92 - 10}{24} = 3.42.$$

$$NKG = 0 \text{ (som før)}, \quad ØKG = 3.42 + 3 \cdot \sqrt{3.42} = 8.96.$$

Ingen af de resterende punkter falder uden for de reviderede kontrolgrænser. ◆

OPGAVER

Opgave 2.1

Ved en fabrikation af solbærsyltetøj tilstræbtes et gennemsnitligt nettoindhold af ca. 455 g solbærsyltetøj pr. glas. På glasetiketten anførtes: Nettoindhold 450 g.

Man havde på fabrikken mistanke om, at en af de automatiske fyldemaskiner havde svært ved at fastholde den ønskede nettovægt, og overvejede at foretage et hovedeftersyn af maskinen.

For at afgøre, om et sådant burde foretages, udtog man med 5 minutters mellemrum i alt 18 gange 4 successivt producerede glas fra den omhandlede maskines produktion, og nettoindholdet af solbærsyltetøj bestemtes. Følgende resultater fandtes:

Stikprøve nr	1	2	3	4	5	6	7	8	9
Nettoindhold	452	452	458	456	460	457	452	455	454
	455	457	455	458	456	457	450	456	459
	454	458	458	456	454	452	454	456	456
	453	456	455	456	457	456	453	457	459
\bar{x}	453.5	455.75	456.5	456.5	456.8	455.5	452.3	456.0	457.0
Stikprøve nr	10	11	12	13	14	15	16	17	18
Nettoindhold	454	453	453	453	455	448	450	452	454
	456	457	452	453	449	453	455	448	453
	457	457	455	458	455	456	450	454	454
	456	457	453	455	455	455	453	453	450
\bar{x}	455.8	456.0	453.3	454.8	453.5	453.0	452.0	452.0	452.8

- 1) Lav et $\bar{x} - R$ - kontrolkort.
- 2) Estimér, hvor stor en del af produktionen der ville være fejlproduktion (underfyldte glas), såfremt maskinen i statistisk kontrol med den i 1) konstaterede variation producerede med et middelinhold på 455g .

Opgave 2.2

Med henblik på indførelse af en proceskontrol for en løbende produktion er observeret følgende data:

Maskine nr	1	2	3	4	5	6	7	8	9	10
Data	12.9	11.8	12.8	12.9	11.9	12.1	13.6	13.8	12.0	13.1
	11.6	12.4	12.8	14.4	12.3	12.7	12.0	11.6	11.4	12.2
	13.0	12.7	12.5	13.4	12.5	12.0	12.4	12.4	11.3	12.2

Foretag en kontrolkortanalyse med opstilling af \bar{x} - kort og s - kort.
Punkter der falder udenfor bortkastes

Opgave 2.3

Viskositeten af en polymer (stort kompliceret molekyle) måles en gang i timen.

Målinger foretaget de sidste 20 timer ses i følgende tabel.

test nr	1	2	3	4	5	6	7	8	9	10
Viskositet	2838	2785	3058	3064	2996	2882	2878	2920	3050	2870
test nr	11	12	13	14	15	16	17	18	19	20
Viskositet	3174	3102	2762	2975	2719	2861	2797	3078	2964	2805

- 1) Vurder først om tallene er rimelig normalfordelt ved at tegne et normalfordelingsplot. I det følgende antages, at fordelingen er rimelig normalfordelt
- 2) Lav et "Moving Range kontrolkort" for processen.
- 3) Vurder om processen er i statistisk kontrol, og beregn et estimat for processens middelværdi og spredning.

Opgave 2.4

På en papirfabrik fabrikeredes en bestemt papirtype ved en af virksomhedens maskiner.

Virksomheden besluttede sig i forbindelse med en eksportaftale til at indføre statistisk kvalitetskontrol og foretog indledningsvis en kontrolkortanalyse af G-værdier fra den omtalte maskine. Da man ikke havde nogen tekniske gruppeinddelingskriterier, lod man tidsmæssig nærhed være kriterium og valgte rationelle undergrupper på to på hinanden følgende observationer. Herved fremkom følgende gruppeinddeling:

Gruppe	1	2	3	4	5	6	7	8	9	10	11	12	13
G-værdi	86.2	85.8	85.8	85.7	86.6	86.1	85.6	86.0	86.1	85.3	85.6	85.6	86.0
	85.4	85.6	85.8	86.4	86.1	86.0	86.0	85.9	86.3	85.5	85.2	85.5	85.3

Gruppe	14	15	16	17	18	19	20	21	22	23	24	25	26
G-værdi	85.3	85.8	85.6	85.5	85.3	85.3	85.8	86.3	86.0	86.0	86.3	86.0	85.9
	86.1	85.9	85.6	85.2	85.8	85.6	86.1	86.3	85.8	85.6	86.2	85.8	86.3

- 1) Opstil R - \bar{x} kontrolkort og vurder om processen efter eventuel fjernelse af enkelte punkter på kortene er i kontrol.
- 2) Vi antager i det følgende at processen er i kontrol med de fundne værdier for kontrolgrænser. - Man benytter nu \bar{x} - kontrolkortet til en løbende kontrol. Området deles op i de $2 \cdot 3$ områder, som kaldes $1 \cdot \sigma$, $2 \cdot \sigma$ og $3 \cdot \sigma$ grænserne. En løbende kontrol skal nu foretages med de nedenfor fundne punkter.

Gruppe	1	2	3	4	5	6	7	8	9	10
G-værdi	86.6	86.3	86.6	86.5	86.2	85.9	85.4	85.2	85.6	85.6
	86.1	86.7	86.5	86.6	86.3	85.3	85.1	85.1	85.0	85.3

Første gang en alarm forekommer stopper man, og angiver hvilken alarmgrænse, der er overtrådt

Opgave 2.5

Ved en fabrikation af gips på basis af kalksten ønskedes indført en løbende proceskontrol med \bar{x} og s - kort. Produktets kvalitet vurderedes bl.a. ved måling af DOP-absorptionen (g DOP/100 g pulver).

Ved den indledende kontrolkortanalyse benyttedes successive råvareleverancer som rationelle undergrupper, idet der af hver råvareleverance udtoges en stikprøve af størrelsen 5. Man fandt, at processen var i kontrol med middelværdi 21.94 og spredning 2.13.

Man opstiller på basis af disse målinger følgende \bar{x} - s kontrolkort:

$$NKG_s = 0, \quad \text{Ø}KG_s = 1.482 \quad NKG_{\bar{x}} = 18.90 \quad \text{Ø}KG_{\bar{x}} = 24.98 \quad .$$

Disse kontrolkort benyttes til den løbende kontrol.

- 1) Der er fastsat følgende tolerancegrænser for produktionen:

$$NTG = 10.0 \quad \text{Ø}TG = 27.0$$

Beregn kapabilitetsindeks C_p

Forklar hvorfor dette tal ikke giver et sandt billede af processens evne til holde sig indenfor tolerancegrænserne.

- 2) Hvilken fejlprocent vil fabrikationen have?

- 3) Hvilket middelniveau bør processen ideelt søges indstillet og fastholdt på.

Beregn fejlprocenten hvis processen er indstillet på dette niveau og spredningen fortsat antages at være den i henhold til 1) estimerede?

Opgave 2.6

Fra en fyldeproces er udtaget 25 stikprøver af størrelsen 400. En enhed siges at være underfyldt, hvis enheden er påfyldt mindre end 95g. Antallet af underfyldte enheder i hver stikprøve er optalt. Følgende data fandtes:

Stikprøve nr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Antal underf. enh.	6	14	10	4	13	9	7	11	15	13	5	14	11	8	15	11	9	21	6	12	6	12	8	15	14

Sum 269

Foretag en kontrolkortanalyse og opstil kontrolkort, der kan benyttes til løbende kontrol.

Opgave 2.7

På en samlefabrik indgik en aksemontering. Som et led i en proceskontrol ønskede man specielt at undersøge risikoen for defekte samlinger.

Ved regelmæssig udtagning af stikprøver af størrelsen 250 fandtes følgende resultater:

Stikprøve nr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Antal defekte	11	12	6	18	10	8	12	10	7	11	11	7	5	8	7
Stikprøve nr	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Antal defekte	7	7	10	7	6	8	5	5	7	8	8	7	6	9	6

Foretag med henblik på kontrol af den fremtidige produktion en kontrolkortanalyse.

Opgave 2.8

Antallet af mindre fejl fundet ved den endelige kontrol af 22 DVD-pladespillere er vist i nedenstående tabel

Stikprøve nr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Antal fejl	1	1	3	7	8	10	5	13	0	19	24	6	9	11	15	8	3	6	7	4	9	20

Kan man konkludere ud fra disse data, at processen er i statistisk kontrol.

Opgave 2.9

Defekte på 26 kvadratiske ståloverflader er optalt. Data er vist i nedenstående tabel:

nr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
antal	1	7	4	3	1	2	5	6	2	1	1	7	8	6	2	1	3	5	4	6	3	1	10	2	4	5

- 1) Opsæt et kontrolkort på basis af disse data.
- 2) Synes processen at være i statistisk kontrol på basis af disse data.

3. STATISTISK GODKENDELSESKONTROL

3.1 INDLEDNING.

Vi vil i dette kapitel betragte problemer af følgende type:

Eksempel 3.1. Problemstilling

Fra en **leverandør** til en **aftager** kommer varer i partier bestående af N emner. Hvert parti kan karakteriseres ved en ukendt procent af fejlemner, som kan variere fra parti til parti. Hvis denne fejlprocent er stor, vil aftageren ikke godkende partiet. Hvorledes afgøres, om et parti skal godkendes eller forkastes? ◆

I visse tilfælde er fejlene ved produktionen så uvæsentlige (måske kun skønhedsfejl) at aftagerne ud fra økonomiske overvejelser i en periode har foretrukket helt (eller næsten helt) at undlade en godkendelseskontrol, dvs. leverancerne accepteres uden inspektion.

Omvendt kan der være situationer, hvor konsekvenserne af godkendelse af fejlproduktion er så alvorlige, f.eks. helbredsmæssige konsekvenser for forbrugere af medicin (kritiske fejl), at 100%-kontrol principielt er nødvendig. Produktionen må tilrettelægges, så kritiske fejl ikke forekommer.

Sædvanligvis vil det imidlertid hverken være økonomisk rentabelt ikke at have nogen godkendelseskontrol (mange fejl giver mange klager og store erstatningsomkostninger) eller have en 100% - inspektion af hvert parti (inspektionsomkostningerne bliver for store).

Inspektion af alle emner i et parti stammende fra en massefabrikation af "ens" artikler vil erfaringsmæssigt alligevel sjældent give 100% sikkerhed. Der hævdes at være erfaringer om, at en 100% inspektion af et meget stort antal emner kan være ned til kun 80% effektiv, dvs. at op til 20% af de defekte emner kan slippe igennem kontrollen ved en sådan 100% inspektion (på grund af kedsommeligt rutinearbejde). Hvis kontrollen er destruktiv (bevirker at emnet ødelægges), er en total inspektion naturligvis umulig.

Det mest økonomiske er sædvanligvis at foretage en stikprøvekontrol, udtage nogle emner fra partiet og på grundlag af en vurdering af disse at afgøre, om partiet skal godkendes eller forkastes.

Dette medfører naturligvis, at uanset at kontrollen i gennemsnit fungerer godt, løber aftageren en vis risiko for, at et parti af dårlig kvalitet bliver accepteret, og leverandøren tilsvarende en risiko for, at et parti af god kvalitet bliver forkastet.

Indirekte kan godkendelseskontrollen i høj grad påvirke kvaliteten af en produktion gennem det pres, der lægges på producenten, om at forbedre kvaliteten af det fremstillede produkt, såfremt mange leverancer bliver forkastet ved kontrollen.

Hvis vurderingen af de udtagne emner baseres på en bedømmelse af hvert emne som fejlfrit eller defekt, taler man om "partikontrol ved alternativ variation".

Hvis vurderingen baseres på en måling af en kvantitativ egenskab ved hvert emne såsom diameteren af en aksel, taler man om "partikontrol ved kontinuert variation".

Vi vil i det følgende kun gennemgå "partikontrol ved alternativ variation".

Med hensyn til kontrol med andre egenskaber, f.eks. middelværdi og spredning, henvises til egentlige bøger om kvalitetskontrol.

3.2. ENKELT-STIKPRØVEPLAN

Denne stikprøveplan er den mest benyttede i praksis på grund af, at den er så let at administrere. n kaldes **stikprøvestørrelsen**. c kaldes **godkendelsestallet**.

Definition af enkelt stikprøveplan.

Lad n være et positivt helt tal, og lad c være et ikke-negativt helt tal, hvor $c < n$.

Ved en enkelt - stikprøveplan (n, c) udtages tilfældigt en stikprøve af størrelsen n , og antallet x af defekte i stikprøven optælles.

Partiet godkendes, såfremt $x \leq c$; partiet forkastes, såfremt $x > c$.

Eksempel 3.2. Enkelt stikprøveplan

Et legetøjsfirma modtager leverancer bestående af 10.000 billige dukker, og ønsker at kontrollere disses kvalitet ved stikprøveplanen $(n, c) = (100, 3)$.

Angiv hvorledes kontrollen skal foregå.

Løsning:

Af hver leverance (på 10.000 dukker) udtages en stikprøve på 100.

Hvis 3 eller færre af disse er defekte godkendes hele leverancen, ellers forkastes det.



OC-kurve for en stikprøveplan

For en given stikprøveplan kan man beregne sandsynligheden for at et parti bliver godkendt (acceptsandsynligheden p_a) som funktion af partiets fejlprocent.

Grafen for denne funktion kaldes funktionens OC-kurve.

Beregning af acceptsandsynlighed.

Der udtages en stikprøve på n emner uden tilbagelægning ud af en leverance på N emner.

Lad c være godkendelsestallet.

X = antal fejlemner blandt n emner. Lad sandsynligheden for fejl i leverancen være p .

Acceptsandsynligheden er $p_a(p) = P(X \leq c)$.

Der er $M = N \cdot p$ fejlemner i partiet, dvs. X er hypergeometrisk fordelt $h(N, M, n)$.

Sædvanligvis vil det gælde, at $\frac{n}{N} \leq \frac{1}{10}$, og derfor kan approksimeres med binomialfordelingen $b(n, p)$.

Eksempel 3.3. Beregning af OC - kurve.

Et legetøjsfirma modtager leverancer bestående af $N = 10.000$ dukker, og ønsker at kontrollere disses kvalitet ved stikprøveplanen $(n, c) = (100, 3)$.

Beregn acceptsandsynligheden p_a for fejlprocenterne 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, og tegn på grundlag heraf stikprøveplanens OC - kurve.

Løsning:

X = antal fejlemner blandt $n = 100$ emner.

Er sandsynligheden for fejl p i leverancen er der $M = N \cdot p = 10000 \cdot p$ fejlemner i partiet.

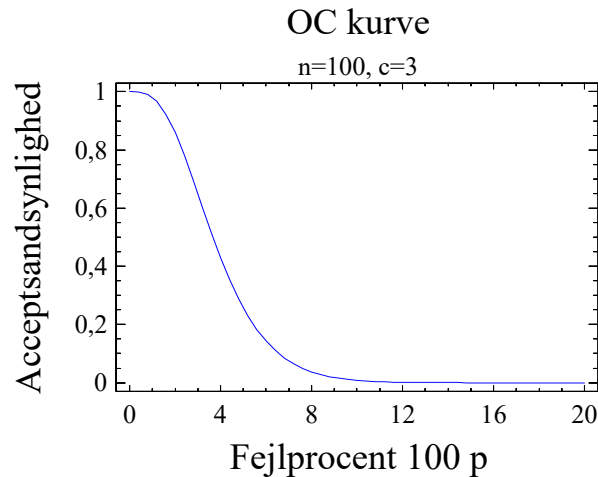
X er derfor hypergeometrisk fordelt $h(10000, 10000 \cdot p, n)$

Da $\frac{n}{N} = \frac{100}{10000} \leq \frac{1}{10}$ kan fordelingen for X approksimeres med binomialfordelingen $b(100, p)$.

Vi finder da eksempelvis, at $P_a(0.02) = P(X \leq 3) = \text{binomCdf}(100, 0.02, 0, 3) = 0.859$.

På tilsvarende måde findes de øvrige værdier.

100p%	1	2	3	4	5	6	7	8	9	10
100 $P_a(p)$ %	98.2	85.9	64.7	42.9	25.8	14.3	7.4	3.6	1.7	0.8



Valg af parametre i enkelt-stikprøveplan.

Til bestemmelse af stikprøvestørrelsen n og godkendelsestal c vælger de to parter, leverandør (producer) og aftager (consumer) to “risikopunkter” som på OC - kurven skal gå igennem.

1) den tilfredsstillende kvalitet AQL (Acceptable Quality Level), og den tilsvarende acceptsandsynlighed $P_a(AQL) = 1 - \alpha$, hvor α kaldes “leverandørens risiko”.

$(AQL, 1 - \alpha)$ kaldes leverandørens risikopunkt.

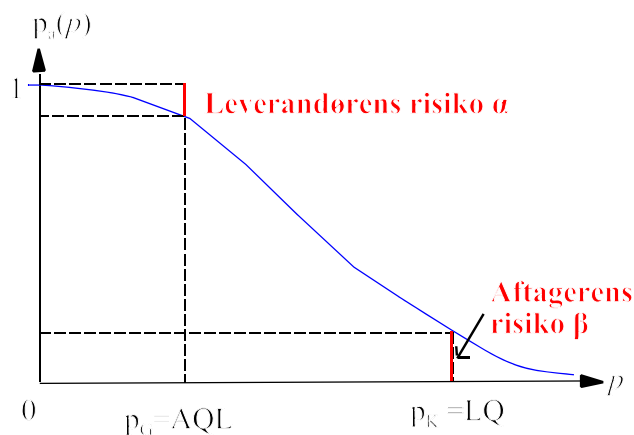
2) den utilfredsstillende kvalitet LQ (Limiting Quality), og den tilsvarende acceptsandsynlighed $P_a(LQ) = \beta$, hvor β kaldes “aftagerens risiko”.

(LQ, β) kaldes aftagerens risikopunkt.

De to risikopunkter må altid vælges ud fra de konkret foreliggende forhold (tekniske, økonomiske m.v.), idet AQL sættes lig en “lille” fejlandel p_G , og LQ sættes lig en noget større fejlandel p_K . (jævnfør den følgende figur)

Traditionelt vælges næsten altid $\alpha = 5\%$ og $\beta = 10\%$. Undertiden vælges dog $\beta = 5\%$.

3. Statistisk godkendelseskontrol



En enkelt stikprøveplan er i princippet fastlagt ved disse specifikationer og kan approksimativt bestemmes ved hjælp af tabel 3.1.

Eksempel 3.4. Bestemmelse af stikprøveplan.

Ved levering af et partier på 10.000 dukker træffer en leverandør og en aftager aftale om at vælge en enkelt - stikprøveplan bestemt ved, at acceptsandsynligheden i tilfælde af 0.6% defekte dukker i partiet skal være 95%, og i tilfælde af 3,1% defekte skal være ca. 10%.

Bestem stikprøvestørrelsen n og godkendelsestallet c .

Løsning:

Tabel 3.1 anvendes

$$\text{Idet } AQL = 0,6\% \text{ og } LQ = 3,1\% \text{ er } R = \frac{LO}{AQL} = \frac{p_k}{p_G} = \frac{3,1}{0,6} = 5,1667.$$

I tabel 3.1 i kolonnen for $\alpha = 5\%$ og $\beta = 10\%$ fås, at det største tal mindre end 5.17 er 4.89.

Dette svarer til, at godkendelsestallet $c = 3$. I kolonnen for $n \cdot p_g$ og $\alpha = 5\%$ findes

$$n \cdot p_G = 137 \Leftrightarrow n = \frac{137}{0,6} = 228,33 \approx 229$$

Vi finder derfor $c = 3$ og $n = 229$

Poissonfordeling anvendes

I stedet for at benytte tabellen kan man da p-værdierne er små (forhåbentlig højst 10%) approksimere binomialfordelingen $b(n,p)$ med Poissonfordelingen $P(\mu)$ hvor $\mu = n \cdot p$.

$$\text{Vi har da, at } R = \frac{P_K}{P_G} = \frac{n \cdot p_K}{n \cdot p_G} = \frac{\mu_K}{\mu_G}$$

Først gættes på $c = 5$:

$$c = 5 \rightarrow 5$$

$$\frac{\text{solve}(\text{poissCdf}(x,0,c)=0.1,x)|x>0}{\text{solve}(\text{poissCdf}(y,0,c)=0.95,y)|y>0} \rightarrow \frac{x}{y} = 3.5494$$

Da $R = 5,167 > 3,549$ sættes c ned indtil man kommer op omkring R .

$$c = 4: \frac{x}{y} = 4,05 \quad c = 3: \frac{x}{y} = 4,87 \quad c = 2: \frac{x}{y} = 6,49$$

Vi vælger nu det nærmeste tal under den ønskede værdi for $R = 0.5167$, dvs $c = 3$

Vi har nu for $c = 3$: $n \cdot p_G = x \Leftrightarrow n = \frac{137}{0006} = 229$ Konklusion: $(n, c) = (229, 3)$ ◆

3.3. REKTIFICERENDE KONTROL.

Undertiden aftales, at stikprøvekontrollen skal udføres på den måde, at leverandøren underkaster alle kasserede partier 100% - inspektion og "rensner" disse partier, dvs. erstatter fejlemner med fejlfrie emner. En sådan kontrol kaldes for "rektificerende kontrol" (rensende kontrol).

En sådan aftale kan eksempelvis ske ved leverancer inden for samme virksomhed (intern kontrol) eller hvis en (dominerende) storkøber ikke er sikker på, at en leverandør kan overholde aftagerens krav til AQL - værdi.

I forbindelse med fastlæggelse af en enkelt stikprøveplan med rektificerende kontrol spiller følgende begreber en stor rolle.

AOQ (Average Outgoing Quality): Fejlandelen af de udgående partier når de indkommende partier produceres af en proces, der med en sandsynlighed på p frembringer fejlemner.

Der gælder formlen $AOQ = p \cdot P_a(p)$.

Bevis: Det modtagne parti har en "fejlsandsynlighed" på p , og sandsynligheden for at dette parti bliver godkendt er $P_a(p)$.

Det udgående parti (efter den rektificerende kontrol), vil derfor have en "fejlsandsynlighed" p_u på enten p eller på 0 afhængig af om det blev godkendt eller ej.

Vi får derfor: $AOQ = E(p_u) = p \cdot P_a(p) + 0 \cdot (1 - P_a(p)) = p \cdot P_a(p)$ ◆

AOQL (Average Outgoing Quality Limit): Den maksimale værdi af AOQ .

ATI (Average Total Inspection): Den gennemsnitlige totale inspektionsstørrelse når de indkomne partier produceres af en proces, der med en sandsynlighed på p frembringer fejlemner.

Der gælder formlen $ATI = n + (N - n) \cdot (1 - P_a(p))$

Bevis:

Lader vi n_T betegne det totale antal inspicerede emner ved en rektificerende kontrol, må det gælde,

$$n_T = \begin{cases} n & \text{med sandsynligheden } P_a(p) \\ N & \text{med sandsynligheden } 1 - P_a(p) \end{cases}$$

Middelværdien bliver $E(n_T) = n \cdot P_a(p) + N \cdot (1 - P_a(p)) = n + (N - n) \cdot (1 - P_a(p))$ ◆

Det følgende eksempel illustrerer beregningerne.

Eksempel 3.5 Rektificerende kontrol.

Vi betragter atter den i eksempel 3.4 omtalte situation, med kontrol af dukker, idet man har valgt enkeltstikprøveplanen $(n, c) = (229, 3)$.

Vi antager nu, at der yderligere er aftalt, at kontrollen udføres som rektificerende kontrol.

Beregn acceptsandsynligheden AQL og ATI for $p = 2\%$.

Løsning:

Da vi har et stort parti kan approksimeres med binomialfordelingen $b(229, 0.02)$

$$P_a(0.02) = P(X \leq 3) = \text{binomCdf}(229, 0.02, 0, 3) = 0.326.$$

og dermed og $ATI = 229 + (10000 - 229) \cdot (1 - 0.326) = 6811.2$ ◆

3.4. DOBBELT STIKPRØVEPLAN

Anvendelsen af dobbelt-stikprøveplaner er administrativt noget besværligere end benyttelsen af enkelt-stikprøveplaner. Er det imidlertid meget dyrt (besværligt) at kontrollere emnerne, er det væsentligt at stikprøvestørrelsen bliver så lille som muligt og man kan i sådanne situationer se en fordel i at gå over til en dobbelt-stikprøveplan. Den gennemsnitlige stikprøvestørrelse vil derved ofte kunne formindskes i forhold til stikprøvestørrelsen n for en enkelt stikprøveplan (n, c) , som går igennem de samme risikopunkter.

Definition af dobbelt stikprøveplan.

Lad n_1 og n_2 være positive hele tal, og lad c_1, c_2 og c_3 være ikke-negative hele tal, hvor

$$c_1 < c_2 < n_1 \text{ og } c_2 \leq c_3 < n_1 + n_2.$$

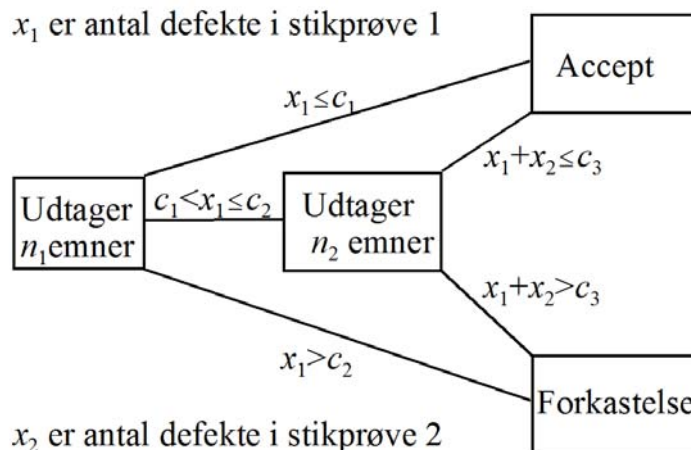
Ved en dobbelt-stikprøveplan $(n_1, n_2, c_1, c_2, c_3)$ udtages tilfældigt en stikprøve af størrelsen n_1 og antallet x_1 af defekte i stikprøven optælles.

Partiet godkendes, såfremt $x_1 \leq c_1$; partiet forkastes, såfremt $x_1 > c_2$.

Såfremt $c_1 < x_1 \leq c_2$ udtages en ny stikprøve af størrelsen n_2 og antallet x_2 af defekte i den anden stikprøve optælles.

Partiet godkendes, såfremt $x_1 + x_2 \leq c_3$; partiet forkastes, såfremt $x_1 + x_2 > c_3$

Skematisk kan en dobbelt stikprøveplan fremstilles på følgende måde:



Som det fremgår, forudsætter fastlæggelsen af en bestemt dobbelt-stikprøveplan valg af stikprøvestørrelserne n_1 og n_2 og de tre tal c_1, c_2 og c_3 .

Hvorledes dette valg påvirker beslutningen om accept eller forkastelse af modtagne varepartier, som underkastes kontrol i henhold til en dobbelt-stikprøveplan, illustreres som ved enkelt-stikprøveplaner af den til planen hørende OC-kurve.

Bestemmelse af en dobbelt-stikprøveplan

OC-kurven for den valgte plan skal på samme måde som ved enkelt-stikprøveplan så vidt muligt gå gennem de 2 risikopunkter $(p_G, 1 - \alpha)$ og (p_k, β) .

Dette er imidlertid ikke tilstrækkeligt til en entydig bestemmelse af en dobbelt-stikprøveplan. En sådan bliver derimod mulig ved valg af ekstra bånd mellem parametrene. Sædvanligvis forlanger man, at $n_2 = n_1$ eller $n_2 = 2 \cdot n_1$. Endvidere vælges ofte $c_3 = c_2$, hvilket har vist sig ikke at hindre, at man i praksis altid kan bestemme en velegnet dobbelt-stikprøveplan.

På samme måde som for enkelte stikprøveplaner er der for dobbelte stikprøveplaner udarbejdet tabeller. Tabel 3.2 er en sådan tabel, hvor, $n_2 = n_1$ eller $n_2 = 2 \cdot n_1$, $c_3 = c_2$, $\alpha = 5\%$ og $\beta = 10\%$.

ASN (Average Sample Number): Værdien af den gennemsnitlige stikprøvestørrelse for en given værdi af "produktionskvaliteten p ."

Der gælder formelen $ASN = n_1 + n_2 \cdot P(c_1 < X_1 \leq c_2)$

Formlen fremgår af, at der altid udtages en stikprøve på n_1 og der så yderligere udtages n_2 med en sandsynlighed på $P(c_1 < X_1 \leq c_2)$.

Ækvivalente stikprøveplaner er stikprøveplaner, der går gennem de samme risikopunkter.

Som nævnt er fordelene ved de dobbelte stikprøveplaner i forhold til de enkelte stikprøveplaner, at for ækvivalente stikprøveplaner, vil den gennemsnitlige stikprøvestørrelse være mindre for de fleste (alle) værdier af p .

Eksempel 3.6 Bestemmelse af en dobbelt-stikprøveplan.

Ved levering af et parti på 10.000 dukker træffer en leverandør og en aftager aftale om at vælge en dobbelt - stikprøveplan $(n_1, n_2, c_1, c_2, c_3)$ som er ækvivalent med den i eksempel 3.4 angivne enkelte stikprøveplan, dvs. at acceptandsynligheden i tilfælde af 0.6% defekte dukker i partiet skal være 95%, og i tilfælde af 3.1% defekte skal være ca. 10%.

- 1) Bestem stikprøveplanens parametre, idet vi ønsker, at $n_2 = 2 \cdot n_1$, og $c_3 = c_2$.
- 2) Skitser OC-kurven på basis af de 4 støttepunkter man kan beregne ud fra tabellen
- 3) Find acceptandsynligheden for $p = 2\%$ ved aflæsning på kurven.
- 4) Find acceptandsynligheden for $p = 2\%$ ved direkte beregning.
- 5) Find ASN for de to risikopunkter, samt for $p = 4\%$.

LØSNING:

$$1) \text{ Idet } AQL = 0,6\% \text{ og } LQ = 3,1\% \text{ er } \frac{LO}{AQL} = \frac{p_k}{p_G} = \frac{3,1}{0,6} = 5,1667.$$

I tabel 3.2 ud for række 9 fås, at det største tal mindre end 5.17 er 5.09.

$$\text{Dette svarer til, at } c_1 = 1 \text{ og } c_2 = 4. \text{ Endvidere fås } n_1 \cdot p_G = 77 \Leftrightarrow n_1 = \frac{77}{0,6} = 128,33.$$

Vi finder derfor $(n_1, n_2, c_1, c_2, c_3) = (129, 258, 1, 4, 4)$.

3. Statistisk godkendelseskontrol

2) $\underline{P_a(0) = 100\%}$

$\underline{P_a(0.006) = 95\%}$:Leverandørens risiko må være ca. 5%, da vi her beregnede n_1 ud fra $p_G = 0.006$.

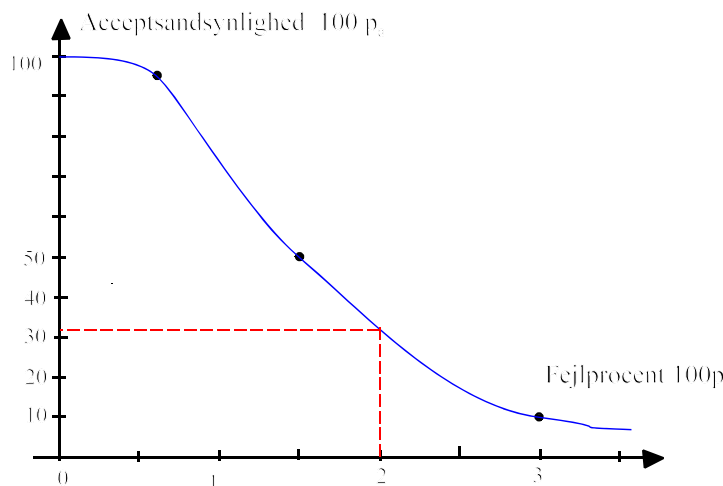
Af række 9, fås, at $p_a(p_K) = 10\%$ for $n_1 \cdot p_K = 392 \Leftrightarrow p_K = \frac{392}{129} = 3.04\%$, dvs.

$\underline{P_a(0.0304) = 10\%}$. Aftagerens risiko på 10% opnås for $p = 3.0\%$, altså lidt mindre end den aftalte på 3.1%. Dette skyldes, at tabellen ikke indeholder forholdet 5.17.

Af række 9, fås, at $p_a(p) = 50\%$ for $n_1 \cdot p = 197 \Leftrightarrow p = \frac{197}{129} = 1.53\%$, dvs.

$\underline{P_a(0.015) = 50\%}$. Vi har altså fundet følgende tabel:

$100 \cdot p \%$	0	0.6	1.5	3
$100 \cdot P_a(p) \%$	100	95	50	10



3) Af kurven fås, at $\underline{P_a(0.02) \approx 31\%}$

4) X_1 =antal fejlemner blandt $n_1 = 129$ emner
 X_2 = antal fejlemner blandt $n_2 = 258$ emner

X_1 og X_2 er hypergeometrisk fordelt, men da $\frac{n}{N} \leq \frac{1}{10}$ kan approksimeres med en binomialfordeling.

Man får en accept på 2 måder:

Ved 1. stikprøve på 129, at få højst 1 defekt.
 eller

Ved 1. stikprøve på 129 at få over 1 og mindre end eller lig 4, og i alt på 1. og 2. stikprøve højst at have fået 4.

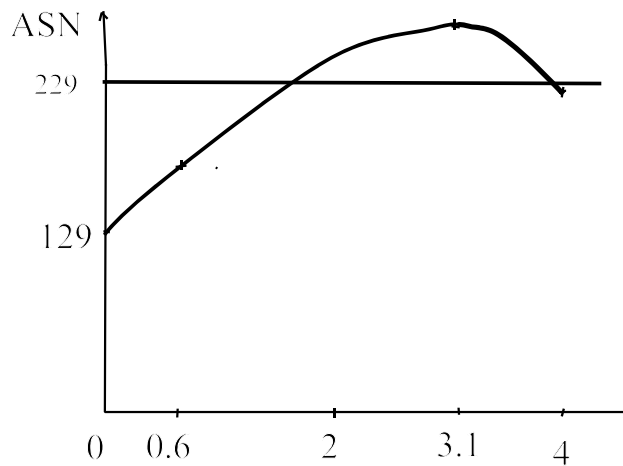
$$P_a(0.02) = P(X_1 \leq 1) + P(X_1 = 2) \cdot P(X_2 \leq 2) + P(X_1 = 3) \cdot P(X_2 \leq 1) + P(X_1 = 4) \cdot P(X_2 = 0)$$

$$P(X_1 \leq 1) = \text{binomCdf}(129, 0.02, 0, 1) = 0.2682$$

$$\begin{aligned}
 &P(X_1 = 2) \cdot P(X_2 \leq 2) + P(X_1 = 3) \cdot P(X_2 \leq 1) + P(X_1 = 4) \cdot P(X_2 = 0) \\
 &= \text{binomPdf}(129, 0.02, 2) \cdot \text{binomCdf}(258, 0.02, 0, 2) \\
 &+ \text{binomPdf}(129, 0.02, 3) \cdot \text{binomCdf}(258, 0.02, 0, 1) \\
 &+ \text{binomPdf}(129, 0.02, 4) \cdot \text{binomCdf}(258, 0.02, 0, 0) = 0.03602 \\
 &P_a(0.02) = 0.2682 + 0.0360 = 0.3042 \approx \underline{30,4\%}
 \end{aligned}$$

5) $ASN = 129 + 258 \cdot P(1 < X_1 \leq 4)$

$100 \cdot p \%$	0	0.6	3.1	4
$P(1 < X_1 \leq 4)$	0	$\text{binomCdf}(129, 0.006, 2, 4) = 0.181$	$\text{binomCdf}(129, 0.031, 2, 4) = 0.541$	$\text{binomCdf}(129, 0.04, 2, 4) = 0.3759$
ASN	129	175.6	$129 + 258 \cdot 0.55 = 268.5$	$129 + 258 \cdot 0.376 = 226.0$



Det ses som forventet, at for meget gode partier og for meget dårlige partier, vil den gennemsnitlige stikprøvestørrelse blive mindre end de 229 vi fandt ved den ækvivalente enkeltstikprøveplan.



Tabel 3.1. Bestemmelse af enkelt - stikprøveplaner

$R = \frac{p_k}{p_g} \quad p_G \text{ og } p_k \text{ regnes i procenter}$												
α												
c	1%			5%			10			α		
	β			β			β			1%	5%	10%
	1%	5%	10%	1%	5%	10%	1%	5%	10%	$n \cdot p_G$		
0	458	298	229	89.8	58.4	45.1	43.2	28.1	21.8	1.0	5.1	10.5
1	44.7	31.9	26.2	18.7	13.3	11.0	12.7	8.97	7.40	14.9	35.5	52.8
2	19.3	14.4	12.2	10.3	7.70	6.50	7.65	5.72	4.84	43.6	81.8	110
3	12.2	9.42	8.12	7.35	5.68	4.89	5.77	4.46	3.83	82.3	137	174
4	9.07	7.16	6.25	5.89	4.65	4.06	4.77	3.77	3.29	128	197	243
5	7.34	5.89	5.20	5.02	4.02	3.55	4.16	3.33	2.95	179	261	315
6	6.25	5.08	4.52	4.44	3.60	3.21	3.74	3.03	2.71	233	329	389
7	5.51	4.52	4.05	4.02	3.30	2.96	3.44	2.81	2.53	291	398	465
8	4.96	4.12	3.71	3.71	3.07	2.77	3.20	2.65	2.39	351	470	544
9	4.55	3.80	3.44	3.46	2.90	2.62	3.02	2.52	2.28	413	543	623
10	4.22	3.56	3.23	3.27	2.75	2.50	2.87	2.42	2.20	477	617	701
11	3.96	3.35	3.06	3.10	2.63	2.40	2.75	2.32	2.12	543	692	784
12	3.74	3.19	2.92	2.97	2.53	2.31	2.64	2.25	2.06	610	769	865
13	3.56	3.05	2.80	2.85	2.44	2.24	2.55	2.19	2.00	678	846	950
14	3.40	2.93	2.69	2.75	2.37	2.18	2.51	2.16	1.96	748	925	1030
15	3.27	2.82	2.60	2.67	2.30	2.12	2.41	2.08	1.91	818	1000	1110

Tabel 3.2: Bestemmelse af dobbelt stikprøveplan.

$\alpha=5\%$, $\beta=10\%$; $c_3 = c_2$. p regnet i %.

Plan nr	$\frac{p_K}{p_G}$	n_2	c_1	c_2	$P_a(p)$		
					95% (p_G)	50%	10% (p_K)
					$n_1 \cdot p$		
1	14.50	$2 \cdot n_1$	0	1	16	84	232
2	11.90	n_1	0	1	21	100	250
3	8.07	$2 \cdot n_1$	0	2	30	107	242
4	7.54	n_1	1	2	52	182	392
5	6.79	n_1	0	2	43	142	296
6	6.48	$2 \cdot n_1$	1	3	60	180	389
7	5.39	n_1	1	3	76	211	411
8	5.39	$2 \cdot n_1$	0	3	49	135	264
9	5.09	$2 \cdot n_1$	1	4	77	197	392
10	4.65	n_1	2	4	116	290	539
11	4.31	$2 \cdot n_1$	0	4	68	164	293
12	4.25	n_1	1	4	104	250	442
13	4.19	$2 \cdot n_1$	1	5	96	218	402
14	3.88	n_1	2	5	143	320	555
15	3.63	n_1	3	6	187	398	678
16	3.60	$2 \cdot n_1$	1	6	116	244	417
17	3.38	n_1	2	6	172	356	582
18	3.26	$2 \cdot n_1$	2	8	168	328	547
19	3.21	n_1	3	7	215	427	691
20	3.09	n_1	4	8	262	502	810
21	2.96	$2 \cdot n_1$	3	10	227	413	672
22	2.85	n_1	4	9	290	533	826
23	2.77	$2 \cdot n_1$	3	11	246	436	682
24	2.62	$2 \cdot n_1$	4	13	307	521	805
25	2.60	n_1	5	11	368	640	956
26	2.46	$2 \cdot n_1$	4	14	329	540	811
27	2.44	n_1	5	12	400	673	977
28	2.32	n_1	5	13	435	706	1008
29	2.22	n_1	5	14	470	752	1045
30	2.21	$2 \cdot n_1$	3	15	341	540	755
31	2.12	n_1	5	16	539	840	1141
32	1.97	$2 \cdot n_1$	4	20	475	702	935

Opgaver

Opgave 3.1

En fabrik fremstiller billige disketter. Disketterne pakkes i kasser, som hver indeholder 100 disketter.

Produktionen er i statistisk kontrol, og det vides, at der i gennemsnit er 1 defekt diskette ud af 125 disketter .

- 1) Angiv øvre og nedre kontrolgrænse for et kontrollkort, baseret på, at man hver time udtager en kasse på 100 disketter, og optæller antallet af defekte disketter.

En forretningskæde med 5 forretninger køber hver måned et stort antal disketter fra fabrikken. Kæden træffer aftale med fabrikanten om at benytte en enkelt-stikprøveplan med følgende specifikationer:

Ved den tilfredsstillende kvalitet: fejlprocent 0.8, skal leverandørens risiko være 10%.

Ved den utilfredsstillende kvalitet: fejlprocent 3, skal aftagerens risiko være 10%.

- 2) Bestem stikprøvestørrelsen n (afrund til nærmeste med 100 delelige tal) og godkendelsestallet c .

- 3) Beregn aftagerens risiko ved en fejlprocent på 3.

- 4) Hver af de 5 forretninger foretager ovennævnte stikprøvekontrol.

Hvis produktionen er i statistisk kontrol således, at der som nævnt i gennemsnit forekommer 1 defekt diskette ud af 125 disketter, beregn da sandsynligheden for, at fabrikanten får et parti godkendt,

- a) i 2 forretninger ud af de 5,
- b) i alle 5 forretninger .

Opgave 3.2

Ved en statistisk partikontrol ved alternativ variation fastsættes en enkelt-stikprøveplan ved følgende specifikationer:

Ved den tilfredsstillende kvalitet: fejlprocent 3, skal leverandørens risiko være 5%.

Ved den utilfredsstillende kvalitet: fejlprocent 8, skal aftagerens risiko være 5%.

- 1) Bestem stikprøvestørrelsen n (afrund til nærmeste med 10 delelige tal) og godkendelsestallet c .
- 2) Beregn for den valgte stikprøveplan acceptsandsynlighederne $P_a(0.04)$, $P_a(0.05)$, $P_a(0.06)$, $P_a(0.07)$, $P_a(0.10)$.
- 3) Skitsér stikprøveplanens OC-kurve.

Opgave 3.3

En kemisk fabrik bestemte sig for at indføre kontrol med kvaliteten af en produktion, som virksomhedens halvfabrikatafdeling leverede til færdigvareafdelingen.

Kontrollen udførtes som partikontrol ved alternativ variation med en enkelt-stikprøveplan fastlagt ved følgende specifikationer:

Ved den tilfredsstillende kvalitet, fejlprocent 2, skal leverandørens risiko være 1%.

Ved den utilfredsstillende kvalitet, fejlprocent 10, skal aftagerens risiko være 1%.

- 1) Bestem stikprøveplanens parametre. stikprøvestørrelsen n (der rundes op til et med 10 deleligt tal) og godkendelsestallet c .

Halvfabrikatafdelingen havde en dagsproduktion på 2500 emner. og stikprøvekontrollen foretoges én gang daglig på den samlede dagsproduktion.

Efter nogen tids forløb bestemte man sig til med bibeholdelse af den valgte enkelt-stikprøveplan at udføre kontrollen som en rektificerende kontrol.

2) Idet $100p'$ angiver de til kontrollen indgående partiers fejlprocent og $100p_u$ de fra kontrollen udgående partiers fejlprocent, ønskes $E(p_u)$ bestemt for følgende værdier af p' :

p'	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.10
------	------	------	------	------	------	------	------	------

3) Angiv grafisk $E(p_u)$ som funktion af p' .

4) Bestem, idet n_T betegner det totale antal emner, som inspiceres ved den rektificerende kontrol, $E(n_T)$ for den værdi af p' , for hvilken $E(p_u)$ er størst.

Opgave 3.4

En enkelt -stikprøveplan ønskes fastlagt ved følgende specifikationer:

Tilfredsstillende kvalitet AQL : $p_G = 1.0\%$. Leverandørens risiko: $\alpha = \text{ca } 5\%$.

Utilfredsstillende kvalitet LQ : $p_k = 7.0\%$. Aftagerens risiko: $\beta = \text{ca } 10\%$.

- 1) Bestem stikprøveplanens parametre.
- 2) Bestem den ækvivalente dobbelte stikprøveplans parametre, idet der yderligere er givet, at stikprøvestørrelserne: $n_2 = 2 \cdot n_1$, og godkendelsestallene: $c_3 = c_2$
- 3) Skitsér OC - kurven på basis af de 4 støttepunkter man kan beregne på basis af tabellen, samt for $p = 4\%$.
- 4) Beregn den gennemsnitlige stikprøvestørrelse ASN for de to risikopunkter, samt for $p = 4\%$ og for $p = 10\%$ og afbild resultatet grafisk.
- 5) Giv på basis heraf en vurdering af om det ville være en fordel at gå over til den dobbelte stikprøveplan.

Opgave 3.5

Ved en kontrol af en bestemt type installerede husholdningsgasmålere udførtes en stikprøvekontrol af disse ved anvendelse af en dobbelt-stikprøveplan.

Man valgte en procedure, hvorved målere fra samme årgang (partistørrelse ca. 2000) med 4 års mellemrum kontrolleredes ved stikprøveudtagning.

En dobbelt-stikprøveplan ønskes fastlagt ved følgende specifikationer:

Stikprøvestørrelserne: $n_2 = n_1$, Godkendelsestallene: $c_3 = c_2$

Tilfredsstillende kvalitet AQL : $p_G = 2.4\%$. Leverandørens risiko: $\alpha = \text{ca } 5\%$.

Utilfredsstillende kvalitet LQ : $p_k = 9.0\%$. Aftagerens risiko: $\beta = \text{ca } 10\%$.

- 1) Bestem dobbelt-stikprøveplanens parametre.
- 2) Skitsér OC - kurven på basis af de 4 støttepunkter man kan beregne på basis af tabellen.
- 3) Bestem den gennemsnitlige stikprøvestørrelse for de ovenanførte p -værdier.

4 PLANLÆGNING AF FORSØG

4.1. INDLEDNING

Forsøg er en naturlig del af ingeniørmæssig og anden videnskabelig metode til at træffe beslutninger.

Antag eksempelvis, at en ingeniør skal studere virkningerne af 4 hærdningsmetoder på trykstyrken af et materiale.

Forsøget ville bestå i, at man fremstillede en række testmaterialer baseret på de 4 hærdningsmetoder, og derefter målte trykstyrken. På basis af disse data kunne man så anvende en statistisk metode til at finde den af de 4 metoder der i middel gav den største trykstyrke.

Alle forsøg er planlagte forsøg, men desværre er nogle forsøg særdeles dårlig planlagt, og resulterer i at kostbare ressourcer bliver benyttet ineffektivt.

Statistisk planlagte forsøg giver effektivitet og økonomi i den eksperimentelle proces, og brug af statistiske metoder i undersøgelsen af data resulterer i en **videnskabelig objektivitet** når man skal drage konklusioner.

Statistisk baserede forsøg er særlig nyttige til at forbedre en fremstillingsproces eller til at udvikle nye metoder.

Ved at benytte statistisk planlagte forsøg, kan ingeniøren bestemme hvilke af de mange procesvariable, såsom temperatur, tryk, hærdningsmetoder osv. der har den største betydning for udfaldet af processen.

Brugen af statistisk baserede forsøg kan derfor resultere i produkter, der er lettere at producere, produkter der har en bedre "performance" og stabilitet (mindre spredning) end konkurrenternes produkter, og kan blive udviklet og produceret på mindre tid, hvilket reducerer udviklingsomkostningerne.

4.2. NOMENKLATUR

I de følgende kapitler benyttes ord, som faktor, niveauer, behandlinger osv.

For at forstå hvad disse ord betyder, vil vi forklare dem ud fra følgende forsøg:

Eksempel 4.1 Nomenklatur

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

Forsøgsresultaterne var:

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

Beskriv forsøget.

Løsning:

Forsøget har to **faktorer**: Cementblander og Cementknuser.

Faktoren "Cementblander" har 3 **niveauer** A₁, A₂, A₃. (niveau hedder på engelsk "level")

Faktoren "Cementknuser" har 4 **niveauer** B₁, B₂, B₃, B₄

Forsøget har 12 **behandlinger** (engelsk treatment)

A₁B₁, A₁B₂, A₁B₃, A₁B₄, A₂B₁, A₂B₂, A₂B₃, A₂B₄, A₃B₁, A₃B₂, A₃B₃, A₃B₄
da der er 12 kombinationer af niveauerne (12 celler)

Hver behandling har 3 **gentagelser**, eksempelvis har behandlingen A₁B₁ 3 **delforsøg**, der resulterede i **forsøgsresultaterne** 147 175 130



Faktorer kan enten være **kvalitative** eller **kvantitative**.

En faktor som temperatur er kvantitativ, da den jo er en talvariabel, der kan antage alle mulige talværdier (indenfor et givet talområde).

En faktor som "Cementblander" i eksempel 2.1 er kvalitativ, da den kun har nogle fastlagte niveauer, og man ikke kan tale om eksempelvis cementblander 1.5.

4.3 KRAV TIL STATISTISK GYLDIGT FORSØG

For at nogle forsøgsresultater skal være statistisk gyldige, skal målingerne være statistisk **uafhængige** og være **repræsentative** for det man skal undersøge.

Ved statistisk uafhængighed forstås, at resultatet af et delforsøg ikke må afhænge af hvad der skete i de øvrige delforsøg.

Det er således ikke korrekt, hvis det arbejds hold, der foretager forsøgene først laver forsøgene med den ene cementblander- derved bliver dygtigere- og så laver forsøgene med de øvrige cementblandere.

Det er heller ikke korrekt, at man eksempelvis i eksempel 1.1 først havde målt holdbarheden af 10 blokke, - derefter foretager en test-opdager at man ikke kan vise signifikans.

Så tager man 10 blokke mere - testet på de 20 blokke osv., indtil man opnåede signifikans.

Dette er ikke "statistisk gyldige" forsøg.

4. Planlægning af forsøg

Til belysning af hvad der er et "statistisk" gyldigt forsøg tages udgangspunkt i følgende eksempel.

Eksempel 4.2. Planlægning af forsøg.

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.

Alle andre delforsøgsbetingelser end behandlingerne sammenfattes i et begreb, der kaldes forsøgsenheden.

I eksempel 4.2 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 "forsøgsenhederne" 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 4.2 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 4.2 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

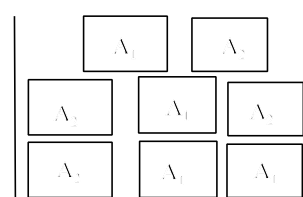
Ét delforsøg hvor apparatur nr 1 og laborant B benyttes, og

Ét delforsøg hvor apparatur nr 3 og laborant C benyttes.

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

Forsøgsenheder			Behandlinger (additiv)
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 lommeregnerne 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 laborant A skal anvende additiv A_1 osv. Resultaterne kunne eksempelvis være som angivet på følgende skema:



Forsøgsenheder			Behandlinger (additiv)
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

Man kan fristes til at tro, at jo flere gentagelser jo bedre.

Da spredningen på et gennemsnit er $\frac{s}{\sqrt{n}}$, er det klart, at hvis antal forsøg n er stort bliver spredningen lille, og så kan man finde, at der er en signifikant forskel selv om denne forskel er lille.

Imidlertid risikerer man med mange gentagelser at opdage så små forskelle, at de ikke har praktisk betydning, og så er de mange delforsøg jo spild af arbejdskraft og penge.

Endvidere gælder det jo, at hvis man laver 25 forsøg, så er spredningen formindsket med en faktor 5, mens hvis man laver 100 forsøg så er spredningen formindsket med en faktor 10. Der skal derfor særdeles mange forsøg til for alvor at formindske spredningen på gennemsnittet.

Analogt med forklaringen i "Statistiske Grundbegreber" 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 ikke 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 forsøgshederne jo er valgt ved randomisering. Når forsøget så er lavet, kan man (lidt sent) se, om man har skønnet rigtigt.

Dimensioneringen har kun betydning hvis man får en accept, da man så ved, at en eventuel forskel ikke har praktisk betydning. Hvis man får en forkastelse, så ved man der er en signifikant forskel, men om den er af praktisk betydning må en nærmere undersøgelse vise.

Formler for dimensionering af 2 variable findes i oversigten i kapitel 5, afsnit 5.5.1

5 1 Faktor på to niveauer

5.1. NORMALFORDELTE VARIABLE

5.1.1. Test af differens mellem middelværdier

I dette afsnit benyttes et eksempel til at forklare metode, teststørrelse osv. Derefter vises hvorledes det samme eksempel regnes med TI-Nspire/TI89.

Eksempel 5.1. Sammenligning af 2 normalfordelte variable

To produktionsmetoder M1 og M2 ønskes sammenlignet. Der udvælges tilfældigt 20 personer, hvoraf de 10 bliver sat til at arbejde med den ene metode, og de 10 andre med den anden. Efter 2 ugers forløb, beregnede man for hver person det gennemsnitlige tidsforbrug pr. enhed. Da metode 1 er mere kostbar end metode 2, ønsker man kun at gå over til den, hvis tidsforbruget pr. enhed ved metode 1 er mindst 2 minutter mindre end ved metode 2.

Man fik følgende resultater.

M_1	87.8	91.9	89.8	89.0	92.6	89.4	91.4	88.7	90.1	92.4
M_2	92.4	94.6	93.0	94.0	92.4	92.9	96.4	92.1	92.8	93.4

Undersøg på basis af disse resultater, om det på et signifikansniveau på 5% kan påvises at tidsforbruget ved metode M_1 er mindst 2 minutter mindre end ved metode M_2 .

Løsning:

a) Lad X_1 = tidsforbrug ved anvendelse af metode M_1 og

X_2 = tidsforbrug ved anvendelse af metode M_2 .

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

$H_0: \mu_2 - \mu_1 \leq 2$ Begrundelse: Nulhypotesen udtrykker, at påstanden ikke er opfyldt (nul virkning).

$H: \mu_2 - \mu_1 > 2$ Begrundelse: Den alternative metode udtrykker det vi ønsker at bevise, så den angiver, at differensen i middeltidsforbruget er større end 2.

I beregningerne antager man, at $H_0: \mu_2 - \mu_1 = 2$, da en forkastelse så sikrer, at H gælder.

Ti-maskinerne anvender et færdigt program, der anvender en testmetode (Satterthwaites metode), som er robust overfor mindre afvigelser fra kravet om normalitet, når blot antallet af gentagelser er den samme.

Er det ikke tilfældet kan man stadig foretage testen, men så stilles der større krav til, at de variable X_1 og X_2 virkelig er normalfordelte.

Formlen for Satterthwaites metode kan findes i afsnit 5.5 oversigt 5.5.1.

Er den beregnede P-værdi $<$ signifikansniveauet α forkastes H_0 , dvs. vi har bevist den alternative hypotese H er sand. (sandsynligheden for vi dermed kommer med en forkert konklusion er mindre end α).

Er P-værdien $>$ signifikansniveauet α accepteres H_0 , dvs. vi kan ikke på dette grundlag bevise, at H er sand.

Får man en accept og er P - værdien ikke meget større end signifikansniveauet α , så er det muligt at en "stærkere" t - test kunne give en forkastelse.

Denne stærkere t-test kræver imidlertid at de to spredninger kan antages at være ens.

3. 1 Faktor på 2 niveauer

Dette er ofte tilfældet på grund af randomiseringen, men er man i tvivl herom kan man først foretage en test af om spredningerne er ens.

Får man en accept heraf, har man naturligvis ikke hermed vist, at varianserne er ens, men da den følgende t -test af middelværdier er robust overfor mindre forskelle i varianserne, blot vi har samme antal gentagelser, er det tilladeligt i den følgende test af middelværdierne, at antage at varianserne er ens. ◆

Forklaring på formler

For hver af de 2 metoder udføres en række delforsøg. Lad antallet af forsøg være henholdsvis n_1 og n_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 .

Nulhypotese $H_0: \mu_1 - \mu_2 = d$ (d er i eksemplet -2), eller $H_0: \mu_1 - \mu_2 - d = 0$,

Testproceduren baseres på fordelingen af differensen $Y = \bar{X}_1 - \bar{X}_2 - d$.

Ifølge additionssætningen (se eventuelt "Statistiske Grundbegreber) er Y normalfordelt og fra regnereglerne fås

$$E(\bar{X}_1 - \bar{X}_2 - d) = E(\bar{X}_1) - E(\bar{X}_2) - d = \mu_1 - \mu_2 - d \quad \text{og} \quad V(\bar{X}_1 - \bar{X}_2 - d) = V(\bar{X}_1) + V(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

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.

Teststørrelsen $U = \frac{\bar{X}_1 - \bar{X}_2 - d}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ gælder kun, hvis spredningerne σ_1 og σ_2 er kendt eksakt.

Kendes kun deres estimater s_1 og s_2 må der anvendes andre testprocedurer.

Får man ved en F-test en accept af, at varianserne er ens, pooles varianserne sammen til en fælles gennemsnitlig varians s_0^2 (se eventuelt oversigten i afsnit 5.5.1) og størrelsen $t = \frac{\bar{X}_1 - \bar{X}_2 - d}{\sqrt{\frac{s_0^2}{n_1} + \frac{s_0^2}{n_2}}} = \frac{\bar{X}_1 - \bar{X}_2 - d}{s_0 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ er nu t -fordelt.

Hvis stikprøvestørrelserne er store (begge over 30) er det dog tilstrækkelig nøjagtigt at anvende

$U = \frac{\bar{X}_1 - \bar{X}_2 - d}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ som teststørrelse.

I modsat fald kan anvendes en ret kompliceret procedure, der kaldes "Satterhwaite's metode". Denne er beskrevet i afsnit 5.5.1. ◆

Eksempel 5.1. fortsat

Hypoteserne omskrives til

$$H_0: \mu_1 + 2 = \mu_2 \quad H: \mu_1 + 2 < \mu_2$$

TI-Nspire: Lister og regneark ► Udfyld lister med overskrift m1 og m2 ► Statistik ► statistiske test ► t-test for 2 middelværdier ► menu: data ► ok ► menu: List1: benyt pil til at vælge "m1" og skriv +2 ► List 2: Vælg "m2" ► "alternative Hyp" $\mu_1 < \mu_2$ ► samlet: nej ► ok

TI89: APPS, STAT/LIST, indtast data i list1 og list 2 ► F6, 4: ► 2 - SampTtest ► ENTER ► I den fremkomne menu vælg Data ► ok I menu for "list 1" skrives list1+2, I menu for "list 2" skrives list2 for "alternative Hyp" $\mu_1 < \mu_2$ og pooled til "NO" OK

Titel	t-test for to mi...
Alternativ...	$\mu_1 < \mu_2$
t	-1.62832
PVal	0.060926
df	16.9946
\bar{x}_1	92.31
\bar{x}_2	93.4
sx1	1.66897
sx2	1.30213
n1	10.
n2	10.

Man får P-værdi = 0.0609.

Da P-værdi = 6.09% > 5% accepteres H_0 , dvs. det er ikke muligt på dette grundlag at bevise, at tidsforbruget ved metode M_1 er 2 minutter mindre end ved metode M_2 .

Da P-værdien var så tæt ved 5%, vil vi nu forsøge med den stærkere t-test, hvor kravet er en fælles "poolet" spredning

Vi gør som før, men retter nu samlet (poolet) til ja (yes).

Vi får som forventet en lidt mindre P-værdi = 0.0604, men det giver stadig en accept, så konklusionen er den samme.

Da P-værdien er så tæt ved 5% er der en god mulighed for, at tidsforbruget ved metode 1 faktisk er 2 minutter mindre end ved metode 2 (begår en fejl af type 2).

Det ville derfor være rimeligt at bede om at få foretaget flere forsøg.

Havde vi fået en P-værdi < 0.05, så ville næste træk være at teste om spredningerne var ens (se eventuelt eksempel 5.2 hvordan), da vi ellers ikke måtte have anvendt den sidste metode. ◆

Eksempel 5.2 (oprindelige data ikke kendt)

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.

Man valgte at udføre 100 forsøg med hvert additiv.

Man fandt for A_1 : $\bar{x}_1 = 118.6$ og $s_1 = 12.13$, og for A_2 : $\bar{x}_2 = 133.2$ og $s_2 = 14.2$

- 1) 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) Forudsat man får et bevis herfor, skal man angive et konfidensinterval for differensen.

Løsning:

$$1) H_0: \mu_1 + 10 = \mu_2 \quad H: \mu_1 + 10 < \mu_2$$

Statistik ► statistiske tests ► t-test for 2 middelværdier ► menu:statistik ► ok ► menu: i xbar1skriv 118.6 + 10 og udfyld resten ► "alternative Hyp" $\mu_1 < \mu_2$ ► samlet: nej ► ok

3. 1 Faktor på 2 niveauer

tTest_2Samp 118.6+10,12.13,100,133.2,14.2,100,-1,0: stat.results

"Titel"	"t-test for to middelværdier"
"Alternativ Hyp"	" $\mu_1 < \mu_2$ "
"t"	-2.46311
"PVal"	0.007324
"df"	193.28
" \bar{x}_1 "	128.6
" \bar{x}_2 "	133.2
"sx1"	12.13
"sx2"	14.2
"n1"	100.
"n2"	100.

P-værdi = 0.00632 = 0.63%

Da P-værdi < 5% forkastes H_0 , dvs vi har bevist, at man skal anvende A_2 .

2) Et 95% konfidensinterval for differensen:

TI-Nspire:

Statistik ► Konfidensintervaller ► t-interval for 2 middelværdier ► Statistik ► Menu udfyldes herunder "samlet" til nej

tInterval_2Samp 118.6,12.13,100,133.2,14.2,100,0.95,0: stat.results

"Titel"	"t-interval for to middelværdier"
"CLower"	-18.2834
"CUpper"	-10.9166
" \bar{x} Diff"	-14.6
"ME"	3.68341
"df"	193.28
" \bar{x}_1 "	118.6
" \bar{x}_2 "	133.2
"sx1"	12.13
"sx2"	14.2
"n1"	100.
"n2"	100.

95% konfidensinterval for differens mellem middelleveid for $A_2 - A_1$: [10.92 ; 18.28] ◆

Test af Varians

Som det ses af eksempel 5.1 kan det sjældent blive nødvendigt at benytte den lidt stærkere metode, da P-værdien nok bliver mindre, kun sjældent så meget, at det har betydning for konklusionen.

Imidlertid vil det i de næste kapitler være nødvendigt at teste om spredningen på 2 stikprøver er ens.

Dette sker ved en såkaldt F - test hvor teststørrelsen er $F = \frac{s_1^2}{s_2^2}$

F-fordelingen er beskrevet i afsnit 5.5.2, hvor også fremgangsmåden ved testningen er angivet. Som det fremgår af eksempel 5.2 har TI-maskinerne færdige programmer til testningen, så det er ikke nødvendigt at anvende teststørrelsen direkte.

Eksempel 5.2 . Test af varians

Samme problem som i eksempel 5.1

Undersøg ved en test på signifikansniveau på 5% om de to metoders varians er ens.

Løsning:

$$H_0: \sigma_1^2 = \sigma_2^2 \text{ mod } H: \sigma_1^2 \neq \sigma_2^2$$

TI-Nspire: Statistik ► Statistiske test ► F-test for 2 spredninger ► data ► menu udfyldes ► ENTER

TI89: F6 ► 9:2 - SampFtest ► ENTER, ► I menuen vælg Data Input Method= Data ► ENTER

Menuen udfyldes, og man vælger "alternative Hyp = $\sigma_1 \neq \sigma_2$ ", ENTER

Titel	F-test for to spred...
Alternativ...	$\sigma_1 \neq \sigma_2$
F	1.64279
PVal	0.471108
dfNumer	9.
dfDenom...	9.
sx1	1.66897
sx2	1.30213
$\bar{x}1$	90.31
$\bar{x}2$	93.4
n1	10.
n2	10.

I udskrift findes P - værdi = 0.4711

Da P - værdi = 0.471 > 0.05 accepteres H_0 , dvs. vi vil i den følgende test antage, at spredningerne er ens.

**Dimensionering.****Eksempel 5.3 (fortsættelse af eksempel 4.2)**

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.

a) Hvor mange delforsøg 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 $\sigma = 12$ 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. 12 minutter.

$$a) \text{ Af formlerne i oversigten i kapitel 5.5.1 fås } n \geq 2 \cdot \left(\frac{u_{1-\alpha} + u_{1-\beta}}{\frac{\Delta}{\sigma}} \right)^2 = 2 \cdot \left(\frac{u_{0.95} + u_{0.90}}{\frac{10}{12}} \right)^2 = 2 \cdot \left(\frac{1.645 + 1.282}{\frac{10}{12}} \right)^2 = 24.67$$

dvs. der skal udføres i alt $n = 25$ delforsøg af hver behandling

$$b) \text{ Idet første led er 24.67 fås } n = 24.67 \cdot \left(\frac{t_{1-\alpha}(2n-2)}{u_{1-\alpha}} \right)^2$$

$$\text{TI89+TI-Nspire: solve}(x=24.67 \cdot (\text{inv. t}(0.95, (2 \cdot x - 2)) / \text{invNorm}(0.95))^2, x) \mid x > 24$$

Der kan nu gå op til 1 minut før der findes en løsning $x = 25.6$ svarende til $n \approx 26$

**5.1.2 Parvise observationer (blokforsøg)**

Hvis observationerne i stikprøven udviser meget stor spredning, kan en test sandsynligvis ikke vise signifikans. Imidlertid er det måske så muligt at nedsætte spredningen ved at sammenligne observationerne to og to (i par). Man siger også, at man opdeler i blokke.

Det følgende eksempel illustrerer fremgangsmåden

Eksempel 5.5. Parvise observationer

En producent af malerverer har laboratorieresultater, der tyder på, at en ny lak A, har en større slidstyrke end den sædvanlige lak B. Han ønsker en afprøvning i praksis og aftaler med ejerne af 6 bygninger med mange trapper, at han må lakere deres trapper.

Da der er meget forskelligt hvor mange personer der går på trapperne i de forskellige bygninger (sammenlign blot sliddet på en skole og et plejehjem) vælger man at foretage et "blokforsøg", med de 6 bygninger som 6 blokke.

I hver bygning lakeres hverandet trin (valgt ved lodtrækning) med lak A og resten med lak B. Efter 3 måneders forløb måles graden af slid (i %) i hver bygning.

De målte værdier af slid efter valg af plan var

Bygning nr	1	2	3	4	5	6
Ny lak	20.3	25.1	21.8	19.6	18.9	23.5
Sædvanlig lak	19.5	28.4	21.6	22.0	20.9	25.8

Undersøg om observationerne leverer et eksperimentelt bevis for, at den nye lak er mere slidstærk end den sædvanlige lak.

Løsning

Vi ser nu på differensen mellem sliddet i en bygning. (hvorved den store forskel mellem bygningerne elimineres)

Lad $D = X_{\text{gammel}} - X_{\text{ny}}$

D antages normalfordelt $n(\mu, \sigma)$, hvor såvel μ som σ er ukendte.

Da vi ønsker at teste om ny lak er mere slidstærk end gammel lak, dvs. den mest slidstærke lak slides mindst, bliver testen en ensidet t -test.

Nulhypotese $H_0: \mu = 0$ Alternativ hypotese $H: \mu > 0$.

TI-Nspire: Lister og regneark ► Udfyld to lister "ny" og "gammel" med de givne data ► kald tredje liste d ► skriv nedenunder gammel - ny ► Statistik ► t-interval for 1 middelværdi ► menu:data ► ok ► udfyld menu ► ok

TI89: Indtaster data i to lister n og g ► kald tredje liste for z og skriv (nederst) g-n ► ► APPS, STAT/LIST ► F6, ► 2: t-test ► vælg Data ► menu udfyldes

A ny	B gammel	C d	D	E	F
		=gammel-ny			=tTest(0,'d,1,1): CopyVar S
20.3	19.5	-0.8		Titel	t-test for én middelværdi
25.1	28.4	3.3		Alternat...	$\mu > \mu_0$
21.8	21.6	-0.2		t	2.26822
19.6	22	2.4		PVal	0.0363
18.9	20.9	2.		df	5.
23.5	25.8	2.3		\bar{x}	1.5
				sx := S _n -...	1.61988
				n	6.

Man finder $\bar{x} = 1.5$ og P -værdi = $P(T > 1.5) = 0.0363$

Da P -værdi = 3.63% < 5% forkastes H_0 (svagt), dvs. den ny lak er mere slidstærk end den gamle.



5.2 BINOMIALFORDELTE VARIABLE.

Man kan ofte approksimere en binomialfordeling med en normalfordeling.

Det er en sådan approksimation, som såvel formlerne i oversigten i afsnit 5.5.3 bygger på som TI.. programmerne

Forudsætningen er:

X_1 og X_2 er **binomialfordelt** henholdsvis $b(n_1, p_1)$ og $b(n_2, p_2)$

Observerede stikprøveværdier x_1 og x_2 .

Lad $\hat{p}_1 = \frac{x_1}{n_1}$, $\hat{p}_2 = \frac{x_2}{n_2}$, $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$. **Forudsætning:** $n_1 \cdot \hat{p} \in [5; n_1 - 5]$, $n_2 \cdot \hat{p} \in [5; n_2 - 5]$

I praksis vil disse forudsætninger for approksimation sædvanligvis være opfyldt.

Vi belyser beregningerne ved følgende eksempel.

Eksempel 5.6. 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 er 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) $H_0: p_1 = p_2$ mod $H: p_1 < p_2$ (da vi ønsker at vise, at $p_1 < p_2$).

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

$$\text{Da } n_1 \cdot \hat{p} = 139 \cdot \frac{48}{279} = 23,9 \in [5; 139 - 5] \text{ og } n_2 \cdot \hat{p} = 140 \cdot \frac{48}{279} = 24,1 \in [5; 140 - 5]$$

er forudsætningerne for at approksimere med normalfordelingen opfyldt.

TI-Nspire: Lister Og regneark ► Statistik ► Statistiske test ► z test for 2 andele ► Udfyld menu:

Succes $x_1=17, n_1=139$, succes $x_2=31, n_2=140$, Alt. hyp

: $p_1 < p_2$ ► ok

TI89: F6, 6 ► 2-Prop-ZTest ► Udfylder menu: Succes $x_1=17, n_1=139$, succes $x_2=31, n_2=140$,

Alt. hyp : $p_1 < p_2$

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

=zTest_2Prop(17,139,31,140,-1):	
Titel	z-test for to andele
Alternativ Hyp...	$p_1 < p_2$
z	-2.19349
PVal	0.014136
\hat{p}_1	0.122302
\hat{p}_2	0.221429
\hat{p}	0.172043
n1	139.
n2	140.

3. 1 Faktor på 2 niveauer

Konklusion:

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

2) 95% konfidensinterval;

TI-Nspire: Statistik ► Konfidensintervaller ► z-interval for 2 andele ► menu som under punkt 1.

TI89: F7, 6 ► 2-Prop-ZInt , menu udfyldes som under punkt 1.

$$C_int = [-0.187 ; -0.0111] \quad \text{eller} \quad [0.011; 0.187]$$

Titel	z-interval for...
CLower	-0.186859
CUpper	-0.011394
pDiff	-0.099126
ME	0.087733
p1	0.122302
p2	0.221429
n1	139.
n2	140.



5.3 POISSONFORDELTE VARIABLE.

Man kan ofte approksimere en Poissonfordeling med en normalfordeling.

Det er en sådan approksimation, som formlerne i oversigten i afsnit 5.5.4 bygger på. I praksis vil disse forudsætninger for approksimation sædvanligvis være opfyldt. Vi belyser anvendelsen af oversigten ved følgende eksempel.

Eksempel 5.7. 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 20 TV-apparater, og man optalte antallet af loddefejl i de 20 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 $p(\mu_2)$.

$H_0: \mu_1 = \mu_2$ mod $H: \mu_1 \neq \mu_2$ (da vi ønsker at vise, at $\mu_1 \neq \mu_2$)

Oversigten i kapitel 5.5.4 anvendes :


$$\bar{x}_1 = \frac{12}{20}, \bar{x}_2 = \frac{7}{19}, \text{ og } \bar{x} = \frac{20 \cdot \frac{12}{20} + 19 \cdot \frac{7}{19}}{20 + 19} = \frac{12 + 7}{20 + 19} = \frac{19}{39}$$

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

er forudsætningerne for at approksimere med normalfordelingen opfyldt.

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

Da $P\text{-værdi} = P(Y > 0.2316) = \text{normCdf}(0.2316, \infty, 0, 0.2236) = 0.1501 > 0.025$ accepteres nulhypotesen.

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

5.4. FORDELING UKENDT (Rangtest)

De testprocedurer vi har benyttet i de forrige kapitler har alle været baseret på, at i det mindste approksimativt kendte fordelingen (normal-, binomial- eller Poisson-fordelt) og testen vedrørte parametre i fordelingen såsom μ , σ eller p . Denne form for statistik kunne kaldes "parametrisk statistik". Kendes fordelingen ikke, og kan man heller ikke approksimere den til en kendt fordeling, så må man benytte de såkaldte "ikke-parametriske test". Disse forudsætter ikke, at fordelingen er kendt, og kunne derfor også kaldes "fordelingsfri test".

Det ligger udenfor denne bogs centrale emner, men der findes en beskrivelse heraf i kapitel 10.

5.5. OVERSIGT OVER CENTRALE FORMLER I KAPITEL 5

5.5.1. Test og konfidensinterval af differens $\mu_1 - \mu_2$ mellem middelværdier μ_1 og μ_2 for 2 normalfordelte 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 . Lad d være en given konstant.

Satterthwaites test: Forudsætning σ_1 og σ_2 ukendte

Forkortelser: $a = \frac{s_1^2}{n_1}$, $b = \frac{s_2^2}{n_2}$, $c = a + b$, $g = \frac{c^2}{\frac{a^2}{n_1-1} + \frac{b^2}{n_2-1}}$ T er t - fordelt med frihedsgradstallet f .				
Forudsætninger	Alternativ hypotese H	P - værdi	Beregning	H_0 forkastes
σ_1, σ_2 ukendte $t = \frac{\bar{x}_1 - \bar{x}_2 - d}{\sqrt{c}}$ f er det nærmeste hele tal, som er større end g	$\mu_1 > \mu_2 + d$	$P(T \geq t)$	$tCdf(t, \infty, f)$ eller TI:-sampTtest se eksempel 5.1	$P - værdi < \alpha$
	$\mu_1 < \mu_2 + d$	$P(T \leq t)$	$tCdf(-\infty, t, f)$ eller TI:-sampTtest se eksempel 5.1	
	$\mu_1 \neq \mu_2 + d$	$P(T \geq t)$ for $\bar{x}_1 > \bar{x}_2 + d$ $P(T \leq t)$ for $\bar{x}_1 < \bar{x}_2 + d$	som række 1 som række 2	$P - værdi < \frac{1}{2} \alpha$ dog hvis Ttest $P - værdi < \alpha$
100 · (1 - α)% konfidensinterval for differens $\mu_1 - \mu_2$: $\bar{x}_1 - \bar{x}_2 - t_{1-\frac{\alpha}{2}}(f) \cdot \sqrt{c} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{1-\frac{\alpha}{2}}(f) \cdot \sqrt{c}$ t-interval for 2 middelværdier eller 2-SampTint. se eksempel 5.4				

Z-test (eller U-test): Forudsætning: σ_1 og σ_2 kendt eksakt

Forkortelser: $\bar{x} = \bar{x}_1 - \bar{x}_2 - d$ $\sigma = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ Y er normalfordelt $n(\mu, \sigma)$				
Forudsætninger	Alternativ hypotese H	P - værdi	Beregning	H_0 forkastes
σ_1, σ_2 kendte	$\mu_1 > \mu_2 + d$	$P(Y \geq \bar{x})$	$normCdf(\bar{x}, \infty, \mu, \sigma)$ eller: t-interval for 2 middelv. 2-sampZtest	$P - værdi < \alpha$
	$\mu_1 < \mu_2 + d$	$P(Y \leq \bar{x})$	$normCdf(-\infty, \bar{x}, \mu, \sigma)$ eller :t-interval for 2 middelv 2-sampZtest	
	$\mu_1 \neq \mu_2 + d$	$P(Y \geq \bar{x})$ for $\bar{x}_1 > \bar{x}_2 + d$ $P(Y \leq \bar{x})$ for $\bar{x}_1 < \bar{x}_2 + d$	som række 1 som række 2	$P - værdi < \frac{1}{2} \alpha$ dog hvis Ztest $P - værdi < \alpha$
100 · (1 - α)% konfidensinterval for differens $\mu_1 - \mu_2$: $\bar{x}_1 - \bar{x}_2 - u_{1-\frac{\alpha}{2}} \cdot \sigma \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + u_{1-\frac{\alpha}{2}} \cdot \sigma$ 2-SampZint				

Parvise observationer (blokforsøg)**t-Test: Forudsætning: Forsøgene udføres "parvist"** (jævnfør eksempel 5.5).

Man danner differensen mellem tallene i hvert par (blok)

Lad D være denne differens, og lad D have middelværdi μ og spredning σ T er en stokastisk variabel der er t -fordelt med $f = n - 1$.

Forudsætninger	Alternativ hypotese H	P -værdi	Beregning	H_0 forkastes
σ ukendt. $t = \frac{(\bar{x} - \mu_0) \cdot \sqrt{n}}{s}$	$H: \mu > \mu_0$	$P(T \geq t)$	$tCdf(t, \infty, n-1)$ eller t-test for 1 middelværdi. eller F6: t-test	P -værdi $< \alpha$
	$H: \mu < \mu_0$	$P(T \leq t)$	$tCdf(-\infty, t, n-1)$ eller t-test for 1 middelværdi eller F6: t-test	
	$H: \mu \neq \mu_0$	$P(T \geq t)$ for $\bar{x} > \mu_0$ $P(T \leq t)$ for $\bar{x} \leq \mu_0$	som række 1 som række 2	P -værdi $< \frac{1}{2}\alpha$ dog hvis t-test P -værdi $< \alpha$

Dimensionering: $\Delta = |\mu_1 - \mu_2|$ er den mindste ændring der har praktisk interesse. $\alpha = P(\text{type I fejl})$, $\beta = P(\text{type II fejl})$ **Spredninger ukendt**, men antages at være nogenlunde ens og højst af størelsen σ

$$\text{Ensided test: } n \geq 2 \cdot \left(\frac{u_{1-\alpha} + u_{1-\beta}}{\frac{\Delta}{\sigma}} \right)^2 \cdot \left(\frac{t_{1-\alpha}(2n-2)}{u_{1-\alpha}} \right)^2.$$

Tosidet test: I ovenstående formel erstattes α med $\frac{\alpha}{2}$ **Spredninger kendt**: Samme formel, idet man blot stryger sidste faktor med t -leddet.

3. 1 Faktor på 2 niveauer

5.5.2. Test af differens mellem varianser for to normalfordelte 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 .

Forkortelser: $F = \frac{s_1^2}{s_2^2}$, Q er F - fordelt $F(n_1 - 1, n_2 - 1)$				
Forudsætninger	Alternativ hypotese H	P - værdi	Beregning	H ₀ forkastes
μ_1, μ_2 ukendte	$H: \sigma_1^2 > \sigma_2^2$	$P(Q \geq F)$	FCdf($F, \infty, n_1 - 1, n_2 - 1$) eller F-test for 2 spredninger eller F6: 2Sampl F test	$P - værdi < \alpha$
	$H: \sigma_1^2 < \sigma_2^2$	$P(Q \leq F)$	FCdf($-\infty, F, n_1 - 1, n_2 - 1$) eller F-test for 2 spredninger eller F6: 2Sampl F test	
	$H: \sigma_1^2 \neq \sigma_2^2$	$P(Q \geq F)$ for $F > 1$ $P(Q \leq F)$ for $F < 1$	som række 1 som række 2	$P - værdi < \frac{1}{2} \alpha$ dog hvis Ftest $P - værdi < \alpha$
<p>100 · (1 - α)% Konfidensinterval for forhold $\frac{\sigma_1^2}{\sigma_2^2}$:</p> $\frac{F}{F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq F \cdot F_{1-\frac{\alpha}{2}}(n_2 - 1, n_1 - 1)$				

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

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 $F(n_1, n_2 - 1)$

F - fordelingen

Til enhver F - fordeling er knyttet to hele positive tal f_T og f_N kaldet henholdsvis tællerfrihedsgradstallet og nævnerfrihedsgradstallet. En sådan fordeling benævnes $F(f_T, f_N)$.

Det kan vises (se evt "afsnit til statistiske grundbegreber" 3C), at den variable $F = \frac{s_1^2}{s_2^2}$ er F - fordelt $F(f_1, f_2)$, hvor

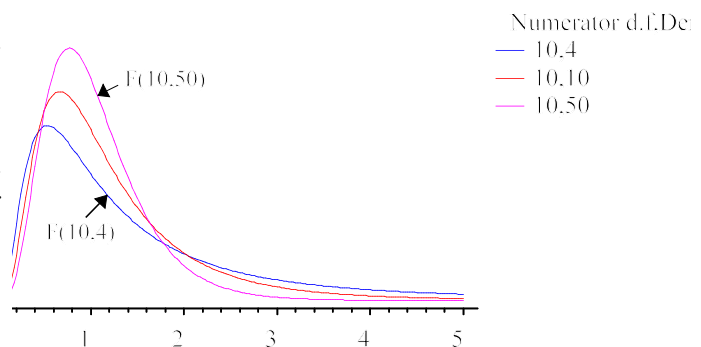
f_1 og f_2 er de til s_1^2 og s_2^2 hørende frihedsgradstal.

Dette benyttes til testning af hypoteser vedrørende forholdet mellem to varianser.

På figuren er afbildet tæthedsfunktionen for F - fordelingerne

$F(10,4)$, $F(10,10)$ og $F(10,50)$.

Det ses, at F fordelinger kun er defineret for tal større end eller lig nul, og at F-fordelinger ikke er symmetriske.



5.5.3. Test og konfidensinterval af $p_1 - p_2$ for 2 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 .

Forkortelser: $\hat{p}_1 = \frac{x_1}{n_1}$, $\hat{p}_2 = \frac{x_2}{n_2}$, $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$, $s = \sqrt{\hat{p} \cdot (1 - \hat{p}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ Y er normalfordelt variabel $n(0,s)$				
Forudsætninger	Alternativ hypotese H	P - værdi	Beregning	H_0 forkastes
$n_1 \cdot \hat{p} \in [5; n_1 - 5]$ $n_2 \cdot \hat{p} \in [5; n_2 - 5]$	$H: p_1 > p_2$	$P(Y > \hat{p}_1 - \hat{p}_2)$	$NormCdf(\hat{p}_1 - \hat{p}_2, \infty, 0, s)$ eller z-test for 2 andele eller F6: 2-Prop Z test	P - værdi $< \alpha$
	$H: p_1 < p_2$	$P(Y < \hat{p}_1 - \hat{p}_2)$	$NormCdf(-\infty, \hat{p}_1 - \hat{p}_2, 0, s)$ eller z-test for 2 andele eller F6: 2-Prop Z test	
	$H: p_1 \neq p_2$	$P(Y > \hat{p}_1 - \hat{p}_2)$ for $\hat{p}_1 \geq \hat{p}_2$ $P(Y < \hat{p}_1 - \hat{p}_2)$ for $\hat{p}_1 < \hat{p}_2$	som række 1 som række 2	P - værdi $< \frac{1}{2} \alpha$ dog hvis 2PropZtest, så P - værdi $< \alpha$
100 · (1 - α)% konfidensinterval for differens $p_1 - p_2$ $\hat{p}_1 - \hat{p}_2 - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ eller z interval for 2 andele eller 2 prop- z- Interval				

5.5.4 Test og konfidensinterval af $\mu_1 - \mu_2$ for 2 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 .

Forkortelser: $\bar{x} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2}{n_1 + n_2}$ og $s = \sqrt{\bar{x} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ Y er normalfordelt variabel $n(0,s)$				
Forudsætninger	Alternativ hypotese H	P - værdi	Beregning	H_0 forkastes
$n_1 \cdot \bar{x} \geq 5$ $n_2 \cdot \bar{x} \geq 5$	$H: \mu_1 > \mu_2$	$P(Y > \bar{x}_1 - \bar{x}_2)$	$NormCdf(\bar{x}_1 - \bar{x}_2, \infty, 0, s)$	P - værdi $< \alpha$
	$H: \mu_1 < \mu_2$	$P(Y < \bar{x}_1 - \bar{x}_2)$	$NormCdf(-\infty, \bar{x}_1 - \bar{x}_2, 0, s)$	
	$H: \mu_1 \neq \mu_2$	$P(Y > \bar{x}_1 - \bar{x}_2)$ for $\bar{x}_1 \geq \bar{x}_2$ $P(Y < \bar{x}_1 - \bar{x}_2)$ for $\bar{x}_1 < \bar{x}_2$	som række 1 som række 2	P - værdi $< \frac{1}{2} \alpha$
100 · (1 - α)% konfidensinterval for differens $\mu_1 - \mu_2$. $\bar{x}_1 - \bar{x}_2 - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{x}_1}{n_1} + \frac{\bar{x}_2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\bar{x}_1}{n_1} + \frac{\bar{x}_2}{n_2}}$				

OPGAVER

Opgave 5.1

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 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 5.2

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 diæten 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:

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

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

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

Opgave 5.3

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 5.4

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
--------	--------	--------	--------	--------	--------	--------	----------	--------	--------

- a) Analyser forsøgsresultaterne.
- b) Hvis der er en signifikant forskel skal der opstilles et 95%-konfidensinterval for differensen $\mu(M) - \mu(N)$ mellem de to middelværdier.

Opgave 5.5

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.

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. (bemærk: antal forsøg er forskellige)
- For en sikkerheds skyld skal man også kontrollere om varianserne er rimelig ens.
- Hvis forslaget er velmotiveret skal der opstilles et 95% - konfidensinterval for differensen mellem middeludbytteerne ved benyttelse af metoderne I og II.

Opgave 5.6

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 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

1) og man ved, at spredningen $\sigma = 0.2$ ohm.

2) og man har en forhåndsformodning om, at spredningen er ca. $\sigma = 0.2$ ohm.

3) Hvilke konklusioner vedrørende behandlingernes virkning kan gøres, såfremt man ved testning af forsøgsresultaterne finder

- signifikans
- ingen signifikans

4) Hvilke yderligere analyser af forsøgsresultaterne bør foretages, såfremt testningen

- viser signifikans
- ikke viser signifikans.

Opgave 5.7

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 $n =$ "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 5.8

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 5.9

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 5.10

I forbindelse med måling af luftforurening benyttes to apparater A og B til måling af mængden svovlmonoxid. Følgende målinger er foretaget daglig i en periode på 2 uger.

dag	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	0.96	0.82	0.75	0.61	0.89	0.64	0.81	0.68	0.65	0.84	0.59	0.94	0.91	0.77
B	0.91	0.78	0.67	0.59	0.80	0.74	0.73	0.61	0.57	0.92	0.55	0.83	0.88	0.67

Man ønsker at undersøge om de to apparater i middel giver forskellige resultater på et signifikansniveau på 5%..

Angiv forudsætningerne for at foretage en relevant test og udfør denne.

Opgave 5.11

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 5.12

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 5.13

- 1) Mange forbrugere tror, at såkaldte "mandagsbiler", dvs. biler produceret om mandagen, har flere alvorlige fejl (alvorlig fejl kan der kun være en af pr. bil) 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 (enkeltfejl kan der godt være mange af på en bil), 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 5.14

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 5.15

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.

6 1 FAKTOR PÅ MERE END 2 NIVEAUER

6.1 INDLEDNING

I kapitel 3 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 6.1 (én faktor på 4 niveauer).

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	116	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	116	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 naturligtvis ikke har mening at se på mellemliggende værdier (såsom tilsætningsstof nr. $T_{1.53}$).

Problemer, hvor faktorerne er kvalitative analyseres ved en "variationsanalyse. 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 variationsanalyseteknikken, men da findes der mere effektive metoder, som dog ikke behandles i denne bog.

6.2 ENSIDET VARIANSANALYSE (normalfordelte variable)

6.2.1. Indledning

Vi vil i dette afsnit behandle problemer, af den type, som er vist i eksempel 6.1, dvs. med én faktor på mere end 2 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 6.1 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 3, 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.

Som altid tilstræber man, at antallet af gentagelser af hvert niveau er det samme.

I appendix 6.1 er angivet de generelle formler, som gælder, selv om eksempelvis et forsøg er mislykket, så man ikke har samme antal gentagelser.

TI-programmerne anvender naturligvis disse generelle formler til beregningerne.

I afsnit 6.2.2 gives en forståelse for den teoretiske baggrund for variansanalyser.

Forklaringen understøttes af eksempel 6.1, hvor der vises, hvorledes man kan foretage beregningerne med en lommeregner eller PC, der kun har de sædvanlige sandsynlighedsfunktioner som normal- t- og F-fordeling, men ikke et egentligt statistikprogram.

Vi nøjes dog her med at se på hovedtilfældet, hvor antallet af gentagelser på hvert niveau er den samme.

6.2.2. Forklaring af metode og formler

Eksempel 6.1 "håndregning"

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 af urenhed	108	105	116	117
	110	110	111	119
	112	109	113	112

Find om muligt det tilsætningsstof der i middel giver den mindste urenhed og angiv i bekræftende fald et 95% konfidensinterval for middelværdien. ◆

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 en af de øvrige".

Forsøgets udførelse. Forsøget skal udføres som et **fuldstændigt randomiseret forsøg**. (jævnfør kapitel 2 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 forsøgsheder.

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	113.33	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.

2) De 4 variable T_1, T_2, T_3 og T_4 skal være tilnærmelsesvis normalfordelte.

Mindre afvigelser er tilladeligt. Man siger, at analysen er robust overfor afvigelser fra normalitet. Mest robust er den, hvis hvert niveau har samme antal gentagelser (som i eksempel 6.1, hvor alle har 3 gentagelser).

3) Der skal være varianshomogenitet.

Det betyder i eksempel 6.1, at målingerne af urenheder med tilsætningsstof T_1 , skal have tilnærmelsesvis samme varians som målingerne af urenheder med T_2 , T_3 og T_4 .

Også her gælder det, at analysen er robust overfor mindre afvigelser, og mest robust, hvis antallet af gentagelser er det samme.

Sædvanligvis er forudsætningerne rimeligt opfyldt, hvis man har i forsøgsplanen har foretaget en randomisering.

Kun hvis man er i alvorlig tvivl vil man foretage en kontrol.

Kontrol af normalitet.

I appendix 6.1 er vist hvorledes man kan lave et normalfordelingsplot af residualerne. Hvis disse ligger nogenlunde på en ret linie, er betingelsen opfyldt.

Kontrol af varianshomogenitet

Kan man få en accept af nulhypotesen

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$ $H: \text{Mindst en varians er forskellig fra de øvrige}$

på et signifikansniveau på 5% (nogle mener 1% er ok hvis der er samme antal gentagelser) så er kravet rimeligt opfyldt.

Der findes forskellige forslag til hvorledes man tester dette. I appendix 6.1 er angivet formlerne for 2 af disse, nemlig "Bartletts test" og "Levines test".

De er imidlertid så beregningsmæssigt kompliceret, at man nok altid vil benytte et egentlig statistikprogram som eksempelvis SAS.JMP til beregningerne.

I bogen "Videregående statistik" hvor SAS.JMP anvendes kan man i eksempel 6.2 bl.a. se beregnet Bartletts test, sammen med et par andre test.

"Simplificerede F-test"

Denne er ret nem at beregne. Man danner forholdet mellem den største varians $\sigma_{\max}^2 = 3.6056^2$ og den mindste varians $\sigma_{\min}^2 = 2^2$

$H_0: \sigma_{\max}^2 = \sigma_{\min}^2$ $H: \sigma_{\max}^2 \neq \sigma_{\min}^2$

$F = \frac{3.6056^2}{2^2} = 3.2501$ P-værdi = $P(F > 3.2501) = \text{FCdf}(3.2501, \infty, 2, 2) = 0.235$

Da testen er en tosidet test, sammenlignes P-værdien med $\frac{\alpha}{2} = 0.025$

Da P-værdi = $0.235 > 0.025$ accepteres H_0 , dvs., vi kan ikke afvise der er varianshomogenitet.

Vil vi derfor i det følgende gå ud fra, at der er en rimelig varianshomogenitet.

Får man en forkastelse betyder det ikke, at der ikke er varianshomogenitet. Metoden tager nemlig ikke hensyn til de mellemliggende varianser.

Giver alle disse test en forkastelse, kan man eksempelvis prøve at tage logaritmen til alle tal i et forsøg på at nedsætte spredningen (vil ikke blive gennemgået her).

Er forudsætningerne ikke opfyldt kan man i stedet udføre en såkaldt "Kruskal Wallis" "rangtest" (se evt. appendix 11.1).

Når man ikke altid gør det, skyldes det at rangtest er svagere test, dvs. man vil ofte ikke opdage en signifikant forskel.

6. En faktor på mere end 2 niveauer

Pooling.

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

$$s_{error}^2 = \frac{s_1^2 + s_2^2 + s_3^2 + s_4^2}{4} = \frac{2.000^2 + 2.6458^2 + 2.5166^2 + 3.6056^2}{4} = 7.583$$

s_{error}^2 er variansen for "forsøgsfejlen" eller på engelsk "error".

s_{error}^2 har $N - r = 12 - 4 = 8$ 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_{error}^2 har derfor $2 \cdot 4 = 8$ frihedsgrader.

F - test. Antages nulhypotesen $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ 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 $s_R^2 \approx s_{error}^2$ eller $\frac{s_R^2}{s_{error}^2} \approx 1$, mens hvis nulhypotesen er falsk (middel-

værdierne er forskellige) er $s_R^2 > s_{error}^2$ dvs. forholdet $F = \frac{s_R^2}{s_{error}^2}$ være signifikant større end 1.

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

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

$$P\text{-værdi} = P(F > 4.62) = FCdf(4.62, \infty, 3, 8) = 0.0371 = 3.71\%$$

Da P - værdi = 0.037 < 0.05 forkastes nulhypotesen (svagt).

Konklusion: De fire tilsætningsstoffer har ikke samme virkning.

Mindst 2 af middelurenhederne er forskellige.

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.

Konfidensintervaller for middelværdi.

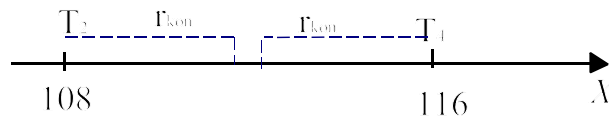
De "sædvanlige" konfidensintervaller for hvert niveau bestemmes ved (jævnfør eventuelt appendix 6.1):

$\bar{x}_i \pm r_{kon}$ hvor $r_{kon} = t_{1-\frac{\alpha}{2}}(f_{error}) \cdot \sqrt{\frac{s_{error}^2}{n}}$, hvor n er det antal gentagelser der blev anvendt til beregning af gennemsnittet.

I vort tilfælde er $r_{kon} = t_{0.975}(8) \sqrt{\frac{7.583}{3}} = 2.31 \cdot 1.589 = 3.673$.

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 måske hinanden en smule, så man ikke opdager det. Idet $2 \cdot r_{kon} = 2 \cdot 2.673 = 7.346$ ses ikke overraskende, at der er forskel på T_2 og T_4 dvs. mellem største og mindste værdi.

Da vi ønske at finde det stof der giver mindst forurening sammenligne nu T_2 med T_1 og T_3 .

Da afstandene er henholdsvis 2 og 5.3 kan de ikke adskilles.

LSD-intervaller (Least Significance Intervals)

Man kunne så foretage parvise sammenligninger svarende til de konfidensintervaller vi fandt i kapitel 3.

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.

LSD - konfidensintervaller, der beror på parvise sammenligninger er bestemt ved, at deres

radius er $r_{LSD} = r_{kon} \frac{\sqrt{2}}{2} = \frac{r_{kon}}{\sqrt{2}}$

Bevis:

Lad $Y = \bar{X}_1 - \bar{X}_2$

Vi har da (af kvadratsætning) at variansen $V(Y) = V(\bar{X}_1) + V(\bar{X}_2) = \frac{s_0^2}{n} + \frac{s_0^2}{n} = \frac{2 \cdot s_0^2}{n}$

Heraf følger, at Y har konfidensintervallet $\bar{x}_1 - \bar{x}_2 \pm r_1$ hvor $r_1 = t_{0.975}(f_0) \sqrt{\frac{2 \cdot s_0^2}{n}}$

Et LSD-interval omkring et enkelt punkt er følgelig $r_{LSD} = \frac{r_1}{2}$

Heraf fås $r_{LSD} = \frac{t_{0.975}(f_0)}{2} \sqrt{\frac{2 \cdot s_0^2}{n}} = \frac{t_{0.975}(f_0)}{2} \sqrt{2} \sqrt{\frac{s_0^2}{n}} = \frac{t_{0.975}(f_0)}{\sqrt{2}} \sqrt{\frac{s_0^2}{n}} = \frac{r_{kon}}{\sqrt{2}}$ ◆

6. En faktor på mere end 2 niveauer

Her vil man kunne opdage en forskel, hvis middelværdiernes afstand er større end

$$2 \cdot r_{lsd} = 2 \cdot \frac{r_{kon}}{\sqrt{2}} = \sqrt{2} \cdot r_{kon} .$$

Da $\sqrt{2} \cdot 3.673 = 5.19$ ses nu, at T_2 giver mindre forurening end T_3 mens T_1 og T_2 stadig ikke kan adskilles.

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.

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

Konklusion: LSD-intervallerne viser, at man får den mindste urenhed, hvis man vælger enten T_1 eller T_2 (de kan ikke adskilles).

Et 95% konfidensinterval for T_2 er $[108 - 3.673; 108 + 3.673] = [104.3 ; 111.7]$

6.2.3. Beregning af ensidet variansanalyse

Eksempel 6.2 (eksempel 6.1 fortsat)

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	116	117
	110	110	111	119
	112	109	113	112

Det forudsættes, at betingelserne om uafhængighed, normalitet og varianshomogenitet er opfyldt.

- Test på signifikansniveau på 5% om der er forskel på middelværdierne for de 4 tilsætningsstoffer
- Find om muligt det tilsætningsstof der i middel giver den mindste urenhed og angiv i bekræftende fald et 95% konfidensinterval for middelværdien.

Løsning:

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

TI-Nspire: Lister og regneark ► Lav 4 lister (husk overskrift) ► Statistik ► statistiske test ► Ensidede variansanalyse ► Data ► 4 ► indsæt navne i menu ► ok

TI89: APPS ► StatList ► Indtast data i 4 lister ► F6, C:ANOVA ► Data Input : Data ► Antal grupper = 4, ENTER, Udfyld listnavne (VAR-Link osv.) ENTER, ENTER.

A	t1	B	t2	C	t3	D	t4	E	F	G	stat2...	H	stat1...	I	stat2...
									=ANOVA('t1,'t2,'t3,'t4						
108	105	116	117	Titel					Ensidet variansana...	110.	106.334	113.666			
110	110	111	119	F					4.96703	108.	104.334	111.666			
112	109	113	112	PVal					0.031084	113.333	109.667	117.			
				df					3.	116.	112.334	119.666			
				SS					113.						
				MS					37.6667						
				dfError					8.						
				SSError					60.6667						
				MSError					7.58333						
				sp					2.75379						
				CLowerLis...					{106.33368709762,...						
				CUpperLis...					{113.66631290238,...						
				xList					{110.,108.,113.333...						

Der fremkommer nu en række resultater, hvor den vigtigste naturligvis er P - Value = 0.031
Da P - værdi = 0.031 < 0.05 forkastes nulhypotesen $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ (svagt)

Konklusion: De fire tilsætningsstoffer har ikke samme virkning.

Mindst 2 af middelurenhederne er forskellige.

- c) **TI-Nspire:** Ud for xbar vises de 4 gennemsnit, tilsvarende for Lower- og Upper list
Ønskes de placeret som lister : Anbring cursor på øverste navnecelle i en søjle og højreklik ►
Vælg variabel ► Kæd til ► vælg xbar ► Enter
Tilsvarende gøres med de to andre lister. Herved fremkommer listerne øverst til højre .

I **TI-Nspire-lommeregner** : Marker en liste foroven, CTRL menu osv.

TI89: De sædvanlige konfidensintervaller findes som ekstra søjler efter list6 (se nedenfor)

xbar	lowlist	uplist
110	106.33	113.67
108	104.33	111.67
113.33	109.67	117.00
116	112.33	119.67

Ønskes beregnet LSD-intervaller må man udnytte, at $2 \cdot r_{kon} = 113.67 - 106.33 = 7.34$

Heraf fås, at $2 \cdot r_{lsd} = \frac{7.34}{\sqrt{2}} = 5.19$

Da afstanden mellem mindste og tredjemindste værdi er $113.33 - 108 = 5.33 > 5.19$ giver T_3 en større forurening end T_2 , mens det umiddelbart ses, at T_1 og T_2 ikke kan adskilles.

Konklusion: Man får den mindste urenhed, hvis man vælger enten T_1 eller T_2 (de kan ikke adskilles).

Et 95% konfidensinterval for T_2 er [104.3 ; 111.7]



Eksempel 6.3 Data ikke opgivet

Et firma, der udvikler elektroniske komponenter, skal anvende nogle batterier til de færdige apparater. Der købes 25 batterier fra hver af 4 forskellige producenter.

Levetiden for hvert enkelt batteri bestemmes og gennemsnit og spredning for hver producent findes i nedenstående tabel.

	Antal batterier	Gennemsnitlig levetid \bar{x} (i timer)	Spredning s i timer
Producent	A	25	12.24
	B	25	14.236
	C	25	11.096
	D	25	9.984

- a) Vis, at levetiden for batterierne afhænger af hvilken producent der vælges. Signifikansniveau 5%. Forudsætningerne for analysen antages opfyldt.
- b) Firmaet ønsker at købe hos den producent hvis batterier har længst levetid. Angiv den/de producenter hvis batterier har længst levetid og angiv et 95% konfidensinterval for disse batterier

Løsning

- a) $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ mod H : "mindst én middelværdierne er forskellig fra en af de øvrige".

TI-Nspire: Lister og regneark ► Indtast n , gennemsnit og spredning i 4 lister a,b,c,d
► Statistik ► statistiske test ► Ensided variansanalyse ► Statistik ► Antal grupper 4 ► udfyld grupper med listenavne ► ok

TI89: APPS ► StatList ► Indtast n , gennemsnit og spredning i 4 lister ► F6, C:ANOVA ► Data Input : Stats ► Antal grupper = 4 ► ENTER, Udfyld grupper med listenavne ► ENTER

Der fremkommer nu en række resultater, hvor den vigtigste naturligvis er P - Value = 0.0002

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

Konklusion: De fire producenters batterier har ikke samme virkning.

Mindst 2 af producenternes batterier er forskellige.

- b) I **TI-Nspire** findes resultaterne som vandrette "lister" nederst i udskriften. Ønskes de som mere overskuelige lister: Marker en liste foroven, højre muse-tast ► variable ► Kæd til ► vælg \bar{x} -list osv.

I **TI-Nspire-lommeregner** : Marker en liste foroven, CTRL menu osv.

I **TI89** beregner nogle lommeregnerne umiddelbart konfidensintervaller som søjler efter listerne (som i eksempel 6.2) mens man for andres vedkommende må gå ned i VarLink og her under STATVARS finde "lowlist" osv. (man kan se indhold ved tryk på F6)

Man finder, at

B har størst gennemsnit på 14.236, A har næststørst på 12.24

Endvidere ses, at diameter i konfidensinterval er $d_{kon} = 13.5633 - 10.9167 = 2.6466$

Da afstanden $AB = 14.236 - 12.24 = 1.99 < 2.647$ overlapper konfidensintervallerne.

G stat1.clowe...	H stat1.cupp...	I stat1...
10.9165	13.5635	12.24
12.9125	15.5595	14.236
9.77254	12.4195	11.096
8.66054	11.3075	9.984

Diameter i LSD-interval er $\frac{2.6466}{\sqrt{2}} = 1.8714$ Da $1.99 > 1.87$ overlapper LSD-intervallerne

ikke, dvs. producent B.s batterier har en signifikant større levetid end de øvrige producenters.
Konfidensinterval [12.91 ; 15.56]

6.3 FULDSTÆNDIGT RANDOMISERET BLOKFORSØG.

I forbindelse med planlægningen af et forsøg, kan man blive tvunget til at benytte forsøgsenheder, som er ret uensartede.

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.

Eksempler på blokforsøg kan være

- 1) Af tidsmæssige grunde (forsøget skal udføres indenfor en given tidsramme) må man benytte 4 forskellige apparater. Påvirker apparaterne forsøgsresultaterne, kan det give en stor spredning (stor "støj"). Dette kan betyde, at man skal op på et urealistisk stort antal gentagelser for at kunne opnå den ønskede information.
I stedet indfører man de 4 apparater som blokke.
- 2) Man er nødt til at benytte 5 forskellige råvareleverancer. Dette kan måske også give for stor spredning. Man indfører så de 5 råvareleverancer som 5 blokke.
- 3) I et landbrugsforsøg samles tilstødende arealer i blokke. Begrundelsen herfor er naturligvis, at arealer der ligger tæt ved hinanden er mere ensartede end arealer der ligger længere væk.

Eksempel 6.4 (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 betyde, 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.

Lad der findes tre fodertyper A, B og C med forskelligt vitaminindhold.

Fra hvert af 4 forskellige kuld grise udtages nu 3 grise.

Et kuld vælges, og ved lodtrækning bestemmes hvilke af de 3 grise, der bliver fodret med fodertype A, hvilken med fodertype B og den sidste får naturligvis type C.

Et nyt kuld udtages, og man randomiserer igen foderet indenfor kullet (blokken), osv.

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

- a) Test, om der er nogen væsentlig virkning af ændringen i foderets vitaminindhold.
- b) Hvis der er en forskel, så skal man angive hvilken foderblanding, der giver den største vægtforøgelse.

6. En faktor på mere end 2 niveauer

Løsning:

- a) H_0 : Foderblanding har ingen virkning på vægtforøgelsen
 H: Foderblanding har virkning på vægtforøgelsen

Da vi jo har 2 faktorer, "kuld" og "fodertype", er analysen en speciel "tosidet variansanalyse".

TI-Nspire: Lister og regneark ► Lav 3 lister (husk overskrift) ► Statistik ► statistiske test ► Tosidet variansanalyse ► Blok, 3 ► indsæt navne i menu ► Enter

TI89: APPS, STAT/LIST hvorefter data indtastes med første søjle (A) i list1, 2 søjle (B) i list 2 osv.
 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:

Man får for faktor en P-værdi på 0.0402

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.

Det ses iøvrigt, at P-værdi for blok er lav 0.028 , dvs. det var en god ide at indføre blokke.

A a	B b	C c	D	E
				=ANOVA2way(a,'
7	14	8.5	Titel	Tosidet variansa...
16	15.5	16.5	F	5.75628
10.5	15	9.5	PVal	0.040217
13.5	21	13.5	df	2.
			SS	54.125
			MS	27.0625
			FBlock	6.22009
			PValBloc...	0.028476
			dfBlock	3.
			SSBlock	87.7292
			MSBlock	29.2431
			dfError	6.
			SSError	28.2083
			MSError	4.70139
			s	2.16827

- b) TI - programmerne beregner ikke konfidensintervaller når der er et blokforsøg+1 faktor.

I stedet bruges formelen: Radius i 95% konfidensinterval er $t_{0.975}(f_{error}) \sqrt{\frac{s_{error}^2}{n}}$

Under error findes $MS = s_{error}^2$ og $df = f_{error}$. n er antal gentagelser af faktoren.

Her er $MS = 4.701$ $df = 6$ og $n = 4$

Radius i konfidensinterval er $t_{0.975}(6) \sqrt{\frac{4.701}{4}} = 2.653$

Radius i LSD-interval er $\frac{2.653}{\sqrt{2}} = 1.876$

Gennemsnit af de 3 foderblandinger $\text{mean}(\mathbf{a}) \blacktriangleright 11.75$ $\text{mean}(\mathbf{b}) \blacktriangleright 16.375$ $\text{mean}(\mathbf{c}) \blacktriangleright 12.$

A: 11.75 B: 16.375 C 12.0

TI-Nspire(lommeregner) : menu, data, listematematik, mean

TI89:brug F3.Math, mean

Da $16.375 - 12 = 4.375 > 2 \cdot 1.875$ giver **foderblanding B den største vægtforøgelse.**

Da det er et blokforsøg , kan vægtforøgelsen afhænge af hvilket kuld der er det bedste, så konfidensintervaller kan kun bruges til relative sammenligninger



Metode uden brug af færdigt "blokprogram"

Først foretages en ensidet variansanalyse med de tre fodertyper

Man noterer sig under faktor: $SS = 54.125$, $df = 2$ og $MS = s^2 = \frac{SS}{df} = 27.0625$

samt under error : $SS_{\text{error}} = 115.9375$ med $df = 9$

Derefter en ensidet variansanalyse med de 4 kuld.

Man noterer sig $SS_{\text{blok}} = 87.7291$ med $df = 3$.

Det blokkene har forminsket støjen (error) findes ved $SS = SS_{\text{error}} - SS_{\text{blok}} = 115.9375 - 87.7291 = 28.2084$ med $df = 9 - 3 = 6$

Resultaterne indtastes i en såkaldt variansanalysetabel

Variation	SAK=SS	df	$MS = s^2 = \frac{SS}{df}$	F	P-værdi
Factor: foder	54.125	2	27.0625	$\frac{27.06}{4.70} = 5.756$	$FCdf(5.756, \infty, 2, 6) = 0.0402$
Block: Kuld	87.7291	3	29.2431	6.22	0.0284
Error	28.2084	6	4.70139		

H_0 : Foder = 0 (Foder har ingen virkning) forkastes, da $P\text{-værdi} = 0.0402 < 0.05$

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

Da $P\text{-værdi}$ for kuld er mindre end 5% har det haft betydning, at man har opdelt i blokke

b) Radius i 95% konfidensinterval findes som under løsning for TI89 og TI-Nspire

6.4. Binomialfordelte variable

Ved analysen anvendes formlerne i afsnit 6.6.2.

Eksempel 6.5 (binomialfordelt variabel).

For hver af 6 leverancer af billige legetøjsbiler udtages en tilfældig prøve på 100 biler, og antallet af defekte biler tæltes. 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 opfyldt (se evt. oversigten i afsnit 6.6.2)

$$\chi^2 = \frac{1}{0.08 \cdot (1 - 0.08)} 100((0.06 - 0.08)^2 + (0.14 - 0.08)^2 + \dots + (0.07 - 0.08)^2) = \frac{1.08}{0.08 \cdot 0.92} = 14.67$$

χ^2 er χ^2 -fordelt med frihedsgradstallet $f = n - 1 = 7$

Da $P\text{-værdi} = P(\chi^2 > 14.67) = \text{chi2Cdf}(14.67, \infty, 7) = 0.0405 < 0.05$ forkastes nulhypotesen (svagt), dvs. vi har et (svagt) statistisk bevis for, at procenten af defekte biler i leverancen ikke er den samme. ◆

6.5. Poissonfordelt variable

Ved analysen anvendes formlerne i oversigt 6.6.3.

Eksempel 6.6 (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 6 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 kanyl 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 oversigten i afsnit 6.5.3 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 i afsnit 6.6.3 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.



6.6. OVERSICHT OVER CENTRALE FORMLER I KAPITEL 6

6.6.1 Oversigt over fremgangsmåde ved ensidet variansanalyse

Givet følgende skema:

	A ₁	A ₂		A _q
Observationer	x ₁₁	x ₁₂		x _{1q}
	x ₂₁	x ₂₂		x _{2q}

	x _{n1}	x _{n2}		x _{nq}

Idet de q variables middelværdier kaldes $\mu_1, \mu_2, \dots, \mu_q$ ønsker vi at teste nulhypotesen

$H_0: \mu_1 = \mu_2 = \dots = \mu_q$, mod H : "mindst én middelværdierne er forskellig fra en af de øvrige".

TI-Nspire: Lister og regneark ► Lav q lister (husk overskrift) ► Statistik ► statistiske test ► Ensidedet variansanalyse ► Data ► q ► indsæt navne i menu ► ok

TI89: APPS ► StatList ► Indtast data i 4 lister ► F6, C:ANOVA ► Data Input : Data ► Antal grupper = 4, ENTER, Udfyld listnavne (VAR-Link osv.) ENTER, ENTER.

P-værdi aflæses. Hvis P-værdi < α forkastes H_0 .

Konfidensintervaller

I **TI-Nspire** findes resultaterne som vandrette "lister" nederst i udskriften

I **TI89** findes de som ekstra søjler efter de normale lister

xbar	lowlist	uplist
\bar{x}_1	a	b
\bar{x}_2		
osv.		

Diameter i 95% konfidensinterval $d_{kon} = b - a$ Diameter i 95% LSD-interval $d_{LSD} = \frac{d_{kon}}{\sqrt{2}}$

Er $|\bar{x}_1 - \bar{x}_2| > d_{LSD}$ antages niveauerne A_1 og A_2 at være signifikant forskellige (forudsat antallet af niveauer er lille (< ca 6)).

Blokforsøg

		Faktor			
		A ₁	A ₂	...	A _q
Blokke	b ₁	x ₁₁	x ₁₂		x _{1q}
	b ₂	x ₂₁	x ₂₂		x _{2q}
	...	-	-		-
	b _n	x _{n1}	x _{n2}		x _{nq}

TI-Nspire: Lister og regneark ► Lav q lister (husk overskrift) ► Statistik ► statistiske test ► Tosidet variansanalyse ► Blok, q ► indsæt navne i menu ► Enter

TI89: APPS, STAT/LIST hvorefter data indtastes med første søjle (A) i list1, 2 søjle (B) i list 2 osv.

F6, ANOVA2-Way, ENTER

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

Næste skema udfyldes med List1, List 2 osv, ENTER

P-værdi aflæses. Hvis P-værdi < α forkastes $H_0: \mu_1 = \mu_2 = \dots = \mu_n$

6. En faktor på mere end 2 niveauer

Konfidensintervaller

Radius i 95% konfidensinterval er $r_{kon} = t_{0.975}(f_{error}) \sqrt{\frac{s_{error}^2}{n}}$

Under error findes $MS = s_{error}^2$ og $df = f_{error} \cdot n$ er antal gentagelser af faktoren.

Diameter i 95% konfidensinterval $d_{kon} = 2 \cdot r_{kon}$ Diameter i 95% LSD-interval $d_{LSD} = \frac{d_{kon}}{\sqrt{2}}$

Er $|\bar{x}_1 - \bar{x}_2| > d_{LSD}$ antages faktorerne R_1 og R_2 at være signifikant forskellige.

6.6.2 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	P-værdi
Aproksimativ metode $n_1 \cdot \hat{p} \in [5; n_1 - 5]$ $n_2 \cdot \hat{p} \in [5; n_2 - 5]$ $\dots \quad n_k \cdot \hat{p} \in [5; n_k - 5]$ hvor $\hat{p} = \frac{x_1 + x_2 + \dots + x_k}{n_1 + n_2 + \dots + n_k}$	$H_0: p_1 = p_2 = \dots = p_k$	$\chi^2 = \frac{1}{\hat{p}(1 - \hat{p})} \sum_{i=1}^k n_i (\hat{p}_i - \hat{p})^2$ hvor $\hat{p}_1 = \frac{x_1}{n_1}, \hat{p}_2 = \frac{x_2}{n_2}, \dots, \hat{p}_k = \frac{x_k}{n_k}$	$P(Y > \chi^2)$

6.6.3. 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
Aproksimativ 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 $\bar{x} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2 + \dots + n_k \cdot \bar{x}_k}{n_1 + n_2 + \dots + n_k}$	$P - \text{ værdi} < \alpha$ hvor $P - \text{ værdi} = P(Y > \chi^2)$

OPGAVER

Opgave 6.1

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 6.2

I følgende tabel er angivet resultaterne af gentagne bestemmelser af blodets alkoholkoncentration (i promille) hos 6 forskellige personer efter indtagelsen af 6 cl. alkohol.

Person					
1	2	3	6	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 alkoholmængde.

Opgave 6.3

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 spredning:

Spole nr.	Antal gentagelser	Gennemsnit	Spredning
1	4	15.2	0.1871
2	4	14.95	0.1225
3	4	14.8	0.1472
4	4	15,15	0.1826
5	4	14.7	0.2562

- 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.

6. En faktor på mere end 2 niveauer

Opgave 6.4.

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 6.5.

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.

Opgave 6.6

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 studentereksamensgenemsnit.

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 6.7.

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 velbegrundet.?

Opgave 6.8

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.

7 2 FAKTORER PÅ 2 ELLER FLERE NIVEAUER TOSIDET VARIANSANALYSE

7.1 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.

7.2 PLANLÆGNING AF FORSØG.

I dette afsnit benyttes følgende eksempel som illustration af begreberne.

Eksempel 7.1. 2 faktorer .

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.



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

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

I alt er der altså 6 **behandlinger**.

7.2.1 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 , hvor man har valgt, at hvert niveau skal gentages mindst 4 gange.

Endvidere er det væsentligt, at hver af de indgående behandlinger har lige mange gentagelser.

		Karburator	
		K_1	K_2
Olieblending	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 olieblending O_1 , 4 andre biler olieblending O_2 , og de sidste 4 biler olieblending O_3 .

Efter at have kørt en udvalgt strækning måles benzinføbruget. Derefter varieres den anden faktor (her karburator), mens olieblendingen fastholdes på O_2 , dvs. de sidste 4 biler udstyres med karburator K_2 og olieblending O_2 . Igen gennemkøres den udvalgte strækning, og benzinføbruget måles. Indtegnes for hver karburator det gennemsnitlige benzinføbrug mod olie-blendingen kunne vi eksempelvis få tegningen på fig. 7.1.

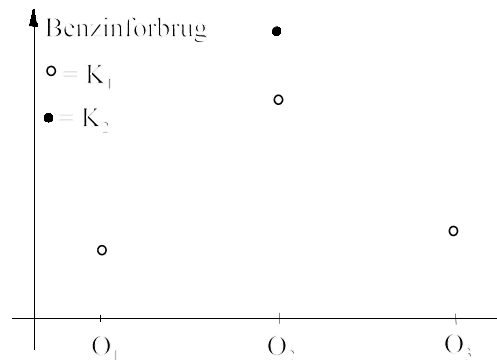


Fig 7.1 Skitse af benzinføbrug

Umiddelbart ses, at K_1 giver lavest benzinføbrug, og O_1 (eller O_3) skal foretrækkes. Hvad nu med benzinføbruget i karburator K_2 , hvis vi anvender olieblending O_1 eller O_3 ? Kan man slutte, at benzinføbruget 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.

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

7.2.2 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
Olieblending	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.

Det er vigtigt, at hver behandling har lige mange gentagelser

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 benzinforgbrug mod olie-blandingen, kan det eksempelvis vise sig, at man får figur 7.2.

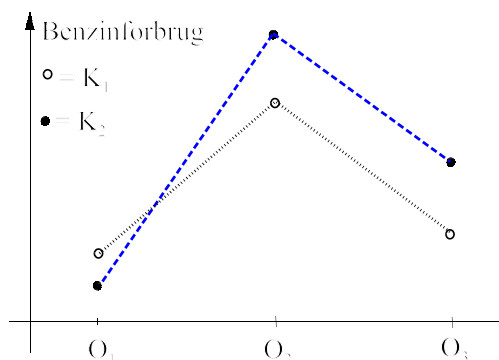


Fig 7.2. Vekselvirkning

Vi ser, i modstrid med hvad vi antog ud fra "en faktor ad gangen forsøget", at kombinationen af katalysator K₂ og olieblending O₁ giver det laveste benzinforgbrug.

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**.

7.3. 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 7.2. 2 faktorer (fortsat fra eksempel 7.1) .

En bilfabrikant ønsker at finde ud af, hvorledes 3 olieblandinger O₁, O₂, og O₃, og 2 karburator typer K₁ og K₂ påvirker benzinforgbruget.

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 olieblending 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 følgende $N = r \cdot q \cdot n$.

I eksempel 7.4 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.

Skitse af fremgangsmåde ved testning

Først testes om modellen er additiv, dvs. om den er uden en signifikant vekselvirkning.

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

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

For at kunne beregne en P-værdi beregnes først et estimat for den fælles (poolede) varians s_0^2 på samme måde som det skete i forrige kapitel.

Derefter beregnes et estimat for en varians s_{RC}^2 som vil være lig s_0^2 hvis H_0 er sand, men væsentlig større hvis H er sand

P-værdien beregnes som en F-test ud fra brøken $F = \frac{s_{RC}^2}{s_0^2}$.

1) Er P-værdi $< \alpha$ (signifikansniveau) forkastes H_0 dvs. faktorerne vekselvirker.

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

2) Er P-værdi $> \alpha$ accepteres H_0 dvs. faktorerne vekselvirker ikke (model er additiv).

Man tester nu nulhypoteserne

$H_0: R = 0$ (rækkefaktor har ikke en virkning)

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

Da man må antage, at de to varianser s_{RC}^2 og s_0^2 er nogenlunde ens, pooles de sammen til et nyt (bedre) estimat s_m^2 for forsøgsfejlen (error).

Variansen s_m^2 benyttes nu ved beregning af P-værdierne for de to hypoteser.

Hvis frihedsgraden for s_0^2 er "stor" og P-værdien for en faktor beregnet ud fra s_0^2 er langt fra signifikansniveauet α , vil en pooling ikke ændre P-værdien så meget, at vi får en anden konklusion. Man behøver derfor i en sådan situation ikke at poole.

a) Finder man at såvel $H_0: R = 0$ forkastes, som $H_0: C = 0$ forkastes vil man sædvanligvis opstille konfidensintervaller til bestemmelse af de to faktors optimale niveau.

Hertil benyttes s_m^2 , da det jo giver et sikrere konfidensinterval.

b) Finder man at eksempelvis $H_0: R = 0$ forkastes, mens $H_0: C = 0$ accepteres pooles varianserne s_{RC}^2 , s_0^2 og s_c^2 sammen til en fælles varians, og denne benyttes ved opstillingen af konfidensintervaller for R.

Beregninger: Hvorledes man foretager testen ved hjælp af en lommeregner der kun kan beregne gennemsnit og spredning er beskrevet i appendix 7.1. Da specielt beregningerne af SAK'erne er temmelig omfattende, og næppe giver en dybere forståelse, vil vi dog sædvanligvis benytte programmer som findes i TI-modellerne hertil. Programmerne dækker dog ikke alle muligheder, så i enkelte tilfælde må man derfor benytte formlerne for eksempelvis konfidensintervaller.

7 To faktorer på 2 eller flere niveauer

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 s² sammen til et fælles skøn s₀ 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 (R₂ = R₁ + 7, R₃ = R₁ + 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 R₂ + C₃ - 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 (s₀), vil vi kunne konkludere at der ikke kan konstateres nogen vekselvirkning.

7.4 BEREGNING AF TOSIDET VARIANSANALYSE

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

Er modellen additiv er der igen 2 tilfælde, nemlig om begge faktorer har en signifikant virkning eller kun den ene faktor har en signifikant virkning.

Vi ser her bort fra den situation, at ingen faktorer har en signifikant virkning. Det kan naturligvis sagtens forekomme, men den situation kræver ingen nærmere forklaring.

I afsnit 7.7 er givet en oversigt over hvorledes beregningerne skal udføres med TI89.

I det følgende anvendes denne oversigt til at regne 3 eksempler svarende til hver af de tre situationer.

Eksempel 7.2 (fortsat) Model med vekselvirkning.

En bilfabrikant ønsker at finde ud af, hvorledes 3 olieblandinger O₁, O₂, og O₃, og 2 karburator typer K₁ og K₂ 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

Det forudsættes at betingelserne om uafhængighed, normalitet og varianshomogenitet er opfyldt.

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

Løsning:

Først testes $H_0: R * C = 0$ mod $H: R * C \neq 0$

TI-Inspire: Lister og regneark ► Lav 2 lister K1 og K2 ► Statistik ► statistiske test ► Tosidet variansanalyse(ANOVA) ► Ligningsrep med 2 faktorer ► 2 ► indsæt navne i menu ► antal rækkefaktorer r ► ok

TI89: APPS ► StatList ► Indtast data i 2 lister (k1 og K2) ► F6, ANOVA2-Way ► Enter ► DESIGN=2 ► Factor, EqReps, Levels of Col Factor =2, Levels of Row Factor =3, ENTER ► Næste skema udfyldes med List1 og List 2, ENTER

A k1	B k2	C	D	E
				=ANOVA2way('k1','k2,3): CopyV
830	810		Titel	Tosidet variansanalyse ANOV...
860	840		FCol	10.125
940	1050		PValCol	0.019028
990	1020		dfCol	1.
855	930		SSCol	6075.
815	910		MSCol	6075.
			FRow	48.9306
			PValRow	0.000193
			dfRow	2.
			SSRow	58716.7
			MSRow	29358.3
			FInteract	5.375
			PValInteract	0.045963
			dfInteract	2.
			SSInteract	6450.
			MSInteract	3225.
			dfError	6.
			SSError	3600.
			MSError	600.
			s	24.4949

Under "Pvalinteraction" findes P - værdi = 0.04596.

Med et 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.

For at finde den kombination af olieblending og karburator, der giver det mindste benzinformbrug beregnes gennemsnittene i hver celle.

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

Det ses, at det mindste gennemsnit er 825.

Beregning af 95% konfidensintervaller og LSD-intervaller (jævnfør afsnit 7.7)

Under "error" findes $s_{error}^2 = MS = 600$ og $df = 6$

Radius i 95% konfidensinterval er $r_{kon} = t_{0,975}(6) \sqrt{\frac{600}{2}} = 2.446 \cdot \sqrt{300} = 42.38$

Diameter i 95% konfidens- og LSD - intervaller er $d_{kon} = 2 \cdot r_{kon} = 84.76$ og $d_{lsd} = \frac{84.76}{\sqrt{2}} = 59.94$

7 To faktorer på 2 eller flere niveauer

Det ses, at umiddelbart giver $K_2 O_1$ det laveste benzinförbruk, men da $K_1 O_3$ og $K_1 O_1$ begge har et benzinförbruk der afviger mindre end 59.9 herfra, er der ingen signifikant forskel mellem $K_2 O_1$, $K_1 O_3$ og $K_1 O_1$.

Konklusion:

Man får det laveste benzinförbruk ved at vælge en af kombinationerne $K_2 O_1$, $K_1 O_3$ og $K_1 O_1$.

Et estimat for det laveste förbruk er 825 liter

Heraf ses, at kombinationen $K_2 O_1$ giver det laveste benzinförbruk (825), men, at der ingen signifikant forskel er mellem $K_2 O_1$, $K_1 O_3$ og $K_1 O_1$.



Eksempel 7.3 Additiv model: To signifikante hovedvirkninger

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

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

Försögsresultaterne var:

		Cementknusere											
		1			2			3			4		
Cementblandere	1	147	175	130	99	85	75	67	23	35	215	97	180
	2	221	155	173	141	110	155	85	55	81	161	167	177
	3	123	85	153	137	143	82	67	25	83	135	91	129

Forudsætningerne for en variansanalyse antages opfyldt.

a) Angiv hvilke kombinationer af cementblender og cementknuser, der giver den største brudstyrke

b) Giv et estimat og et 95% konfidensinterval for denne største middelbrudstyrke.

Lösning:

TI-Nspire og **TI89**: Indtastning foregår som i eksempel 7.2

Resultatet fremgår af omstående skema

A	B	C	D	E	F
ck1	ck2	ck3	ck4		
147	99	67	215	Titel	Tosidet ...
175	85	23	97	FCol	19.5557
130	75	35	180	PValCol	0.000001
221	141	85	161	dfCol	3.
155	110	55	167	SSCol	51995.2
173	155	81	177	MSCol	17331.7
123	137	67	135	FRow	4.91158
85	143	25	91	PValRow	0.016292
153	82	83	129	dfRow	2.
				SSRow	8706.06
				MSRow	4353.03
				FInteract	1.33942
				PValInterac...	0.278669
				dfInteract	6.
				SSInteract	7122.61
				MSInteract	1187.1
				dfError	24.
				SSError	21270.7
				MSError	886.278
				s	29.7704

a) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning) mod $H: R * C \neq 0$

For "interaction" findes P -værdi ($P_{\text{værdi}}(\text{interact}) = 0.2787$.

Da P -værdi $= 0.2787 > 0.05$ accepteres H_0 .

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

Vi skal nu teste

$H_0: C = 0$ (Cementknuser har ingen virkning) mod $H: C \neq 0$

Da P -værdi for C er 0.000001 vil en pooling af vekselvirkningen ikke ændre, at $H_0: C = 0$ forkastes, dvs. Cementknuserne har en (stærk) virkning

$H_0: R = 0$ (Cementblander har ingen virkning) mod $H: R \neq 0$

Da P -værdi for $R = 0.016$ kunne en pooling af vekselvirkningen muligvis øge P -værdi til mere end 0.05. Det er ikke muligt at give nogle regler herfor, så for en sikkerheds skyld foretages pooling.

For overskueligheden skyld opskrives beregningerne i nedenstående variansanalysetabel.

Mere detaljerede beregninger ses nedenfor

Variation	SAK=SS	df	$MS = s^2 = \frac{SS}{f}$	F	P -værdi
Column:Cementknuser : C	51995.2	3			
Row: Cementblander : R	8706.06	2	4253.03	4.5994	0.0181
Interaction:R*C	7122.61	6			
Error	21270.7	24	886.278		
error (pool)	$7122.61+21270.7=28393.3$	30	946.443		

Da vi lige har konkluderet, at der ingen vekselvirkning er, må et sikrere estimat for variansen være en pooling af de to varianser

$$SAK_{\text{pool}} = SAK_{\text{error}} + SAK_{\text{vekselvirkning}} = 7122.61 + 21270.7 = 28393.3.$$

$$f_{\text{pool}} = f_{\text{error}} + f_{\text{vekselvirkning}} = 24 + 6 = 30 \quad s_{\text{pool}}^2 = \frac{28393.3}{30} = 946.443$$

$$F_{\text{blander}} = \frac{4253.03}{946.44} = 4.5994. \quad P\text{-værdi} = P(F > 4.5994) = FCdf(4.5994, \infty, 2, 30) = 0.0181$$

Da P -værdi $= 0.0181 < 0.05$ forkastes H_0 (P -værdien steg dog ved pooling)

Konklusion: Cementknuserne har en (stærk) virkning.

Cementblandere har en virkning

For at finde hvilken cementblander der giver den største middelbrudstyrke beregnes gennemsnit for blanderne:

$$1: \text{mean}(\{147, 175, 130, 99, 85, 75, 67, 23, 35, 215, 97, 180\}) = 110.67$$

$$2: \text{mean}(\{221, 155, 173, 141, 110, 155, 85, 55, 81, 161, 167, 177\}) = 140.08$$

$$3: \text{mean}(\{123, 85, 153, 137, 143, 82, 67, 25, 83, 135, 91, 129\}) = 104.42$$

Det ses, at cementblander 2 giver den største brudstyrke

7 To faktorer på 2 eller flere niveauer

Konfidensinterval for de 4 knusere findes ved at benytte de i afsnit 7.7 angivne formler.
 Bemærk: Her er det vigtigt, at man for en sikkerheds skyld altid bruger det poolede estimat)

$$r_{kon} = t_{0.975}(f_{pool}) \sqrt{\frac{s_{pool}^2}{n \cdot q}} = t_{0.975}(30) \sqrt{\frac{946.443}{3 \cdot 4}} = 2.04 \cdot 8.881 = 18.14$$

$$d_{kon} = 2 \cdot 18.14 = 36.3 \quad d_{lsd} = \sqrt{2} \cdot 18.14 = 25.7$$

Da $140.08 - 110.67 = 29.41 > 25.7$ ses af LSD-intervallet,
at cementblander 2 må foretrækkes

Gennemsnit for knuserne findes (skriv i celle mean(ck1) osv.)

1: 151.3 2: 114.11 3: 57.89 4: 150.22

$$r_{kon} = t_{0.975}(f_{pool}) \sqrt{\frac{s_{pool}^2}{n \cdot r}} = t_{0.975}(30) \sqrt{\frac{946.443}{3 \cdot 3}} = 2.04 \cdot 10.255 = 20.92$$

$$d_{kon} = 2 \cdot 20.92 = 41.9 \quad d_{lsd} = \sqrt{2} \cdot 20.92 = 29.6$$

Benyttes LSD, ses, at Cementknuser 1 eller 4 må foretrækkes

Konklusion: Størst middelbrudstyrke fås i kombinationen
 cementknuser 1 og cementblander 2 eller cementknuser 4 og cementblander 2

- b) Et estimat $\tilde{\mu}_{12}$ for største middelbrudstyrke på basis af cementknuser 1 og cementblander 2: (se oversigt i kapitel 7.7)

$$\bar{x}_1 = 151.3 \quad \bar{x}_2 = 140.08 \quad \bar{x} = \frac{110.67 + 140.08 + 104.42}{3} = 118.4$$

$$\tilde{\mu}_{12} = \bar{x}_1 + \bar{x}_2 - \bar{x} = 151.3 + 140.1 - 118.4 = \underline{172.9}$$

$$\text{Radius i konfidensinterval: } r_{kon} = t_{0.975}(f_m) \sqrt{\frac{3+4-1}{36} s_m} = t_{0.975}(30) \frac{\sqrt{946.443}}{\sqrt{6}} = 25.6$$

$$95\% \text{ konfidensinterval: } [172.9 - 25.6; 172.9 + 25.6] = \underline{[146.6; 197.8]}$$



Eksempel 7.4 Additiv model: Een signifikant hovedvirkning

Samme problem som i eksempel 7.3, blot er tallene for cementblander 2 formindsket med 10.
 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

Forudsætningerne for en variansanalyse antages opfyldt.

Angiv hvilke kombinationer af cementblander og cementknuser, der giver den største brudstyrke, og giv et estimat og et 95% konfidensinterval for denne største middelbrudstyrke.

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

Beregningerne er ganske analoge til eksempel 7.3.

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. $H_0: C = 0$ (Cementknuser har ingen virkning) mod $H: C \neq 0$ Da P -værdi for C er 0.000001 vil en pooling af vekselvirkningen ikke ændre, at $H_0: C = 0$ forkastes, dvs. Cementknuserne har en (stærk) virkning $H_0: R = 0$ (Cementblander har ingen virkning) mod $H: R \neq 0$ Da P -værdi for $R = 0.1097$ kunne en pooling af vekselvirkningen ændre P -værdi til en lidt anden værdi, men da frihedsgradet for error på 24 er så stort, at en forøgelse med 6 til 30 næppe vil ændre resultatet meget, er en pooling unødvendig $H_0: R = 0$ (Cementblander har ingen virkning) accepteres, da P -værdi = 0.1097 > 0.05**Konklusion:** Cementknuserne har en (stærk) virkningCementblandere har ingen virkning

For at finde hvilken cementknuser der giver den største middelbrudstyrke kunne vi nu beregnes gennemsnit og konfidensinterval for de 4 knuser ved at benytte de i oversigt 7.7 angivne formler.

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

Konfidensintervaller findes som ekstra søjler (se evt. eksempel 4.2)

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

Cementknuser 1 og 4 må foretrækkes, selv om man ikke helt kan afvise at cementknuser 2 kan være lige så god. For at afgøre dette beregnes diameter i LSD-interval

$$d_{\text{kon}} = 169.7 - 126.3 = 43.4 \quad d_{\text{LSD}} = \frac{43.4}{\sqrt{2}} = 30.7$$

Konklusion: Cementknuser 1 og 4 må foretrækkes

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



7.5 FULDSTÆNDIGT RANDOMISERET BLOKFORSØG.

Som beskrevet i kapitel 4.3, kan man blive tvunget til at benytte forsøgsenheder, som er ret uensartede. Derved får den tilfældige forsøgsfejl en relativ stor spredning (stor "støj").

For at "dæmpe støjen" kan man inddele forsøgsenhederne i grupper (blokke), hvor de forsøgsenheder der ligger i samme blok er væsentlig mere ensartede end forsøgsenhederne i forskellige blokke.

Til illustration heraf, så betragter vi igen forsøget beskrevet i eksempel 7.1.

Eksempel 7.5 (randomiseret blokforsøg). 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 benzinforsøget. 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) Beskriv hvorledes en randomisering kunne tænkes at foregå.

2) Opdelingen i 2 blokke resulterede i følgende plan med tilhørende benzinforsøg

bil 1	dag 1 O_1K_2 810	dag 2 O_3K_2 910	dag 3 O_2K_2 1020	dag 4 O_1K_1 830	dag 5 O_2K_1 940	dag 6 O_3K_1 815
bil 2	dag 1 O_1K_1 860	dag 2 O_3K_2 930	dag 3 O_1K_2 840	dag 4 O_2K_1 990	dag 5 O_3K_1 855	dag 6 O_2K_2 1050

Skitsér udseendet af en variansanalysetabel og beregn om man på dette grundlag kan vise, at der er vekselvirkning mellem olieblending og karburator.

Løsning:

1) Randomisering: To dåser mærkes henholdsvis bil1 og bil2. 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

Derefter fortsættes med at trække sedler fra dåsen med bil 2.

2) Analyse: Der udføres en tosidet variansanalyse med faktorerne "olieblending" og "karburator"

Da tallene er de samme som i eksempel 7.2 fås samme resultater.

Der udføres en ensidet variansanalyse med blokke (biler) som faktor.

Resultat: $SS = 3333.33$ $df = 1$ $MS = 3333.33$

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 benzinforsøg end bil 2).

Der udføres en tosidet variansanalyse.

Resultat: $S_{\text{error}} = 3600$, $df = 6$

Idet blokvirkningen trækkes fra fås $SS = 3600 - 3333.33 = 266.67$ med $df = 6 - 1 = 5$.

Resultaterne indskrives i en variansanalysetabel

Variansanalyse	SAK	f	MS	F	P-værdi
Blokke (biler)	3333.33	1	3333.33		
Olieblending		2			
Karburator		1			
Olie * karburator	6450	2	3225	60.469	$FCdf(60.469, \infty, 2, 5) = 0.0031$
"Error"	266.67	5	53.333		
Total		11			

Idet vi som sædvanlig antager at signifikansniveauet er 5 % fås, at da P -værdi = 0.31% < 5% forkastes H_0

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

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. ◆

7.6. To binomialfordelte eller Poissonfordelte 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. I appendix 7.2 er vist, hvordan man kan transformere data så det er tilladeligt at bruge variansanalyseteknikken på de transformerede data

7.7. OVERSIGT OVER FREMGANGSMÅDE VED TOSIDET VARIANSANALYSE

Givet følgende skema

		Søjlefaktor Q			
		Q ₁	Q ₂	...	Q _q
Rækkefaktor R	R ₁	X ₁₁₁ X ₁₁₂ ... X _{11n}	X ₁₂₁ X ₁₂₂ ... X _{12n}		X _{1q1} X _{1q2} ... X _{1qn}
	R ₂	X ₂₁₁ X ₂₁₂ ... X _{21n}	X ₂₂₁ X ₂₂₂ ... X _{22n}		X _{2q1} X _{2q2} ... X _{2qn}
	...				
	R _r	X _{r11} X _{r12} ... X _{r1n}	X _{r21} X _{r22} ... X _{r2n}		X _{rq1} X _{rq2} ... X _{rqn}

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

H: $R * C \neq 0$

TI-Inspire: Lister og regneark ► Lav q lister Q₁, Q₂ osv. ► Statistik ► statistiske test ► Tosidet variansanalyse (ANOVA) ► Ligningsrep med 2 faktorer ► q ► indsæt navne i menu ► antal rækkefaktorer r ► ok

TI89: Apps ► Stats ► Data indtastes: Q₁ søjle i list1 Q₂ søjle i list2 osv. F6, ANOVA2-Way, ENTER DESIGN=2 Factor, EqReps, Levls of Col Factor =q, Levls of Row Factor =r, ENTER

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

Resultat fremkommer som en lang række tal

Under "interaction" findes "P-value"

1) P -værdi $< \alpha$: H_0 forkastes, dvs. Faktorerne R og C vekselvirker

Konfidens og LSD intervaller opstilles til afgørelse af hvilken kombination af faktorer, der eksempelvis giver mindst resultat.

Man udregner de $r \cdot q$ gennemsnit og finder den mindste \bar{x}_{\min}

Radius $r_{\text{kon}} = t_{0,975}(f_{\text{error}}) \cdot \sqrt{\frac{s_{\text{error}}^2}{n}}$, hvor man under "ERROR" finder s_{error}^2 og $f_{\text{error}} = \text{df}$

Beregner diameter $d_{\text{kon}} = 2 \cdot r_{\text{kon}}$ og $d_{\text{lsd}} = \sqrt{2} \cdot r_{\text{kon}}$

De gennemsnit der ligger indenfor en afstand d_{lsd} fra \bar{x}_{\min} er ikke signifikant forskellige.

2) P -værdi $> \alpha$: H_0 accepteres, dvs, der kan ikke påvises en signifikant vekselvirkning

For overskuelighedens skyld opstilles en variansanalysetabel ud fra udskrifterne

Variation	SAK=SS	f = df	MS=s ²	F
Rækkefaktor R :	SAK _R	$f_R = r - 1$		
Søjlefaktor C :	SAK _C	$f_C = q - 1$		
Vekselvirkning R*C	SAK _{RC}	$f_{RC} = (r-1)(q-1)$		
Gentagelser (residual, error)	SAK _{error}	$f_{\text{error}} = r \cdot q \cdot (n-1)$		

7.7. Oversigt over fremgangsmåde ved tosidet variansanalyse

Der foretages en "pooling" af vekselvirkning ned i Error, og man udregner nye P-værdier.

Variation	SAK=SS	f = df	MS=s ²	F	P-værdi
Rækkefaktor R	SAK _R	$f_R = r - 1$	$s_R^2 = \frac{SAK_R}{f_R}$	$F_R = \frac{s_R^2}{s_{pool}^2}$	FCdf(F _R , ∞, f _R , f _{pool})
Søjlefaktor C:	SAK _C	$f_C = q - 1$	$s_C^2 = \frac{SAK_C}{f_C}$	$F_C = \frac{s_C^2}{s_{pool}^2}$	FCdf(F _C , ∞, f _C , f _{pool})
Error(poolet)	SAK _{pool} = SAK _{RC} + SAK _{error}	f _{pool} = f _{RC} + f _{error}	$s_{pool}^2 = \frac{SAK_{pool}}{f_{pool}}$		

2a) Begge P-værdier < α

H₀: R = 0 (Ingen signifikant virkning af rækkefaktor)

H₀ forkastes, dvs. rækkefaktor har en signifikant virkning

Konfidens- og LSD intervaller opstilles til afgørelse af, hvilket niveau af rækkefaktor der eksempelvis giver mindst resultat.

Man udregner de r rækkegennemsnit og finder den mindste $\bar{x}_{r, \min}$

$$\text{Radius } r_{\text{kon}} = t_{0.975}(f_{\text{pool}}) \cdot \sqrt{\frac{s_{\text{pool}}^2}{n \cdot q}}$$

Beregner diameter d_{kon} = 2 · r_{kon} og d_{lsd} = √2 · r_{kon}

De gennemsnit der ligger indenfor en afstand d_{lsd} fra $\bar{x}_{r, \min}$ er ikke signifikant forskellige

H₀: C = 0 (Ingen signifikant virkning af søjlefaktor)

H₀ forkastes, dvs. søjlefaktor har en signifikant virkning

Man udregner de q søjlegennemsnit og finder den mindste $\bar{x}_{q, \min}$

$$\text{Radius } r_{\text{kon}} = t_{0.975}(f_{\text{pool}}) \cdot \sqrt{\frac{s_{\text{pool}}^2}{n \cdot r}}$$

Beregner diameter d_{kon} = 2 · r_{kon} og d_{lsd} = √2 · r_{kon}

De gennemsnit der ligger indenfor en afstand d_{lsd} fra $\bar{x}_{q, \min}$ er ikke signifikant forskellige

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

$\tilde{\mu}_{ij} = \bar{x}_i + \bar{x}_j - \bar{x}..$ (jævnfør betragtningerne i afsnit 7.3)

Konfidensintervaller for hver celle:

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

7 To faktorer på 2 eller flere niveauer

2b) Kun en P-værdi $< \alpha$.

Lad os antage, det er den ud for rækkefaktoren

$H_0: R = 0$ (Ingen signifikant virkning af rækkefaktor)

H_0 forkastes, dvs. rækkefaktor har en signifikant virkning

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

H_0 acceptere, dvs. søjlefaktor har ingen signifikant virkning

Problemet svarer nu til, at man har en ensidet variansanalyse, og beregningerne fortsætter derfor som sådan

OPGAVER

Opgave 7.1

I et forsøg undersøges, 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) Idet det antages at forudsætningerne for at udføre en variansanalyse er tilstede, skal der udføres en test til vurdering af, om middelduudbyttets (eventuelle) afhængighed af de benyttede katalysatorer og apparater kan beskrives ved en additiv model.
- 2a) Hvis man af økonomiske grund vælger apparat A_1 hvilken katalysator skal man så vælge?
- 2b) Hvis man af økonomiske grunde vælger katalysator 1 hvilket apparat skal man så vælge?
- 2c) Hvilken (hvilke) kombinationer af apparat og katalysator giver det største udbytte.

Opgave 7.2

Man ønsker at undersøge den virkning som 2 faktorer (typen af glas og fosfor) har på skarpheden 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	255	220
		235	240	225
		240	235	230

Idet de sædvanlige variansanalyseforudsætninger antages opfyldt, ønskes følgende spørgsmål belyst:

- 1) Har glastype og fosfortype indflydelse på skarpheden?
- 2) Ud fra svaret i spørgsmål 2 skal angives, hvilken glastype og fosfortype der giver den største skarphed (giver den mindste respons)

Opgave 7.3

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

Idet de sædvanlige variansanalyseforudsætninger antages opfyldt, ønskes følgende spørgsmål belyst:

- 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 7.4

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 laboratiemæssig 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 7.5

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.

Opgave 7.6.

Ved en tekstilfabrikation målttes 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 middelantallet af garnbrud afhænger af vævemetoder og/eller materialetyper.

Opgave 7.7.

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 målttes.

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 middelloverlevelsestid for hver enkelt vitaminbehandling og bestem den vitaminbehandling, som må antages at have bedst virkning.

8 FLERE END TO FAKTORER, SCREENINGS-FORSØG

8.1 Beskrivelse af forsøg

Vi har i variansanalysen analyseret forsøg med 1 eller 2 faktorer. Imidlertid vil man i praksis ofte have behov for forsøg, hvori der indgår mange faktorer.

De følgende eksempler illustrerer dette.

Eksempel 8.1. (5 faktorer)

Mængden af et uønsket spildprodukt ved en proces kan muligvis afhænge af en eller flere af følgende 5 faktorer:

A: Mængden af reaktant

B: Syrekoncentrationen

C: Katalysatoren

D: Reaktionsid

E: Reaktionstemperatur.



Eksempel 8.2. (3 faktorer)

Antallet af brud i en stålfjeder er rigelig stor. Følgende 3 faktorer har muligvis betydning:

A: Stålets procentiske indhold af kul.

B: Temperaturen af det oliebad, som fjederen dyppes ned i under hærdeningen

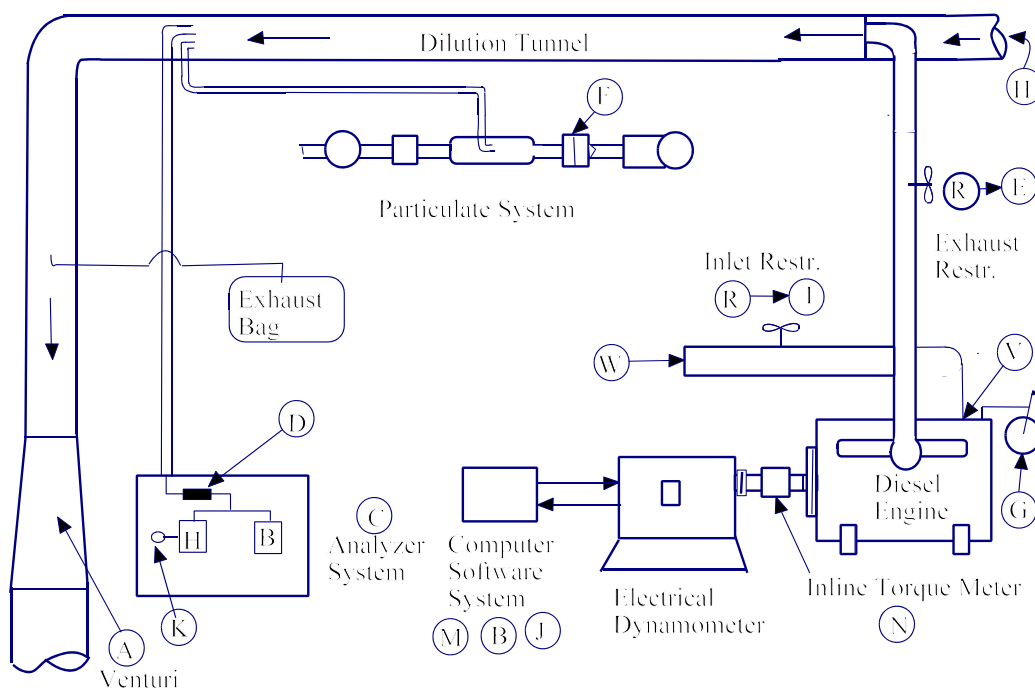
C: Fjederens temperatur lige før den nedsænkes i oliebadet.



Eksempel 8.3 (over 20 faktorer)

Skitsen viser et måleudstyr til måling af emissionen af NO_x fra dieselmotorer

Da der var for stor variation i målingerne søgte man at forbedre udstyret. Efter "brainstorm" blev foreslået en lang række faktorer, som på figuren er angivet med bogstaverne A, B, C,



8.2 Statistikprogrammer, der kan foretage beregningerne.

Det er imidlertid beregningsmæssigt så kompliceret, at TI-Nspire ikke har et program hertil. På samme hjemmeside som denne bog findes der 2 bøger, som har kapitler hvor beregningerne foretages. En bog "Vidergående statistik" som anvender statistikprogrammet SAS.JMP og en anden bog med titel "Statistisk forsøgsplanlægning som anvender statistikprogrammet Statgraphics.

Her er det, at de såkaldte screeningsforsøg er centrale.

Man kan ved denne metode med forholdsvis få delforsøg undersøge mange faktorerers virkning på en produktions slutresultat.

Ved forsøget får man så afklaret, hvilke af faktorerne der har en væsentlig virkning. De sædvanligvis få faktorer, som viser sig at have en virkning, kan man så studere nøjere ved supplerende forsøg.

Den statistiske metode blev i 30-erne udviklet af en berømt statistiker R. Fisher, (F- fordelingen er opkaldt efter ham) .

9. ENKELT REGRESSIONSANALYSE

9.1 INDLEDNING

I dette kapitel betragtes forsøg, hvor man har målt sammenhørende værdier af to **kvantitative** variable. Det følgende eksempel demonstrerer et sådant tilfælde.

Eksempel 9.1. Kvantitative variable

I et spinderi udtrykkes garnets kvalitet bl.a. ved en norm for den forventede trækstyrke. Kvaliteten anses således for at være i orden, hvis middeltrækstyrken mindst er lig med 10 måleenheder (me).

Ved uldgarn opfylder garnets naturlige trækstyrke ikke det nævnte kvalitetskrav, hvorfor der tilsættes en vis mængde kunstfibre, hvilket forøger trækstyrken. Herved sker der dog det, at andre kvalitetsegenskaber, såsom elasticitet og isoleringsevne, forringes. Man har eksperimenteret med forskellige tilsatte mængder kunstfibre x og registreret garnets trækstyrke y ved disse forskellige mængder. Herved fremkom følgende observationsmateriale:

Mængde x (i gram) af kunstfibre pr. kg uld	40	50	55	60	70	75	80	85	90	95	100	105	110	120	130
Trækstyrke (me): Y	4.5	6.5	5.4	7.0	8.2	8.0	7.1	8.9	8.2	10.3	9.6	10.8	10.5	11.2	12.0

Mængden af kunstfibre x er blevet bestemt på forhånd (har fået ganske bestemte værdier), så den er ikke en statistisk variabel. Trækstyrken Y synes derimod udover mængden af kunstfibre også at være påvirket af andre ukendte og ukontrollable “støjfaktorer”. Y må derfor opfattes som en statistisk variabel.

I andre situationer er både X og Y statistiske variable. Dette gælder eksempelvis, hvis man ønsker at undersøge om der er en sammenhæng mellem personers højde Y og vægt X , og derfor for en række personer måler sammenhørende værdier af højde og vægt.

Målet med en regressionsanalyse er at finde en funktionssammenhæng mellem den uafhængige variabel y og de afhængige variable.

I eksempel 9.1 ville man umiddelbart sige, at da man har 15 punktpar, så vil et polynomium af fjortende grad $y = a_{14}x^{14} + a_{13}x^{13} + \dots + a_1x + a_0$ gå igennem alle punkter, og det må derfor være en god model. Dette er imidlertid ikke tilfældet, da y -værdierne jo er resultater af forsøg der er påvirket af ukontrollable støjfaktorer. Polynomiets koefficienter vil derfor afspejle disse tilfældige udsving, og det giver derfor en ganske meningsløs model. Endvidere er modellen alt for matematisk kompliceret til at kunne bruges i praksis. Vi søger derfor i regressionsanalysen en enklere model, som tager rimeligt hensyn til støjen ved målingerne.

Er den ene variabel som i eksempel 9.1 en (kontrolleret) ikke statistisk variabel, så har man mulighed for hver x -værdi, at foretage gentagne målinger af den statistiske variabel Y (randomiseret). Dette giver mulighed for at beregne et estimat for den spredning der skyldes støjen, hvilket (som det vises i afsnit 9.2.3) kan udnyttes ved testning af den foreslåede model.

Lineær model.

Ved en lineær model forstås her en model, der er lineær med hensyn til parametrene.

Eksempelvis er såvel funktionen $y = f(x) = a + bx + cx^2$ som $y = g(P, T) = a + b \cdot P + c \cdot T$ lineære i de 3 parametre a , b og c . Som et eksempel på en model der ikke er lineær i parametrene kan nævnes $y = a + bx^c$.

Vi vil i dette kapitel betragte det ved anvendelserne meget ofte forekomne tilfælde, som kaldes **enkelt regressionsanalyse**, og hvor modellen er lineær i 2 parametre, eller kan gøres lineær ved en passende transformation..

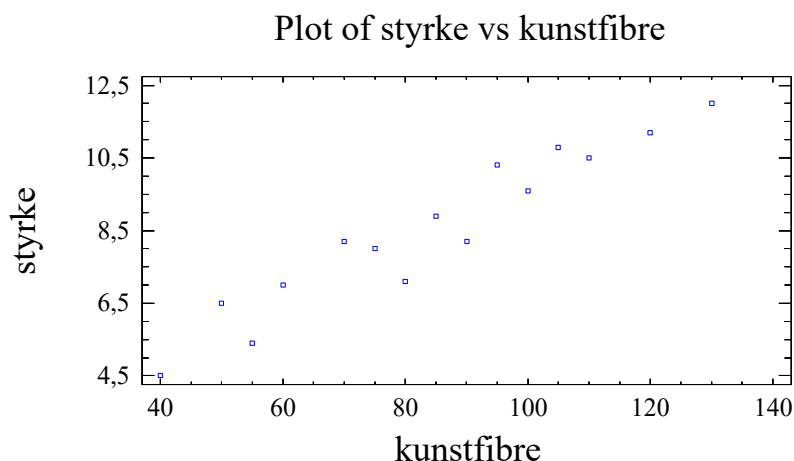
Som eksempler herpå kan nævnes $y = a + bx$ og $\ln y = a + b \cdot \ln x$.

9.2. BESTEMMELSE AF REGRESSIONSLIGNING

Vi betragter igen eksempel 9.1

Regressionslinie og regressionskoefficienter.

Afsættes de målte punktpar (x_i, y_i) i et koordinatsystem for at få et overblik over forløbet, fås følgende tegning:



Punkterne ligger ikke eksakt på en ret linie, men det synes rimeligt at antage, at afvigelserne fra en ret linie kan forklares ved den tilfældige variation (støjen).

Derfor er det nærliggende at antage, at **middelværdien** af den statistiske variable Y er en lineær funktion af x af formen $E(Y|x) = \beta_0 + \beta_1 x$. (1)

$E(Y|x)$ skal læses middelværdien af Y for fastholdt x .

Vi vil i det følgende ofte i ligningen (1) kort skrive Y eller μ fremfor $E(Y|x)$.

Ligning (1) kaldes **regressionsligningen** (eller den teoretiske regressionsligning), grafen kaldes for **regressionslinien** (eller den teoretiske regressionslinie), og konstantledet β_0 og hældningskoefficienten β_1 kaldes **regressionskoefficienterne**.

9. Regressionsanalyse

Mens middelværdien af Y ligger på regressionslinien, kan den aktuelle observerede værdi af Y ikke forventes at ligge på den.

For et punktpar (x_i, y_i) gælder derfor, at $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, hvor ε_i kaldes den i 'te residual.

Bestemmelse af regressionslinien ved mindste kvadraters metode.

På basis af en række sammenhørende værdier af x og y bestemmes estimater $\hat{\beta}_0$ og $\hat{\beta}_1$ for regressionskoefficienterne β_0 og β_1 ved "mindste kvadraters metode". Værdierne $\hat{\beta}_0$ og $\hat{\beta}_1$ kaldes de empiriske regressionskoefficienter.

Kan det ikke misforstås, så kort blot regressionskoefficienterne.

Det følgende eksempel viser metoden anvendt på et lille talekseksempel. Man benytter hertil et matrixprogram. De angivne metoder kan umiddelbart generaliseres til mere komplicerede eksempler.

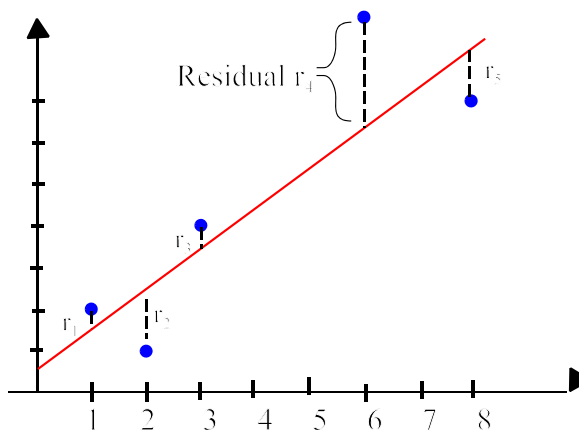
Eksempel 9.2. Bestemmelse af regressionskoefficienter ved mindste kvadraters metode.

I et medicinsk forsøg måles på en forsøgsperson sammenhørende værdier af en bestemt medicin i blodet (i %) og reaktionstiden. Resultaterne var:

x	1	2	3	6	8
y	2	1	4	9	7

Bestem ved mindste kvadraters metode et estimat for regressionslinien. ◆

Residual. Ved et punkts residual til en linie forstås den "lodrette" afstand fra punktet til linien (se tegningen).



Figur 9.1 Residualer

På figur 9.1 er afsat de 5 punkter, og indtegnet en ret linie.

Mindste Kvadraters metode. Regressionslinien $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ bestemmes som den af alle mulige rette linier, for hvilket summen af kvadratet af residualerne til linien er mindst.

I eksempel 9.2 er kvadratsummen $r_1^2 + r_2^2 + r_3^2 + r_4^2 + r_5^2$.

Løsningen af dette optimeringsproblem er angivet nedenfor.

Bestemmelse af regressionsligningen ved mindste kvadraters metode

LØSNING:

I vort tilfælde hvor vi har 5 punkter, indsættes vi disse i ligningen $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$. Dette giver:

$$2 = \hat{\beta}_0 + \hat{\beta}_1 \cdot 1, \quad 1 = \hat{\beta}_0 + \hat{\beta}_1 \cdot 2, \quad 4 = \hat{\beta}_0 + \hat{\beta}_1 \cdot 3, \quad 9 = \hat{\beta}_0 + \hat{\beta}_1 \cdot 6, \quad 7 = \hat{\beta}_0 + \hat{\beta}_1 \cdot 8.$$

De 5 ligninger med 2 ubekendte $\hat{\beta}_0$ og $\hat{\beta}_1$ kan i matrixnotation skrives:

$$Y = X \cdot B \text{ hvor } Y = \begin{pmatrix} 2 \\ 1 \\ 4 \\ 9 \\ 7 \end{pmatrix}, \quad X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 6 \\ 1 & 8 \end{pmatrix} \text{ og } B = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix}$$

De søgte værdier af $\hat{\beta}_0$ og $\hat{\beta}_1$ findes som den løsning til dette overbestemte ligningssystem som giver den mindste RMS - fejl.

Løsningen er (se evt. "Matematik for ingeniører" bind 3 eller Matricer og lineære ligninger afsnit 10) bestemt ved løsningen $B = (X^T \cdot X)^{-1} \cdot X^T \cdot Y$.

I vort taleksempel er

$$B = \left(\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 6 \\ 1 & 8 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 6 & 8 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 4 \\ 9 \\ 7 \end{pmatrix} = \begin{pmatrix} 0.6 \\ 1.0 \end{pmatrix}$$

Regressionsligningen bliver følgelig $\hat{y} = 0.6 + 1.0x$



I praksis vil man benytte et færdigt program til bestemmelse af regressionskoefficienterne. Dataene indtastes enten i et statistikprogram, eller i en lommeregner med regressionsprogram som TI-89 eller TI-Nspire.

9.3 VURDERING AF OM REGRESSIONSLIGNING BESKRIVER DATA GODT.

Vi vil i dette afsnit se på det tilfælde, hvor der for hver x - værdi kun er målt én y - værdi.

Det er altid muligt ved mindste kvadraters metode at finde en sådan "mindste kvadraters linie". Det betyder ikke nødvendigvis, at linien så også er en rimelig model, som kan anvendes til at beskrive sammenhængen.

Til vurdering heraf vil man

1) se på en **tegning**.

"Mindste kvadraters linie" tegnes i et koordinatsystem sammen med punkterne. Hvis den lineære model beskriver dataene godt, skal punkterne fordele sig tilfældigt omkring linien. I mere komplicerede tilfælde, er det nødvendigt i stedet at afsætte residualerne (i et såkaldt residualplot). Residualerne bør så fordele sig tilfældigt omkring den vandrette 0 - linie

- 2) Se på **størrelsen af forklaringsgraden** r^2 (også kaldet determinationskoefficient).

Den angiver et talmæssigt mål for hvor tæt punkterne ligger på linien.

Sædvanligvis finder man, at den fundne model på tilfredsstillende måde beskriver data, hvis forklaringsgraden er på over 70% samtidig med, at tegningen viser, at punkterne fordeler sig tilfældigt omkring den fundne regressionskurve.

Selv om forklaringsgraden er lav, kan der dog godt være en vis afhængighed mellem de to variable.

Eksempelvis, hvis man måler sammenhørende værdier af højde og vægt for alle værnepligtige mænd, vil forklaringsgraden være lav (under 50%). Ser man på den såkaldte korrelationskoefficient ρ vil den dog være positiv.

Hvis det ved en test viser sig at være signifikant større end 0 (svarer til at man tester om liniens hældningskoefficient er større end 0), har vi vist, at i middel vil der gælde, at jo højere man er jo mere vil man i middel veje.

Målepunkterne ligger imidlertid så spredt, at dette blot afspejler en tendens. Man kan ikke benytte regressionslinien til at forudsige noget med blot en rimelig sikkerhed.

- 3) Undersøge om der er **outliers**.

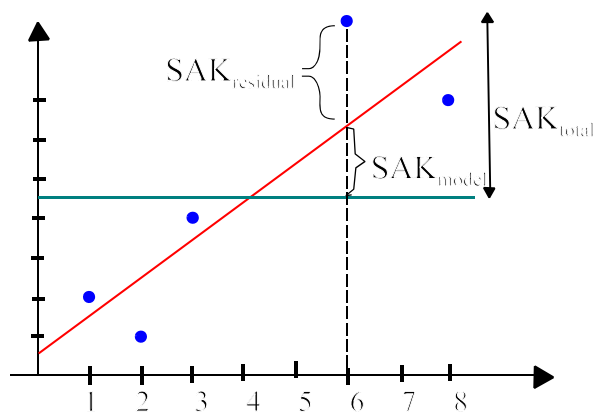
Outliers er en betegnelse for enkelte målinger som afviger så kraftigt fra den almindelige tendens, at det kan skyldes fejlmålinger. Sådanne punkter kan i uheldige tilfælde på grund af et stort bidrag til residualsommen få regressionslinien til at dreje. Det er dog klart, at man ikke blot kan stryge sådanne "ubehagelige" punkter. Det må kun ske, hvis man er sikker på, at punktet skyldes en fejl af en eller anden art ved målingen.

Definition og beregning af forklaringsgrad.

I praksis vil lade en lommeregner eller en PC med et statistikprogram beregne de enkelte statistiske størrelser.

Ved tolkningen af de fremkomne størrelser vil en anskuelig forståelse af størrelserne dog være nyttig.

I det følgende vil vi derfor definere nogle fundamentale definitioner, og søge at anskueliggøre dem dels på figur 9.3 dels ved at foretage beregningerne på tallene i eksempel 9.2.



Figur 9.3. SAK - størrelser

Definitioner:

SAK_{total} = sum af kvadrater af residualerne til den vandrette linie $y = \bar{y}$

$$\text{Data i eksempel 9.2 giver: } \bar{y} = \frac{2 + 1 + 4 + 9 + 7}{5} = 4.6$$

De 5 punkters "residualer" til den vandrette linie $y = \bar{y}$.

$$r_1 = 2 - 4.6 = -2.6, r_2 = 1 - 4.6 = -3.6, r_3 = 4 - 4.6 = -0.6, r_4 = 9 - 4.6 = 4.4, r_5 = 7 - 4.6 = 2.4.$$

$$\text{Vi får } SAK_{total} = r_1^2 + r_2^2 + \dots + r_5^2 = 2.6^2 + 3.6^2 + 0.6^2 + 4.4^2 + 2.4^2 = 45.2.$$

$SAK_{residual}$ = sum af kvadrater på de enkelte punkters afstand fra den fundne regressionslinje.

Af eksempel 9.2 fås følgende residualer til den fundne regressionslinje $y = 0.6 + 1.0x$:

$$r_1 = 2 - (0.6 + 1 \cdot 1) = 0.4, r_2 = 1 - (0.6 + 1 \cdot 2) = -1.6, r_3 = 4 - (0.6 + 1 \cdot 3) = 0.4,$$

$$r_4 = 9 - (0.6 + 1 \cdot 6) = 2.4, r_5 = 7 - (0.6 + 1 \cdot 8) = 0.4.$$

$$SAK_{residual} = r_1^2 + r_2^2 + \dots + r_5^2 = 0.4^2 + (-1.6)^2 + 0.4^2 + 2.4^2 + 0.4^2 = 11.2$$

SAK_{model} = sum af kvadrater af "regressionsliniens afstand" fra det totale gennemsnit \bar{y} .

Af eksempel 9.2 fås residualerne for regressionslinjen $y = 0.6 + 1.0x$'s "afstand" fra det totale gennemsnit $\bar{y} = 4.6$.

$$r_1 = 0.6 + 1 \cdot 1 - 4.6 = -3.0, r_2 = 0.6 + 1 \cdot 2 - 4.6 = -2.0, r_3 = 0.6 + 1 \cdot 3 - 4.6 = -1.0,$$

$$r_4 = 0.6 + 1 \cdot 6 - 4.6 = 2.0, r_5 = 0.6 + 1 \cdot 8 - 4.6 = 4.0.$$

$$SAK_{model} = r_1^2 + r_2^2 + \dots + r_5^2 = (-3.0)^2 + (-2.0)^2 + (-1.0)^2 + 2.0^2 + 4.0^2 = 34$$

Der gælder generelt, at $SAK_{total} = SAK_{residual} + SAK_{model}$ (jævnfør, at $45.2 = 11.2 + 34$)

Forklaringsgraden r^2 er bestemt ved

$$r^2 = \frac{SAK_{model}}{SAK_{total}} = \frac{SAK_{total} - SAK_{residual}}{SAK_{total}} = 1 - \frac{SAK_{residual}}{SAK_{total}} \quad (r^2 = \frac{34}{45.2} = 0.752)$$

Anskuelig forklaring:

Ligger punkterne tæt ved linien, er residualerne små, og $SAK_{residual}$ lille.

Nu er lille jo et relativt begreb, så man sammenligner med SAK_{total} , dvs. med residualerne til den vandrette linie $y = \bar{y}$. Hvis der ingen sammenhæng er mellem y - og x - værdierne (Y er uafhængig af x) vil regressionslinjen stille sig næsten vandret, dvs. $y \approx \bar{y}$.

Måler man eksempelvis sammenhørende værdier af personers højde, og deres IQ (intelligenskvotient) vil man uden tvivl få en ret linie der er praktisk taget vandret, og en forklaringsgrad der ligger tæt på 0.

Det betyder igen at $SAK_{residual} \approx SAK_{total}$ og dermed at $r^2 = 1 - \frac{SAK_{residual}}{SAK_{total}} \approx 0$.

Hvis derimod der er en sammenhæng (Y er lineært afhængig af x) vil regressionslinjen have en hældning forskellig fra nul. Det betyder igen at $SAK_{residual} \ll SAK_{total}$.

Heraf følger, at $r^2 \approx 1$.

Man siger også, at den fundne model "forklarer" $r^2 \cdot 100\%$ af den "totale variation"

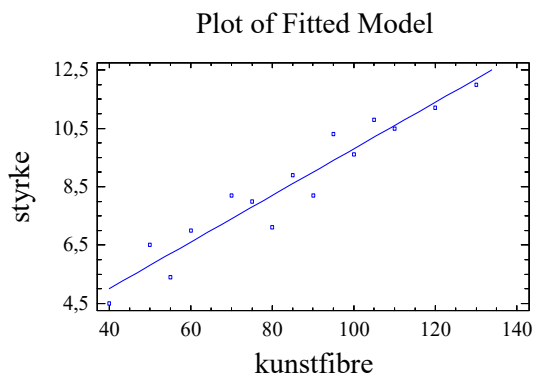
I eksempel 9.2 forklarer den fundne model således "75.2% af den totale variation."

Transformation. Hvis man ikke finder, at en ret linie beskriver data godt nok, så er det jo muligt, at en anden kurve bedre beskriver sammenhængen. Eksempelvis er det jo velkendt fra matematikken, at graferne for eksponentialfunktioner og potensfunktioner ved en passende logaritmisk transformation kan blive til rette linier. Det giver naturligvis lidt mere komplicerede regninger, men statistikprogrammer og også en del lommeregnerne kan dog let foretage en regressionsanalyse også i sådanne tilfælde. I eksempel 9.5 er et sådant eksempel gennemgået.

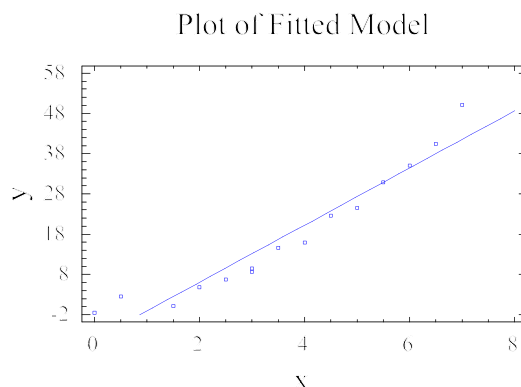
At man ikke alene kan stole på forklaringsgraden illustreres ved følgende eksempel.

Eksempel 9.3 .Grafisk vurdering af model.

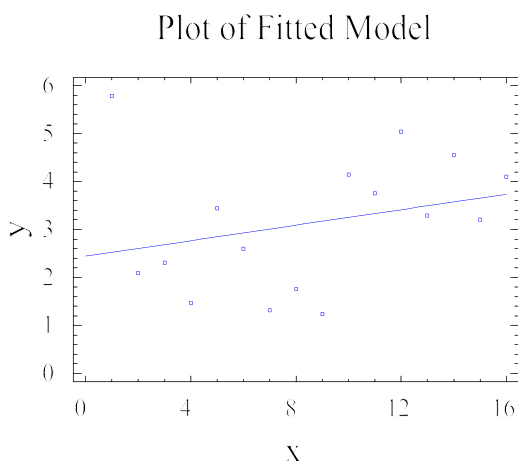
De følgende 4 figurer afspejler forskellige muligheder.



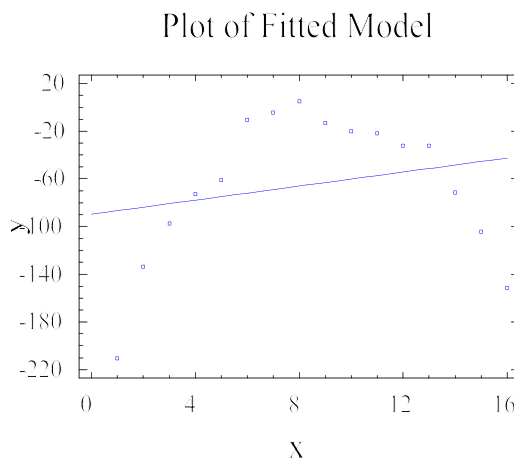
Figur 9.2a: $r = 0.959$ $r^2 = 91.9\%$



Figur 9.2b: $r = 0.962$ $r^2 = 92.6\%$



Figur 9.2c: $r = 0.278$ $r^2 = 7.73\%$



Figur 9.2d: $r = 0.229$ $r^2 = 5.24\%$

I figur 9.2a synes den lineære model at kunne beskrive dataene godt, idet punkterne fordeler sig tilfældigt omkring linien, og forklaringsgraden $r^2 = 91.9\%$ er høj.

I figur 9.2b er forklaringsgraden også høj, og punkterne ligger da også tæt ved linien. Imidlertid ligger punkterne ikke tilfældigt omkring linien. Yderpunkterne ligger over og de midterste punkter under linien, så det er næppe rimeligt at anvende en ret linie som model. I stedet kunne man overveje en eksponentialfunktion eller et andengradspolynomium.

I figur 9.2c er regressionslinien næsten vandret, og forklaringsgraden ringe.

En test må afgøre om hældningen er signifikant forskellig fra 0. Er dette ikke tilfældet, er der ingen relation mellem x og y (de er uafhængige)y.

Er x og y uafhængige vil punkterne fordele sig tilfældigt omkring gennemsnitslinien $y = \bar{y}$, og forklaringsgraden være 0.

I figur 9.2d er forklaringsgraden også lille, men alligevel må vi antage at der er en sammenhæng mellem x og y. Den er blot ikke lineær, men muligvis en parabel. ◆

Sammenhæng mellem korrelationskoefficient og forklaringsgrad.

Hvis både X og Y er normalfordelte statistiske variable (som eksempelvis når man aflæser sammenhørende værdier af højde og vægt for en række personer) angiver korrelationskoefficienten ρ en størrelse mellem -1 og 1 som kan anvendes til at angive, om der er en sammenhæng (korrelation) mellem X og Y . Er korrelationskoefficienten positiv har punkterne en voksende tendens, hvis den er negativ har de en aftagende tendens.

Et estimat for ρ er størrelsen r . Kvadreres den er r^2 den samme som forklaringsgraden.

Ekstrapolation. Selv om modellen synes på tilfredsstillende måde at beskrive data, så er det jo faktisk kun sikkert indenfor måleområdet. Man skal være yderst forsigtig med at ekstrapolere, dvs. på basis af modellen for x - værdier udenfor måleområdet beregne hvad y er.

Årsagssammenhæng. Selv om man finder, at der er en sammenhæng mellem x og y , er det ikke sikkert, at der er en årsagssammenhæng.

Der findes en god korrelation mellem antallet af storke i Sønderjylland i 1930-erne og antallet af børnefødsler (de faldt begge i samme takt), men det ene er nok ikke årsagen til det andet.

Man kender det også fra sammenhængen mellem kræft og tobaksrygning, hvor der i mange år var en diskussion om der ene bevirkede det andet, eller om det var en hel tredje faktor, der fik antallet af lungekræft til at stige.

Eksempel 9.4 Vurdering af model

Tilsætning af en vis mængde kunstfibre forøger et garns trækstyrke. Man har eksperimenteret med forskellige tilsatte mængder kunstfibre x og registreret garnets trækstyrke y ved disse forskellige mængder. Herved fremkom følgende observationsmateriale:

Mængde x (i gram) af kunstfibre p. kg uld	40	50	55	60	70	75	80	85	90	95	100	105	110	120	130
Trækstyrke : Y	4.5	6.5	5.4	7.0	8.2	8.0	7.1	8.9	8.2	10.3	9.6	10.8	10.5	11.2	12.0

- 1) Find r^2 og anvend denne samt en figur på lommeregnerens grafiske display eller residualernes fortegn til vurdering af modellen.
- 2) Opskriv regressionsligningen.

Løsning

- 1) **TI-Nspire** Data indtastes i "lister og regneark" ► Dokumentværktøjer(eller menu) ► Statistik ► Statistiske beregninger ► Lineær regression(mx+b) ► Udfyld menu ► ok

Der fremkommer konstanterne $m = 0.0799$, $b = 1.8087$ og $r^2 = 0.9193$

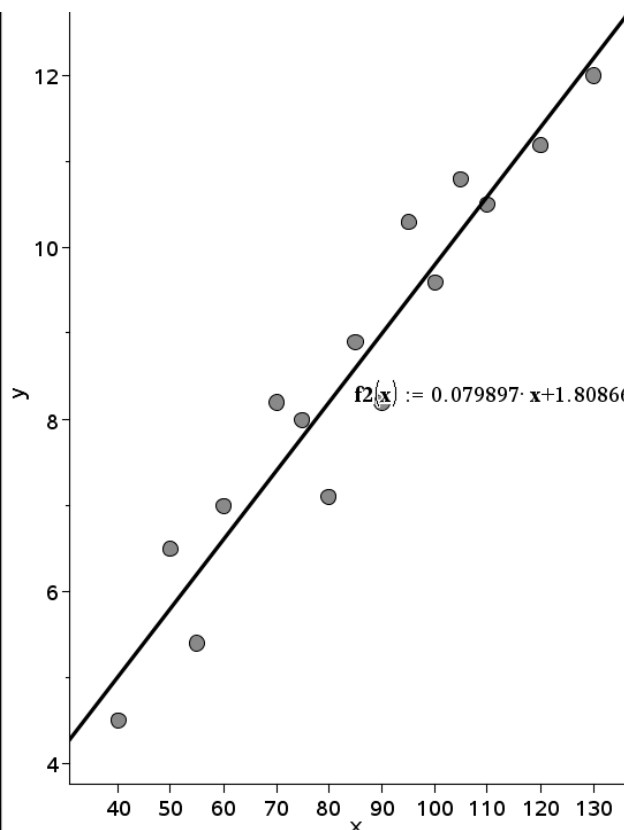
Marker begge lister ved at trykke på listebogstav (A) holde shift nede og gå mod højre ►.

Vælg data ► hurtiggraf ► Der viser sig så et punktplot ► Vælg "Undersøg data, Plotfunktion, skriv funktionen

Grafen tegnes nu sammen med punkterne.

9. Regressionsanalyse

A y	B x	C	D	E
				=LinRegMx(
4.5	40		Titel	Lineær regr...
6.5	50		RegEqn...	m*x+b
5.4	55		m	0.079897
7.	60		b	1.80866
8.2	70		r ²	0.919301
8.	75		r	0.958802
7.1	80		Resid	{-0.504550...
8.9	85			
8.2	90			
10.3	95			
9.6	100			
10.8	105			
10.5	110			
11.2	120			
12	130			



Tegningen viser, at punkterne fordeler sig tilfældigt omkring linien.

Hvis man ikke har mulighed for at vise tegningen kan man (hvis x-værdierne står i stigende rækkefølge) betragte residualerne:

TI-Nspire(PC): Sæt cursor øverst i tom kolonne, højre musetast ► variable ► kæd til ► resid

TI-Nspire lommeregner: Sæt cursor øverst i tom kolonne ► Tryk VAR ► Kæd til ► resid

TI89: residualer står i sidste kolonne i STAT/LIST

Fortegnene er - + - + + - - - - + - - - - -, dvs. de fordeler sig rimelig tilfældigt på begge sider af linien.

Outliers: Af tegningen ses, at ingen af punkterne afviger kraftigt fra linien.

Var dette tilfældet, kunne man teste om det virkelig er tilfældet ved at beregne de såkaldte “Studentized Residuals”

TI-Nspire: Vælg Statistiske tests ► Multipel lineær regressionstest ► udfyld menu ► ok

Vælg som ovenfor “kæd til” ► sresid, hvorved man ser en række tal

TI 89: Vælg F6 ► MultRegTtest ► udfylder menu ► Enter

Nu dukker en række nye lister op, og sresid er “Studentized Residuals”

Hvis ingen af disse transformerede residualer numerisk er større end 3 (det er tilladt, at op til 20% er større end 2) er der ingen outliers.

Vi ser, at ingen er større end 2, så konklusionen er, at der ikke er nogen outliers.

Da forklaringsgraden $r^2 = 0.919$ er tæt på 1, punkterne fordeler sig tilfældigt om linie, og der ikke er nogen outliers er den lineære model acceptabel.

- 2) Regressionskoefficienterne ses i den ovennævnte udskrift eller ved at vælge “Y=” hvoraf man finder $y = 1.8087 + 0.0799x$ ◆

9.4 TEST OG KONFIDENSINTERVALLER

De foregående betragtninger kræver (bortset fra kontrol af Outliers) ingen statistiske forudsætninger, idet man jo altid ved mindste kvadraters metode kan beregne regressionskoefficienterne, beregne forklaringsgrad, tegne kurver og punkter ind i et koordinatsystem og så herfra vurdere om modellen er acceptabel.

Forudsætninger for regressionsanalyse.

Ønsker vi at foretage en nøjere statistisk analyse som eksempelvis at teste "om Y er uafhængig af x , dvs. af om $\beta_1 = 0$ ", eller opstille konfidensintervaller for β_1 må observationerne opfylde de samme krav som ved variansanalyser, dvs.

1) Uafhængige målinger:

De enkelte observationer y_i er indbyrdes uafhængige (eksempelvis hvis der udføres flere målinger for samme mængde medicin skal de være indbyrdes uafhængige, ligesom det også skal gælde målinger baseret på forskellige mængder medicin.

Kravet kan opfyldes ved en hensigtsmæssig forsøgsplan. I eksempel 9.2 skal man således være sikker på at den foregående dosis medicin er ude af blodet inden man foretager en ny indsprøjtning, ligesom forsøgene skal være randomiseret. Man kan nok i dette tilfælde betvivle uafhængigheden, hvis man udfører forsøgene på samme person.

2) Residualerne skal være rimelig normalfordelt.

Analysen er robust overfor afvigelser, men er man i tvivl kan man indtegne residualerne på et normalfordelingsplot.

3) Varianshomogenitet.

Variansen af residualerne (eller af Y) skal være den samme uafhængig af x 's værdi.

Igen er analysen robust overfor afvigelser, så man vil sædvanligvis nøjes med at se på et residualplot, og deraf se, at residualerne fordeler sig jævnt uden store udsving.

Test af om Y er uafhængig af x

Lad os antage, at forudsætningerne er opfyldt, og at vi har fundet (ved at betragte tegning + forklaringsgrad og manglende outliers), at modellen $E(Y|x) = \beta_0 + \beta_1 x$ gælder.

Hvis Y er uafhængig af x betyder det, at regressionslinien er vandret, eller at hældningskoefficienten β_1 er 0.

Vi får altså: $H_0: Y$ er uafhængig af $x \Leftrightarrow H_0: \text{Regressionslinien er vandret} \Leftrightarrow H_0: \beta_1 = 0$.

Metode 1: F - test.

Ifølge definitionen af SAK_{model} angiver denne kvadratsummen af punkter på regressionsliniens afstand til det totale gennemsnit \bar{y} (se side 105)

Det er derfor klart, at hvis regressionslinien er næsten vandret, vil $SAK_{\text{model}} \approx SAK_{\text{residual}}$

Man kan vise, at $s_{\text{model}}^2 = \frac{SAK_{\text{model}}}{1}$ har frihedsgraden 1 og $s_{\text{residual}}^2 = \frac{SAK_{\text{residual}}}{N-1}$ har frihedsgraden $N-1$

Idet $F_{\text{model}} = \frac{s_{\text{model}}^2}{s_{\text{residual}}^2}$ kan vises, at en test for om linien er vandret er en sædvanlig F -test.

Metoden er forklaret mere udførligt i appendix 9.1.

H_0 forkastes, hvis P -værdi $= P(F > F_{\text{model}}) < \alpha$, hvor F_{model} er F -fordelt med en tællerfrihedsgrad på 1 og en nævnerfrihedsgrad på $N-2$.

Metode 2. t - test.

Metoden er forklaret mere udførligt i appendix 9.1.

Lad $t = \frac{\hat{\beta}_1}{s_{\beta_1}}$, hvor $s_{\beta_1} = \frac{|\hat{\beta}_1| \cdot s_{\text{residual}}}{s_{\text{model}}}$ er et estimat for spredningen på β_1 .

Det kan vises, at t er t - fordelt med $N - 2$ frihedsgrader.

H_0 forkastes, hvis P - værdi = $P(T > |t|) < \alpha$

På tilsvarende måde kan man teste $H_0: \beta_1 \geq 0$ og $H_0: \beta_1 \leq 0$ ved ensidede test.

Endvidere kan man analogt teste nulhypotesen $H_0: \beta = k$ hvor k er en konstant

Konfidensintervaller og prædistributionsintervaller.

Et led i analysen kan være, at udregne et 95% konfidensinterval for β_1 .

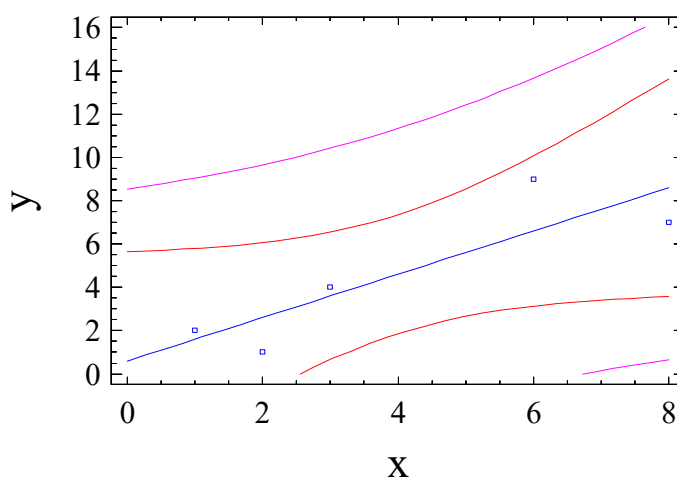
Endvidere vil man ofte være interesseret i en speciel værdi for x , for hvilken man ønsker beregnet såvel den tilsvarende “forventede” y - værdi (predicted value”) som et 95% -konfidensinterval for middelværdien $\tilde{\mu}$ og et 95% prædistributionsinterval for én ny observation.

På figur 9.4 er her tegnet 95% konfidensintervaller for middelværdierne (de inderste buede kurver), og 95% prædistributionsintervaller (de yderste to kurver).

Man ser tydeligt, at konfidensintervallerne er smallest omkring “centrum” (\bar{x}, \bar{y}) .

I appendix 9.1 er angivet formler for bl.a. disse konfidensintervaller.

Plot of Fitted Model



Figur 9.4. Konfidensintervaller og prædistributionsintervaller

Eksempel 9.5 (fortsættelse af eksempel 9.4) Test

Tilsætning af en vis mængde kunstfibre forøger et garns trækstyrke. Man har eksperimenteret med forskellige tilsatte mængder kunstfibre x og registreret garnets trækstyrke y ved disse forskellige mængder. Herved fremkom følgende observationsmateriale:

Mængde x (i gram) af kunstfibre p kg uld	40	50	55	60	70	75	80	85	90	95	100	105	110	120	130
Trækstyrke : Y	4.5	6.5	5.4	7.0	8.2	8.0	7.1	8.9	8.2	10.3	9.6	10.8	10.5	11.2	12.0

Det forudsættes, at forudsætningerne for regressionsanalyse er opfyldt.

I eksempel 9.4 fandt man at ligningen $y = 1.8087 + 0.0799x$ var en god model for data.

- 1) Test om y er uafhængig af x
- 2) Find 95% konfidensinterval for hældningen β
- 3) Find den til $x = 65$ svarende værdi for y , samt et 95% konfidensinterval for y .
- 4) Find 95% prædistributionsinterval for 1 ny observation svarende til x - værdien 65.

Løsning:

- 1) $H_0: Y$ er uafhængig af $x \Leftrightarrow H_0: \text{Regressionslinien er vandret} \Leftrightarrow H_0: \beta_1 = 0$.

Da vi skal teste, vælges nu

TI-Nspire: Vælg statistik ► Statistiske tests ► lineær regressions t-test ► udfyld menu ► ok

TI89: APPS ► STAT/LIST ► F6, A: Lin Reg T-Test ► menu udfyldes herunder Alternate Hyp=
 $\beta \& \rho \neq 0$

Vi får samme værdier af regressionskoefficienterne som i eksempel 9.4

Endvidere ses, at P - værdi $= 2 \cdot 10^{-8}$ svarende til $t = 12.1693$.

Da P - værdi < 0.05 forkastes $H_0: \beta = 0$ (stærkt)

Konklusion: Y er ikke uafhængig af x .

- 2) 95% konfidensinterval for β_1 :

TI-Nspire Vælg statistik ► Konfidensintervaller ► Lin.Reg t-intervaller ► vælg Hældning ► Udfyld menu ► ok

TI89:F7: LinRegTInt: Udfyld menu bl.a Interval = Slope:

Resultat: [0.0657;0.09413]


- 3) Konfidensinterval for y svarende til $x = 65$:

TI-Nspire: Som under punkt 2) men vælg svar

TI89:Som under punkt 2) men Interval=Response, x Value = 65

y 's værdi for $x = 65$: $\hat{y} = 7.002$, 95% konfidensinterval: [6.548 ; 7.456]

- 4) 95% Prædistributionsinterval for 1 ny observation svarende til x - værdien 65.

TI89+ TI-Nspire Som under punkt 3. Se længere nede i udskrift Resultat: [5.53 ; 8.474]. 

9.5. TRANSFORMATION AF DATA.

Da man altid vil foretrække den simplest mulige model, er modellen $y = ax + b$ altid den, man starter med at anvende.

Hvis man ser, at punkterne ikke ligger tilfældigt omkring linien, men dog synes at følge en krum kurve, så må man anvende en anden model.

Blandt de oftest forekommende modeller er

- 1) Eksponentialmodellen $y = a \cdot b^x$

Tages logaritmen på begge sider fås $\ln(y) = \ln(a) + \ln(b) \cdot x$

som er en normal retlinet model, hvis man erstatter $\ln(y)$ med z

- 2) Potensmodellen $y = a \cdot x^b$

Tages logaritmen på begge sider fås $\ln(y) = \ln(a) + b \cdot \ln(x)$

som er en normal retlinet model, hvis man erstatter $\ln(y)$ med z og $\ln(x)$ med v .

- 3) Logaritmemodellen $y = a + b \cdot \ln(x)$

som er en normal retlinet model, hvis man erstatter $\ln(x)$ med v .

9. Regressionsanalyse

Generelt gælder, at man så vidt mulig foretrækker modeller med kun 2 parametre a og b (som (1), (2) og (3)), fremfor modeller som eksempelvis andengradsmodellen $y = ax^2 + bx + c$ der indeholder 3 parametre, da de mest stabile (set fra et statistisk synspunkt).

I appendix 9.3 er angivet en liste med kommentarer over de mest almindelige transformationer.

Har man på forhånd en viden om, at en bestemt transformation skal anvendes, kan man uden større besvær foretage den pågældende transformation på dataene og så udføre regressionsanalysen på de transformerede data, på samme måde som vist i eksempel 9.5.

Dette illustreres ved følgende eksempel.

Eksempel 9.6. Valg mellem lineær og eksponentiel model

I et forsøg undersøgte et ventilationsanlægs effektivitet. Målingerne foretoges ved at fylde et lokale med gas og vente til koncentrationen var stabil. Herefter startedes ventilationsanlægget og gaskoncentrationen C_t målte til forskellige tidspunkter t .

Følgende resultater fandtes:

t (min. efter anlæggets start)	2.67	4.59	6.75	7.67	11.34	14.34	16.25	18.25	23.09
C [ppm]	34	28	26	22	16	14	12	10	8

Følgende 2 modeller for funktioner overvejes:

Model 1 (lineært henfald): $C = a + b \cdot t$

Model 2 (eksponentielt henfald): $C = a \cdot e^{b \cdot t}$

- 1) Vurder hvilken model der er bedst.
- 2) Opskriv regressionsligningen for den model du finder bedst.
- 3) Beregn ud fra den valgte model den værdi af C , for hvilken $t = 12$ minutter, og opskriv et 95% konfidensinterval for C .

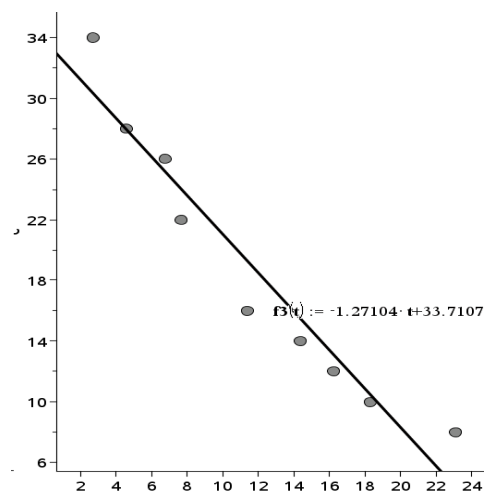
Løsning:

- 1) **TI-Nspire:** Data indtastes i "lister og regneark" ► Dokumentværktøjer ► Statistik ► Statistiske beregninger ► Lineær regression ► Udfyld menu ► ok

TI89: APPS, STAT/LIST ► data indtastes i to tomme lister (kan eventuelt give dem navne som t og c)
F4: Calc, 3. Regressions, 1:linReg(a+bx), Udfylder lister,

Af udskriften fås umiddelbart $r^2 = 0.9293$

Endvidere tegner vi kurve og linie (sker som i eksempel 9.4)

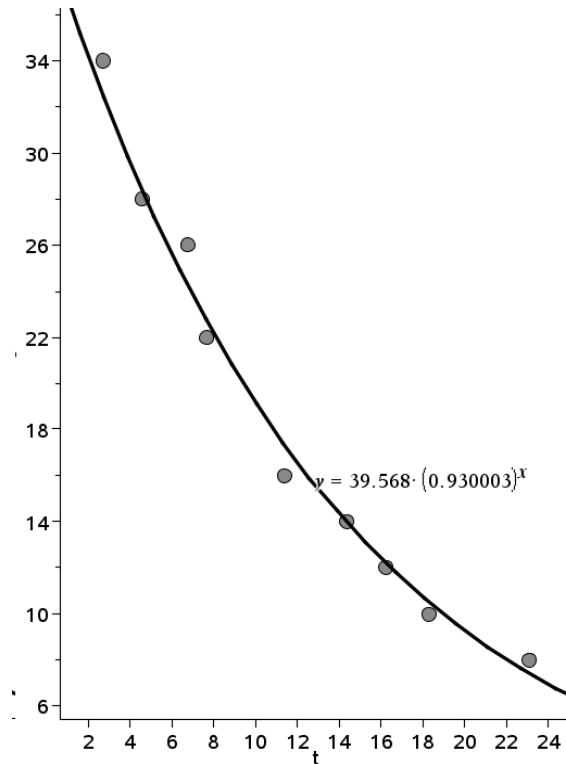


Forklaringsgraden er høj, men punkterne ligger ikke tilfældigt om linien, så modellen kan næppe anvendes udenfor måleområdet.

Vi kunne i stedet betragte residualernes fortegn: + + + - - - - +

Vi gentager nu ovenstående, idet vi nu vælger ExpReg (eksponentiel regression)
(PowerReg er potensmodellen)

Af udskriften fås umiddelbart $r^2 = 0.9883$



Af tegningen ses, at punkterne ligger tæt og "tilfældigt" omkring kurven.

Alternativt kan man betragte residualernes fortegn: + - + - - + - - +

Da forklaringsgraden er højere, og residualerne ligger tilfældigt omkring linien vælges model 2

2) Af udskriften ses, at modellen er $C = 39.57 \cdot 0.93^t$

3) Da vi nu skal til at beregne konfidensintervaller må vi transformere ligningen

$$C = a \cdot b^t \Leftrightarrow \ln(C) = \ln(a) + \ln(b) \cdot t$$

Vi må nu danne en søjle for $\ln(C)$

TI-Nspire: en søjle navngives z ► Placer cursor i rubrikken under z og marker cellen så et lighedstegn fremkommer ► skriv $\ln(c)$ ► ENTER

Vælg statistik ► Konfidensintervaller ► Lin.Reg t-intervaller ► svar ► Udfyld menu med t = 12 ► ok

TI89: APPS, STAT/LIST ► en søjle navngives z ► marker navnet og forinden skrives $\ln(c)$ ENTER
F7: linRegTinteg(a+bx), Udfylder lister med t og z ► Interval: Vælg Response ► x-value =12

Vi får $\hat{y} = 2.80722$, dvs. for $t = 12$ er $c = e^{2.80722} = \underline{\underline{16.56}}$

(Kontrol= Indsættes $t = 12$ i regressionsligningen fås det samme)

95% konfidensinterval $[e^{2.762}; e^{2.853}] = \underline{\underline{[15.83; 17.34]}}$ ◆

9.6. ENKELT REGRESSIONSANALYSE MED FLERE y -VÆRDIER FOR HVER x - VÆRDI.

Hvis man selv fastlægger sine x - niveauer, er det ofte muligt for hver x - værdi, at foretage flere målinger af y - værdien. Vi siger kort at analysen er “med gentagelser”. Sædvanligvis er det bedre at nøjes med en y -værdi for hver x -værdi, og så til gengæld have flere forskellige punkter, men i tilfælde, hvor man er i alvorlig tvivl om forudsætningen om varianshomogenitet gælder, eller hvor man vil teste om en lineær model gælder, kan det være relevant med eksempelvis to gentagelse for hver x -værdi.

Dette er således tilfældet i følgende eksempel:

Eksempel 9.7 Regressionsanalyse med gentagelser

Metalpladers overflader oxideres i en ovn ved 200° C. Med henblik på en undersøgelse af sammenhængen mellem det oxiderede lags tykkelse y (i ångstrøm) og tiden t (i minutter) foretog man følgende målinger:

Tiden t	20	30	40	60	70	90	100	120	150	180
Tykkelse y	4.2 4.9	7.4 6.9	8.8 8.2	13.6 12.0	13.1 12.4	14.9 16.8	20.0 21.2	23.1 25.2	27.5 25.1	32.9 32.4



Fordelen herved er, at man nu kan få et estimat for forsøgsfejls spredning (“støjen”), som kan anvendes til at teste, om den lineære model kan accepteres, når man tager støjen i betragtning.

Endvidere kan man, hvis man finder det nødvendigt, teste om der er varianshomogenitet. Alle andre test udføres på samme måde som beskrevet i forrige afsnit.

Forklaring af metode og formler

Test af model.

For hver x - værdi beregnes gennemsnittet af de dertil hørende y - værdier. Disse “gennemsnitspunkter” bør ligge tæt på linien hvis modellen er god. Hvis modellen er den rigtige, så er den eneste grund til at “gennemsnitspunkterne ikke ligger eksakt på linien, at der er støj.

Vi kan derfor beregne et estimat (kaldet $s_{\text{lack of fit}}^2$) for variansen af denne støj ud fra de afvigelser som gennemsnitspunkterne har. Hvorledes denne beregnes ses i appendix 9,2.

Da vi samtidig ud fra gentagelserne kan beregne et andet estimat for støjen (kaldet s_{error}^2), har vi mulighed for at teste de to varianser mod hinanden, ved en sædvanlig F - test

$$F_{\text{lack of fit}} = \frac{s_{\text{lack of fit}}^2}{s_{\text{error}}^2}$$

Får vi her en forkastelse, kan “gennemsnitspunkternes” afvigelser fra linien ikke forklares alene ved støjen, og vi må derfor forkaste modellen.

Test af varianshomogenitet.

Som tidligere nævnt, er analysen robust overfor afvigelser fra kravet om varianshomogenitet (konstant varians σ^2), hvis der er lige mange gentagelser (som i dette forsøg). Man vil derfor kun foretage en vurdering af dette krav, hvis man ud fra forsøgets natur mener, at varianserne kan tænkes at være voldsomt forskellige.

Testen kan foretages præcist som beskrevet under ensidet variansanalyse (se appendix 4.1).

Eksempel 9.7. Regressionsanalyse (med gentagelser)

Givet følgende målinger

Tiden t	20	30	40	60	70	90	100	120	150	180
Tykkelse	4.2	7.4	8.8	13.6	13.1	14.9	20.0	23.1	27.5	32.9
y	4.9	6.9	8.2	12.0	12.4	16.8	21.2	25.2	25.1	32.4

- 1) Det formodes på forhånd, at der er en lineær sammenhæng mellem x og y .
Undersøg ved en "lack of fit" test, om formodningen kan accepteres.
- 2) Bestem i bekræftende fald ligningen for den fundne regressionslinie.
- 3) Det påstås i litteraturen, at hældningskoefficienten β_1 er 0.15
Test om dette på et signifikansniveau på 5% kan være sandt.
- 4) Angiv et 95% konfidensinterval for middelværdien af tykkelsen y , når $t = 110$ minutter.

Løsning :

Data indtastes vandret, ved at man opretter 10 lister med navnene t1, t2 osv.

Tiden t	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10
Tykkelse	4.2	7.4	8.8	13.6	13.1	14.9	20.0	23.1	27.5	32.9
y	4.9	6.9	8.2	12.0	12.4	16.8	21.2	25.2	25.1	32.4

- 1) H_0 : Lineær model gælder $\Leftrightarrow H_0: (x_i, \mu_i)$ ligger på en ret linie .

TI-Nspire: Data indtastes i "lister og regneark" ► Statistik ► Statistiske test ► Ensidede variansanalyse ► Udfyld menuer ► ENTER

TI89: F6, C:ANOVA, ► Data Input : Data ► Antal grupper = 10, ENTER, Udfyld listnavne t1, t2 osv. ENTER.

Titel	Ensidede variansana..
F	170.934
PVal	9.42255E-10
df	9.
SS	1509.17
MS	167.686
dfError	10.
SSError	9.81
MSError...	0.981
sp	0.990454
CLower...	{2.9895072492971..
CUpper...	{6.1104927507029..
\bar{x} List	{4.55,7.15,8.5,12....

9. Regressionsanalyse

Data indtastes "lodret"

list1	list2
20	4.2
20	4.9
30	7.4
30	6.9
osv.	osv.

Laver regressionstest på de to lister list 1 og list 2

TI-Nspire: Statistik ► Statistiske test ► Lineær Regresions t-test ► udfyld menu ► ENTER

TI89: APPS, STAT/LIST ► F6, A:LinRegTest ► Udfyld list1 = list 1 og list2 = list 2, Alternate Hyp=
 $\beta \& \rho \neq 0$ ► ENTER

Titel	Lineær r...
Alternativ H...	$\beta \& \rho \neq \dots$
RegEqn	a+b*x
t	28.6756
PVal	1.77861...
df	18.
a	1.65415
b	0.17297...
s	1.3445
SESlope	0.00603...
r ²	0.97857...
r	0.98923...
Resid	{-0.9136..

Heraf: $s = s_{\text{residual}} = 1.3445$, $df = f_{\text{residual}} = 18$ og $SAK_{\text{residual}} = (1.3445)^2 \cdot 18 = 32.5382$

I den forrige udskrift fandtes $df_{\text{error}} = 10$, $SK_{\text{error}} = 9.81$ og dermed $s_{\text{error}}^2 = MS_{\text{error}} = 0.981$

$SAK_{\text{lack of fit}} = SAK_{\text{residual}} - SAK_{\text{error}} = 32.538 - 9.81 = 22.728$

Vi kan nu udfylde variansanalysetabellen

Variation	SAK	f	$s^2 = \frac{SAK}{f}$	$F = \frac{s_{\text{lack}}^2}{s_0^2}$	P-værdi
Lack of fit	22.728	8	2.841	2.896	$P(F > 2.896) = FCdf(2.896, \infty, 8, 10) = 0.0591$
Error	9.81	10	0.981		
Residual	32.538	18	1.3445		

Da P -værdi = 0.0591 > 0.05 accepteres H_0 , dvs.

vi vil i det følgende antage, at den lineære model gælder.

3) I ovenstående udskrifter finder man konstanterne

Den empiriske regressionslinie bliver: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot t = 1.654 + 0.1730 \cdot t$

4) $H_0: \beta_1 = 0.15$

TI-Nspire: Statistik ► Konfidensintervaller ► Lineær Reg t-intervaller ► udfyld menu bl.a. hældning ► ENTER

TI89: F7: LinRegTInt ► Udfyld menu: Interval=Slope

95% konfidensinterval for β : [0.160 ; 0.186]

Da 0.15 ikke ligger i intervallet forkastes H_0 , dvs.

data giver ikke den i litteraturen angivne hældningskoefficient..

5) 95% konfidensinterval for middelværdien af tykkelsen y , når $t = 110$ minutter.

TI-Nspire: Statistik ► Konfidensintervaller ► Lineær Reg t-intervaller ► udfyld menu bl.a. svar

TI89: F7: LinRegTInt ► Udfyld menu: Interval=Response, x Value = 100

$\hat{y} = y$'s værdi for $x = 110 = \underline{20.68}$.

95% konfidensinterval [19.98 ; 21.38]

Det skal bemærkes, at hvis man vil beregne forklaringsgraden for modellen, skal det gøres ud fra gennemsnittene dvs. lave to lister med henholdsvis gennemsnit og y -værdier

9.7 MULTIPLE REGRESSIONSANALYSE.

9.7.1 Indledning

Vi vil i dette afsnit behandle det tilfælde, hvor der indgår mere end 1 kvantitativ variabel. Vi vil begrænse os til at se på modeller, hvor de variable indgår lineært. Et eksempel herpå er modellen $Y = \alpha_0 + \beta_1 x_1 + \beta_2 x_2$, hvor parametrene er α_0, β_1 og β_2 .

I appendix 9.4 findes de nødvendige matrixformler til beregningerne. Det er dog langt lettere at have et statistikprogram til rådighed, så eksemplerne er regnet med Ti ...

9.7.2 Multipel regressionsanalyse med én y - værdi for hver x - værdi.

Man vurderer om modellen er acceptabel ved

1) at se på "forklaringsgraden" r^2

2) at se om der er outliers (se om "studentized residuals" alle numerisk er under 3, og kun få er over 2)

Vi vil illustrere metoden ved følgende eksempel.

Eksempel 9.8 (multipel regressionsanalyse uden gentagelser)

Det månedlige elektriske forbrug Y på en fabrik formodes at være afhængig af den gennemsnitlige udendørs temperatur x_1 , antal arbejdsdage x_2 i måneden, den gennemsnitlige renhed x_3 af det fremstillede produkt og det antal tons x_4 , der produceres i den pågældende måned. Det formodes, at Y er en lineær funktion af x_1, x_2, x_3 og x_4 , dvs. på formen

$$Y = \alpha_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4.$$

Følgende observationer fra det forløbne år foreligger

x_1	x_2	x_3	x_4	Y
-4	22	91	100	836
-1	20	90	95	789
7	21	88	110	883
16	19	87	88	790
18	20	91	94	816
23	19	94	99	859
27	23	87	97	831
29	21	86	96	832
24	22	88	110	897
16	23	91	105	872
10	20	90	100	842
3	20	89	98	821

- 1) Vurder ud fra forklaringsgraden og "studentized residualer" om ovennævnte model er rimelig. Det antages i det følgende, at ovenstående model gælder.
- 2) Undersøg om modellen kan reduceres, dvs. kan nogle af koefficienterne antages at være 0.
- 3) Angiv regressionsligningen i den endelige model.
- 4) Angiv 95% konfidensintervaller for de regressionskoefficienter der indgår i ovenstående model
- 5) Angiv et 95% konfidensinterval for Y i punktet $(x_1, x_2, x_3, x_4) = (0, 20, 90, 100)$

Løsning :

- 1) **TI-Nspire:** Data indtastes i "lister og regneark" ► Statistik ► Statistiske tests ► Multipel lineær regressionstest ► udfyld menuer ► ENTER
TI89: APPS, STAT/LIST ► indtast data x_1 i list 1 osv., x_4 i list 4 og y i list 5 ► F6:Test ► B:MultReg-Tests ► Udfyld menuen ► ENTER

Man får følgende udskrift, hvoraf man ser, at $r^2 = 0.9654$

I listen sresid (findes på samme måde som i eksempel 9.4), er "Studentized residuals".

Da kun en enkelt værdi numerisk er større end 2 og ingen er over 3, antages, at der ikke er outliers

Da forklaringsgraden er tæt ved 1 og der ikke er outliers vurderes modellen at være rimelig god.

G	H	I stat1.sresid	J stat1.plist
	=MultRegTests		
Titel	Multipel lineæ...	1.79004	0.167138
RegEqn	b0+b1*x1+b2...	-2.25477	0.002712
F	48.899	0.13626	0.719466
PVal	0.000034	1.4978	0.139262
R ²	0.965448	-0.633566	0.000011
adjR ²	0.945705	0.527484	
s	7.90936	-0.653346	
DW	2.73501	0.114698	
dfReg	4.	-0.330042	
SSReg	12236.1	0.207811	
MSReg	3059.02	0.252362	
dfError	7.	-0.012553	
SSError	437.906		
MSError	62.558		
bList	{175.4994882...		
tList	{1.541326670...		
PList	{0.167138484...		
SEList	{113.8626169...		
ŷList	{826.1022790...		
Resid	{9.897720971...		
sResid	{1.790039690...		

2) $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, H: Mindst en af regressionskoefficienterne er forskellig fra 0.

I Udskriften fra MultRegTests findes en P -værdi på 0.00003

Da P -værdi = 0.00003 < 0.05 forkastes H_0 (stærkt), dvs. mindst en af regressionskoefficienterne er forskellig fra 0.

I P-list findes P -værdierne for de enkelte regressionskoefficienter.

Den er i ovenstående udskrift flyttet op i en søjle ved at anvende "kæd til", mens den i TI89 findes efter listerne.

Man ser, at nr. 3 giver den største P -værdi på 0.7195. Da den første P -værdi svarer til konstantleddet svarer nr. 3 til x_2 .

$H_0: \beta_2 = 0$ accepteres, da P -værdien = 0.7195 > 0.05.

x_2 - leddet bortkastes. **Bemærk, at man kun eliminerer én variabel ad gangen**

TI89+TI-Nspire: Vælg igen MultRegTests, og udfyld menuen med kun 3 variable.

Man ser, at nu er den største P -værdi ud for x_3 og P -værdien = 0.0849.

$H_0: \beta_3 = 0$ accepteres da, da P -værdien = 0.0849 > 0.05

x_3 slettes nu af modellen.

TI89+TI-Nspire: Vælg igen MultRegTests, og udfyld menuen med kun 3 variable.

Nu ses, at alle P -værdier er mindre end 0.05, dvs. modellen kan ikke reduceres mere.

3) I listen blist findes koefficienterne, dvs. $\hat{y} = \underline{\underline{335.65 + 0.9017 x_1 + 4.94 x_4}}$.

$$4) 95\% \text{ konfidensinterval for } \beta_1 : \left[\hat{\beta}_1 - t_{1-\frac{\alpha}{2}}(f_{\text{residual}}) \cdot s_{\hat{\beta}_1} ; \hat{\beta}_1 + t_{1-\frac{\alpha}{2}}(f_{\text{residual}}) \cdot s_{\hat{\beta}_1} \right]$$

$$s_{\hat{\beta}_1} \text{ findes i "selist"} = 0.23495. \quad f_{\text{residual}} = df_{\text{err}} = 9$$

$$[0.9017 - t_{0.975}(9) \cdot 0.23495; 0.9017 + t_{0.975}(9) \cdot 0.23495] = [0.371; 1.432]$$

Tilsvarende findes $s_{\hat{\beta}_4}$ i "selist" = 0.40245 og dermed 95% konfidensinterval for β_4

$$[4.94 - t_{0.975}(9) \cdot 0.40245; 4.94 + t_{0.975}(9) \cdot 0.40245] = [4.03; 5.85]$$

- 5) **TI-Nspire**: Statistik ► Konfidensintervaller ► Multipl Reg Int ► Udfyld menuer ► Enter
TI89: F7:Ints, 8: MultRegInt, Udfyld menuer heraf x Value List={0,100}, ENTER

Titel	Multipel...
RegEqn...	b0+b1*x...
\hat{y}	829.67
dfError	9.
CLower	820.348
CUpper	838.992
ME	9.32211
SE	4.12089
LowerPr...	808.14
UpperPr...	851.199
MEPred...	21.5296
SEPred	9.51728
bList	{335.65...
Resid	{9.9370...

Man finder 95% konfidensintervallet : [820.34 ; 838.99]

9.8 POLYNOMIAL REGRESSIONSANALYSE

9.8.1 Indledning

Ved en polynomial regressionsanalyse er den statistiske model

$$Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \dots + \beta_p x^p.$$

hvor den variable Y skal opfylde de sædvanlige regressionsforudsætninger.

Som det ses, er den i afsnit 9.2 betragtede enkelte regression et specielt tilfælde. Den statistiske analyse da også meget beslægtet hermed.

Det man søger er altid den "enkleste" model der giver en tilstrækkelig god beskrivelse af Y indenfor det foreliggende variationsområde for x . Ud fra et statistisk synspunkt, vil man altid foretrække den model med de færreste parametre, da de på samme datamateriale giver en sikrere bestemmelse af parametrene. At andet lige vil man derfor foretrække de i afsnit 9.3 nævnte "transformerede" modeller som alle kun har 2 parametre fremfor eksempelvis et andengrads-polynomium $Y = \beta_0 + \beta_1 x + \beta_2 x^2$ hvor man skal bestemme tre parametre β_0 , β_1 og β_2 .

Blandt polynomierne vil man naturligvis foretrække et af lavest grad.

Programmerne tegner let polynomier af 2. grad, 3 grad og 4 grad og regner forklaringsgraden ud.

Derved kan man få et indtryk af hvilket polynomium, af lavest grad, der giver en rimelig beskrivelse af data.

Imidlertid er man altid nødt til at teste modellen, og derfor er fremgangsmåden den, at polynomiet omskrives og man går over til multipel regression.

Eksempelvis omskrives polynomiet $T = a_0 + b_1x + b_2x^2 + b_3x^3$
til $T = a_0 + b_1x_1 + b_2x_2 + b_3x_3$ ved at sætte $x_1 = x, x_2 = x^2, x_3 = x^3$

Vi vil i dette kapitel kun betragte tilfældet hvor der til hver x-værdi svarer en y-værdi.

Modificeret forklaringsgrad (adjusted)

Et 17-gradspolynomium vil gå eksakt gennem de 18 punkter, og $r^2 = 100\%$. Det er imidlertid klart, at en sådan model dels er alt for komplicerede til de fleste praktiske formål, dels følger kurven alle de tilfældige variationer, som vi netop ikke bør tage hensyn til.

Betragter man nu eksempelvis polynomiet $Y = \beta_0 + \beta_1x + \beta_2x^2$, beregner forklaringsgraden r^2 , derefter betragter et polynomium af tredje grad, og beregner igen forklaringsgraden, så er den steget, selv om tredjegradsmodellen måske ikke er væsentlig bedre.

Det er altid muligt at øge r^2 ved at addere flere led til modellen.

For hvert led der tilføjes mistes der en frihedsgrad i "error", og hvis SAK for det nye led ikke giver et væsentligt bidrag kan det betyde, at den nye model er ringere end den gamle model. For at tage hensyn til dette, betragtes ofte et modificeret r^2 (R-squared (adjusted for d.f.¹))

Når r^2 ikke stiger væsentligt, og R-squared (adjusted for d.f.) næsten er konstant, eller begynder at falde, er man tæt ved den "bedste" model.

Fremgangsmåde

Man starter med at tegne punkterne, eventuelt sammen med en lineær model af første grad $y = \alpha + \beta_1x$ (se eksempel 9.4)

Er punkterne i et interval stigende og i et andet interval faldende så de eksempelvis kunne beskrive en parabel, så vil det være rimeligt at se på et polynomium af anden eller højere grad.

Der er i princippet to fremgangsmåder:

Forward:

Man ser på en model af anden grad $y = \alpha + \beta_1x + \beta_2x^2$ og sammenligner den med en af tredje grad $y = \alpha + \beta_1x + \beta_2x^2 + \beta_3x^3$

Dette sker ved at man sammenligner de to "adjusted forklaringsgrader", samt teste nulhypotesen $H_0: \beta_3 = 0$

Hvis man får en accept af H_0 og endvidere at "adjusted r^2 " for tredjegradspolynomiet ikke er væsentlig højere end den af anden grad, så vil man vælge andengradspolynomiet.

Man vil dog så for andengradspolynomiet også se om der er outliers (studentized residuals).

Hvis man omvendt får en forkastelse af $H_0: \beta_3 = 0$ vil man sammenligne tredjegradspolynomiet med et polynomium af fjerdegrad.

¹ r^2 (adjusted) = $\frac{(n-1) \cdot r^2 + 1 - k}{n - k} = \frac{s_{total}^2 - s_0^2}{s_0^2}$, hvor k er antal parametre i modellen (incl konstantled).

Backward

Her starter man med en model af fjerde grad $y = \alpha + \beta_1x + \beta_2x^2 + \beta_3x^3 + \beta_4x^4$

Man tester nulhypotesen $H_0: \beta_4 = 0$

Får man en accept af nulhypotesen ser man nu på en model af trediegrad, og tester analogt $H_0: \beta_3 = 0$

Endvidere sammenligner man de to "adjusted r²" for at se om der er sket en væsentlig ændring. Således fortsættes indtil man får en forkastelse af nulhypotesen, samt ser, at "adjusted r²" ikke har ændret sig væsentlig.

De to metoder kan godt give forskelligt resultat, da der jo er et element af personlig vurdering i konklusionen.

For det polynomium man ender med, ser man derfor på outliers, og residualplot for at være sikker på, at modellen er acceptabel..

9.8.2. Beregning af polynomial regressionsanalyse

Eksempel 9.9. Polynomial regressionsanalyse uden gentagelser.

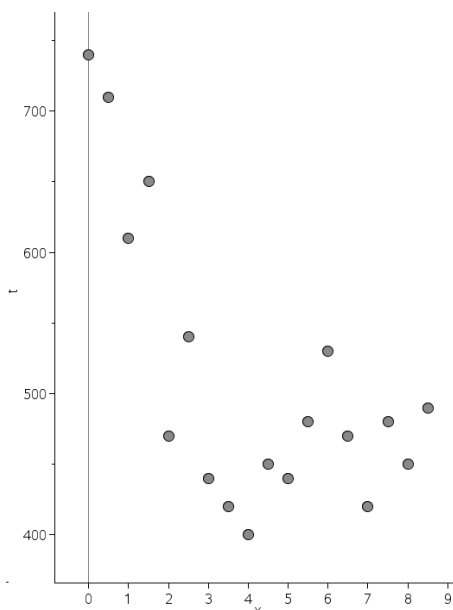
Et forsøg udføres, for at finde hvordan størkningstiden T (i minutter) afhænger af antal gram x af et additiv. Man fik følgende forsøgsresultater:

x g/l	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5
T min.	740	710	610	650	470	540	440	420	400	450	440	480	530	470	420	480	450	490

- 1) Vurder på basis af ovennævnte observationer, hvilket polynomium $T = \alpha_0 + \beta_1x + \beta_2x^2 + \beta_3x^3 + \dots + \beta_px^p$ af lavest mulig grad p, der indenfor måleområdet [0 ; 8.5] giver en tilfredsstillende beskrivelse af T's variation.
- 2) Angiv regressionsligningen for den model, man i spørgsmål 1 har fundet frem til.
- 3) Beregn værdien af T for x = 6.2, og angiv et 95% konfidensinterval for T for x = 6.2.

Løsning:

Først betragtes en tegning af punkterne samt evt. en ret linie (jævnfør eksempel 9.4).



Det ses tydeligt, at punkterne viser både en opadgående og nedadgående tendens, så et polynomium må være den rette model

Vi vælger "Forward-metoden"

For at teste må man omforme regressionsligningen til en lineær funktion i de variable .

For ikke at skulle gentage indtastningen vælger vi for en sikkerheds skyld modellen

$$T = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \text{ ved at sætte } x_1 = x, x_2 = x^2, x_3 = x^3, x_4 = x^4$$

TI-Nspire: Lister og regneark ► dan lister x, y, , x2 = x^2, x3 = x^3, x4=x^4

Statistik ► statistiske tests ► Multipel lineær regresionstest ► Udfyld menuer ► Enter

TI89: APPS, STAT/LIST ► dan nye lister x2 = x^2, x3 = x^3, x4=x^4

F6:tests ► B: MultRegIn ► Udfyld menuer

Først starter vi med en model af anden grad.

TI-Nspire: Statistik ► statistiske tests ► Multipel lineær regresionstest ► Udfyld menuer ► Enter

TI89: F6:tests ► B: MultRegIn ► Udfyld menuer

1)

Af udskriften fås $r^2 = 0.7876$. $\text{adjusted}R^2=0.7592$
Endvidere er koefficienterne vil polynomiet angivet i bList.

Titel	Multipel lin...	727.544
RegEqn...	b0+b1*x1+	-107.975
F	27.8058	9.75748
PVal	0.000009	
R ²	0.787571	
adjR ²	0.759247	
s	49.711	
DW	1.81484	
dfReg	2.	
SSReg	137427.	
MSReg	68713.3	
dfError	15.	
SSError	37067.8	
MSError...	2471.18	
bList	{727.5438...	
tList	{23.03987...	
PList	{4.024491...	
SEList	{31.57758...	
yList	{727.5438...	
Resid	{12.45614...	
sResid	{0.324435...	

Vi ser nu på en model af tredje grad

Vi ser, at $\text{adjusted}R^2=0.8072$, altså lidt bedre

Vi tester nu nulhypotesen $H_0: \beta_3 = 0$

Vi finder, at P-værdi = 0.047

Da P-værdi < 5% forkastes H_0 , dvs vi kan ikke bortkaste trediegradsleddet. (vi er dog tæt på)

J	K	L	M
		stat5....	stat5.plist
	=MultReg		
Titel	Multipel...	770.702	2.40989E-12
RegEqn...	b0+b1*x...	-179.27	0.000213
F	24.7308	31.3364	0.007652
PVal	0.00000...	-1.69247	0.047175
R ²	0.84125...		
adjR ²	0.80723...		
s	44.4811		
DW	2.44637		
dfReg	3.		
SSReg	146794.		
MSReg	48931.5		
dfError	14.		
SSError	27699.9		
MSError...	1978.57		
bList	{770.70...		
tList	{22.324...		
PList	{2.4098...		
SEList	{34.522...		
yList	{770.70...		
Resid	{-30.701...		
sResid	{1.0045...		

9. Regressionsanalyse

Vi ser nu på et polynomium af fjerde grad

Vi ser, at $\text{adjusted}R^2=0.799$, altså lidt mindre

Vi tester nu nulhypotesen $H_0: \beta_3 = 0$

Vi finder, at P-værdi = 0.523

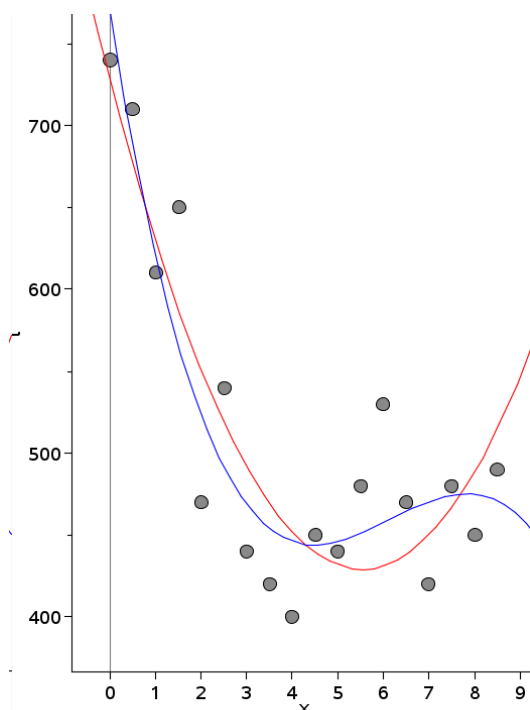
Da P-værdi > 5% accepteres H_0 , dvs. vi kan bortkaste fjerdegradsleddet.

	=MultRegT	
Titel	Multipel l...	6.78661E-11
RegEqn	b0+b1*x1...	0.060733
F	17.9002	0.778843
PVal	0.000034	0.713503
R ²	0.846337	0.523498

Heraf sluttes, at en trediegradsmodel må være det foreløbig bedste bud, men at en andengradsmodel muligvis også kunne være acceptabel

Vi må nu undersøge om der ikke er outliers, og se på residualerne for de 2 modeller

Nedenfor er tegne de to modeller , og angivet sresid for dem.



stat2.sresid	stat3.sresid
-1.09454	0.324436
0.561546	0.792451
-0.273772	-0.425559
2.05316	1.33877
-1.33781	-1.71116
1.19148	0.454375
-0.717879	-1.09667
-0.838798	-1.05366
-1.12375	-1.11313
0.136619	0.231236
-0.150579	0.179877
0.710907	1.09062
1.79419	2.10191
0.13341	0.679975
-1.25003	-0.639496
0.12748	0.29526
-0.672768	-0.890816
0.656724	-0.644302

Da der i begge tilfælde ikke er nogen værdier, der er numerisk over 3 og kun en enkelt numerisk over 2 er der ingen outliers.

Fortegnene skifter også pænt, hvilket også fremgår af tegningen, da punkterne fordeler sig jævnt på begge sider af kurverne.

Begge modeller er acceptable, men da $\text{adjusted}R^2$ er lidt højere vælges trediegradsmodellen.

Man kunde dog med lige så god ret vælge andengradsmodellen, da den jo er en enklere model

2) 2) Af b-list fås koefficienterne

$$T = -1.692 \cdot x^3 + 31.3364 \cdot x^2 - 179.266 \cdot x + 770.702$$

3) Beregn værdien af T for $x = 6.2$, og angiv et 95% konfidensinterval for T for $x = 6.2$.

TI-Nspire: Statistik ► Konfidensintervaller ► Multipl Reg Int ► Udfyld menuer herunder
x Value List={6.2, 6.2^2, 6.2^3}

TI89: F7:Ints ► 8: MultRegInt ► Udfyld menuer heraf x Value List={6.2, 6.2^2, 6.2^3}, ENTER

=MultRegInterval	
Titel	Multipel reg inte...
RegEqn...	$b_0+b_1*x_1+b_2*x_...$
\hat{y}	460.439
dfError	14.
CLower	420.015
CUpper	500.863
ME	40.4237
SE	18.8474
LowerPr...	356.826
UpperPr...	564.052
MEPred...	103.613
SEPred	48.3094
bList	{770.70175438...
Resid	{-30.701754385..

Man får $T = 460.44$ og 95% konfidensinterval [420.0;500.9]



9.9. OVERSIGT OVER FREMGANGSMÅDE VED REGRESSIONSANALYSE

9.9.1 Lineær model:

Valg af model

- Punkter + ret linie tegnes se eksempel 9.4
- 1) Punkter ligger tilfældigt om den rette linie
 - a) Se på forklaringsgrad r^2
 - a1) r^2 "høj" (over ca. 70%), så ok
 - a2) r^2 lav, så test om hældning er 0. se eksempel 9.5
 - Hvis ja, så er y uafhængig af x (ingen sammenhæng)
 - Hvis nej, så viser data en stigende / faldende tendens.
 - b) Se på Outliers. Se evt. på "Studentized Residuals" MultRegTests ► se på slist
 - b1) Ingen outliers, så ok
 - b2) Outliers : Forsøgsomstændighederne ved dette forsøg undersøges.
 - 2) Punkter ligger ikke tilfældigt om linie
 - a) Se på forklaringsgrad r^2
 - a1) r^2 meget høj (over 90%) : Model måske accepteres indenfor måleområdet.
 - a2) r^2 ikke så høj, eller en sikrere model ønskes undersøgt:
 - I: Punkter ligger "krummer monotont op eller ned":
Se på modeller med 2 parametre (eksponentiel, potens, logaritme)
Model skal opfylde krav om: Punkter ligger tilfældigt om kurve, r^2 høj, Ingen outliers
 - II: Punkter ligger ikke monotont
Se på polynomier.
Model vælges af lavest grad. Quadreg, CubicReg, beregne r^2 (adjusted), se på residualplot
evt. teste om koefficient til højstegradsled $\neq 0$

Test

- 1) Ret linie $y = \beta_0 + \beta_1 x$
 - a) H_0 : Y er uafhængig af x $\Leftrightarrow H_0$: Regressionslinien er vandret $\Leftrightarrow H_0: \beta_1 = 0$.
: Lin Reg T-Test ► Alternate Hyp= $\beta \& \rho \neq 0$
 - b) Konfidensinterval for β_1 :
LinRegTInt: ► Interval = Slope:
 - c) Konfidensinterval for y svarende til x =t:
LinRegTInt: ► Interval = Response, x Value = Slope:
 - d) Prædeterminationsinterval for 1 ny observation
Som under punkt c). Se nederst i udskrift
- 2) Andre modeller i 2 parametre:
Ligningen transformeres til et lineært udtryk se eksempel 9.6
Indsætter nye lister med de transformerede variable
Derefter som under punkt 1)
- 3) Polynomier af anden eller højere grad. $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots$
Ligningen omskrives til multipel lineær model $y = \beta_0 + \beta_1 x + \beta_2 x_2 + \dots$ sæt $x_2 = x^2$, $x_3 = x^3$ osv
Se på forklaringsgrad (bør være over 0.7)
Undersøge om der er Outliers (ingen numerisk over 3) ► MultRegTests ► se på sresid

Test

- 1) Reduktion af model:
Hvis højstegradsleddet har en største P-værdi $> 5\%$ se under P-list
Vælg MultRegTests Fjerner leddet, Gntag. Der elimineres én variabel ad gangen
- 2) Koefficienter : Findes i blist (se evt. under multipel regression)

9.9.2 Multipel regression

Model acceptabel

Se på forklaringsgrad (bør være over 0.7)

Undersøge om der er Outliers (ingen numerisk over 3) ► MultRegTests ► se på sresid .

Test

1) Reduktion af model:

Find det led hvis koefficient har den største P-værdi >5% se under P-list

Vælg MultRegTests Fjerner leddet

Gentag med næste P-værdi Der elimineres én variabel ad gangen

2) Regressionskoefficienter : Findes i blist

3) 95% konfidensinterval for β_1 : $\left[\hat{\beta}_1 - t_{1-\frac{\alpha}{2}}(f_{\text{residual}}) \cdot s_{\hat{\beta}_1} ; \hat{\beta}_1 + t_{1-\frac{\alpha}{2}}(f_{\text{residual}}) \cdot s_{\hat{\beta}_1} \right]$

I Statvar findes $s_{\hat{\beta}_1}$ i "selist" $f_{\text{residual}} = df_{\text{err}}$

4) Man finder 95% konfidensintervallet for y for given x-værdi

Ints, 8: MultRegInt

OPGAVER

Opgave 9.1

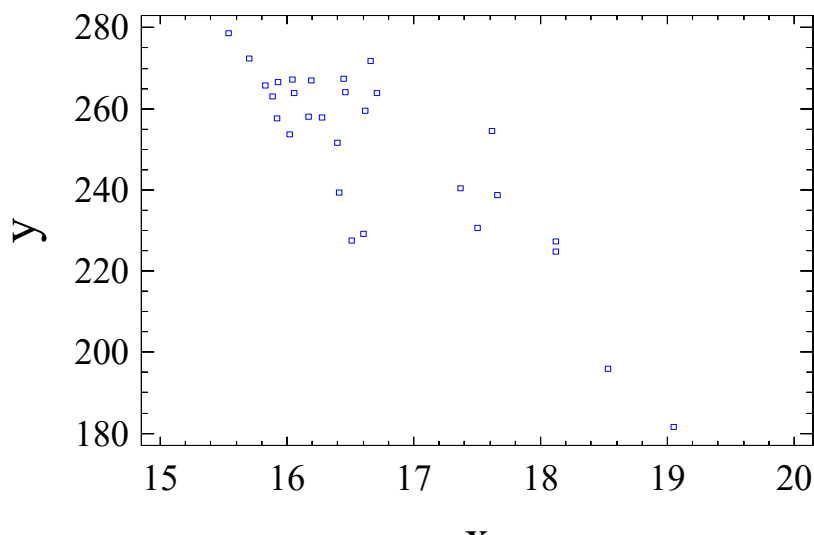
Nedenstående tabel angiver sammenhørende værdier af den "radiale" afbøjning X (i milliradianer) og den totale energiflux Y (i kilowatt) på et solvarmeanlæg.

x	16.66	16.46	17.66	17.50	16.40	16.28	16.06	15.93	16.60	16.41	16.17	15.92	16.04	16.19	16.71
y	271.8	264.0	238.8	230.7	251.6	257.9	263.9	266.5	229.1	239.3	258.0	257.6	267.3	267.0	263.8
x	16.62	17.37	18.12	18.53	15.54	15.70	16.45	17.62	18.12	19.05	16.51	16.02	15.89	15.83	
y	259.6	240.4	227.2	196.0	278.7	272.3	267.4	254.5	224.7	181.5	227.5	253.6	263.0	265.8	

Det oplyses, at korrelationskoefficienten er -0.8488

Punkterne er afsat på nedenstående figur.

Plot of y vs x



- 1) Begrund i ord på baggrund af figuren og forannævnte oplysninger, om du finder det sandsynligt, at der er uafhængighed mellem X og Y .
- 2) Find r^2 og anvend denne, ovenstående figur samt residualernes fortegn og størrelse til vurdering af om en lineær model kan antages at gælde.

I det følgende antages, at forudsætningerne for en regressionsanalyse er opfyldt, og at den lineære model gælder.

- 3) Opskriv regressionsligningen.
- 4) Angiv et 99% konfidensinterval for hældningen β_1 . (Bemærk, der ønskes et 99% interval).
- 5) Beregn et 95% konfidensinterval for middelfluxen y i det tilfælde, hvor den radiale afbøjning x er 16.5 milliradianer.

Opgave 9.2

Man ønskede på et universitet at undersøge om der var en sammenhæng mellem de point de studerende fik ved en indledende prøve i matematik, og de point de fik ved den afsluttende prøve i matematik.

Resultaterne var

Student	1	2	3	4	5	6	7	8	9	10
Indledende prøve x	39	43	21	64	57	47	28	75	34	52
Afsluttende prøve y	65	78	52	82	92	89	73	98	56	75

- 1) Find en ligning for regressionslinien m , og tegn i et koordinatsystem såvel punkterne som linien m .
- 2) Man forventer en positiv korrelation mellem x og y . Finder du at dette er tilfældet? Det antages i det følgende at forudsætningerne for en regressionsanalyse er opfyldt.
- 3) Test om y er uafhængig af x (signifikansniveau $\alpha = 0.01$)
- 4) Find et 95% konfidensinterval for hældningskoefficienten β_1 .
- 5) En student har opnået 50 point ved den indledende prøve. Forudsig indenfor hvilket interval denne student pointtal vil ligge ved den afsluttende prøve. (signifikansniveau $\alpha = 0.05$)
- 6) Angiv et 95% konfidensinterval for middelværdien af det pointtal som studenter opnår ved den afsluttende prøve, når de alle ved den indledende prøve har opnået 50 point.

Opgave 9.3

Ved en kemisk proces vides reaktionshastigheden v at afhænge af mængden x af et bestemt additiv, som virker som katalysator.

Ved et fuldstændigt randomiseret forsøg fandtes følgende observationer:

Tilsat mængde additiv x	0.001	0.2	0.5	1.0	1.5	2.0
Reaktionshastighed v	1	17	29	41	50	58

- 1) Foretag en vurdering, ud fra forklaringsgrad og graf, hvilken af de fire standardmodeller, der bedst approksimerer data.
- 2) Angiv regressionsligningen for den valgte model
- 3) Find ved en additivtilsætning på 1.75, middelværdien af reaktionshastigheden og angiv et 95% konfidensinterval herfor.

Opgave 9.4

Man mener der er en sammenhæng mellem en bilists alder og antallet af alvorlige færdselsulykker, der skyldes for stor hastighed. Man har fra USA, hvor aldersgrænsen for erhvervelse af kørekort er 16 år, følgende data indsamlet gennem en periode:

Alder x	16	17	18	19	20	22	24	27	32	42	52	57	62	72
Antal fart-relaterede ulykker y	37	32	33	34	33	31	28	26	23	16	13	10	9	7

Det fremgår klart, at antallet af ulykker falder med alderen.

- a) Giv en vurdering af, om modellen $y = a + bx$ (antal ulykker aftager lineært med alderen) på rimelig måde kan beskrive denne sammenhæng.
- b) En trafikexpert mener, at modellen $y = \alpha \cdot e^{bx}$ (antal ulykker aftager eksponentielt med alderen) giver en bedre beskrivelse af modellen. Har vedkommende ret?
- c) Bestem ligningen for den model du finder bedst.
- d) Angiv ud fra ovennævnte ligning det forventede antal fart-relaterede ulykker som 50-årige i middel vil forårsage i den givne periode.
- e) Angiv et 95% konfidensinterval for middelværdien af ulykker for 50-årige.

Opgave 9.5

I et organisk-kemisk laboratorium undersøgte man forskellige reaktionskinetiske processer. Ud fra teoretiske overvejelser har man fundet frem til, at "middeludbyttet" (angivet i %-enheder) af en bestemt kemisk forbindelse for $t > 5$ er approksimativt bestemt ved et udtryk af formen

$$y = 100 - \alpha_0 \cdot \beta_1^t, \text{ hvor } t \text{ angiver reaktionstiden og } y \text{ procesudbyttet.}$$

For at efterprøve rigtigheden af de teoretiske overvejelser udførte man et forsøg med følgende resultater:

t	6.5	8.2	11.1	13.6	16.4	18.5	20.7	23.0	25.8	28.5	33.3
y	39.5	64.7	65.6	72.9	88.0	92.7	92.5	95.9	96.3	98.3	99.2

- 1) Omskriv ovennævnte udtryk for modellen således, at regressionsmodellen kan gøres lineær i parametrene ved en logaritmisk transformation.
- 2) Foretag den logaritmiske transformation og vurder såvel grafisk som ud fra forklaringsgraden om den formodede model (1) kan accepteres.
- 3) Foretag, idet det forudsættes, at modellen (1) gælder, en estimation af parametrene α_0 og β_1 .
- 4) Opstil et 95% - konfidensinterval for middelværdien af udbyttet y svarende til $t = 20$.

Opgave 9.6

Ved en standardisering af et bestemt hormonpræparat behandler man et mindre antal mus med doser af forskellig størrelse og registrerer i hvert tilfælde tiden t , indtil musen dør. Fra tidligere undersøgelser ved man, at t er normalfordelt med konstant varians og med en middelværdi, som er en lineær funktion af logaritmen til dosis, dvs. $t = a + b \cdot \ln(\text{dosis})$.

Til brug for standardiseringen af et produktionsparti af præparatet blev foretaget 5 delforsøg, som gav følgende resultater:

dosis (antal enheder)	1585	2239	2884	5248	6918
t (timer)	8.70	6.20	8.22	2.94	3.88

- 1) Angiv et estimat for regressionslinien, hvor t er en funktion af logaritmen til dosis.
- 2) Opstil et 95% - konfidensinterval for koefficienten til logaritmen til dosis.
- 3) Opstil et 95% - konfidensinterval for middelværdien af t for en dosis på 6300 enheder.

Opgave 9.7

Man har erfaring for, at jerns viskositet Y under smeltning afhænger af jernets siliciumindhold x . Man besluttede sig til at foretage et forsøg med henblik på at undersøge denne sammenhæng nærmere. Ved forsøget foretoges 3 viskositetsmålinger for hver af 5 forskellige værdier af siliciumindholdet. Forsøgsresultaterne var:

	x				
	125	150	175	200	225
Y	47.5	60.0	65.0	72.5	77.5
	55.0	55.0	67.5	75.0	85.0
	40	50.0	70.0	77.0	75.0

- 1) Angiv forudsætningerne for at kunne udføre en regressionsanalyse. Det antages i det følgende, at forudsætningerne er opfyldt.

- 2) Test om der er en lineær sammenhæng mellem jerns viskositet og siliciumindholdet, og angiv i bekræftende fald ligningen for den empiriske regressionslinie.
Det antages i det følgende, at der er en lineær sammenhæng mellem x og y .
- 3) Foretag en testning af om regressionslinien er vandret.
- 4) Angiv et 95% konfidensinterval for hældningskoefficienten
- 5) Angiv et 95% konfidensinterval for middelværdien af viskositeten y , når $x = 160$.

Opgave 9.8

Koncentrationsbestemmelse af stoffet aprindin kan foretages ved hjælp af en gaskromatograf. Ved denne metode indsprøjtes en del af prøven indeholdende aprindin i gaskromatografen, og den såkaldte tophøjde bestemmes. Såfremt de laboratorietechniske procedurer er korrekt udført, skal tophøjden, bortset fra tilfældige udsving, være proportional med koncentrationen i prøven. I et eksperiment fremstillede man 12 prøver med kendte koncentrationer af aprindin og målte tophøjderne. Resultaterne fremgår af nedenstående tabel.

	Koncentration x ($\mu\text{g} / \text{ml}$)					
	0.5	1	2	3	4	5
Tophøjde	46	120	232	310	440	550
Y	55	90	241	318	435	539

- 1) Bestem den lineære regressionslinie for Y på x
- 2) Test, om en sådan lineær regression kan beskrive data.
- 3) Test, om tophøjden y kan antages at være proportional med koncentrationen x , dvs. $y = a \cdot x$.

Opgave 9.9

Den tid (y) det tager inden en bestemt maskinkomponent svigter kan tænkes at afhænge af den spænding (x_1), den temperatur (x_2) som komponenten udsættes for under kørslen, samt motorens omdrejningshastighed pr. minut (x_3)

Det forløbne år har givet de data, som er vist i følgende tabel:

(x_1, x_2, x_3)	(110,60,750)	(110,82,850)	(110,60,1000)	(110,82,1100)	(120,60,750)
y	2145	2155	2220	2225	2360

(x_1, x_2, x_3)	(120,82,850)	(120,60,1000)	(130,82,1100)	(115,66,840)	(115,66,880)
y	2266	2334	2340	2212	2180

Det forudsættes, at regressionsforudsætningerne er opfyldt.

- 1) Vurder ud fra forklaringsgraden, og ud fra en vurdering af om der er outliers, om en lineær model i de tre variable, dvs. af formen $Y = \alpha_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ er rimelig.
Det antages i det følgende, at ovenstående model gælder.
- 2) Undersøg om modellen kan reduceres, dvs. om nogle af koefficienterne kan antages at være 0.
- 3) Angiv regressionsligningen i den endelige model.
- 4) Bestem et estimat for Y i tilfældet $x_1 = 125$, $x_2 = 70$ og $x_3 = 900$, og angiv et 95% konfidensinterval for denne værdi.

Opgave 9.10

Ved en given produktion ønskes undersøgt, hvorledes mængden Y af et uønsket biprodukt afhang af mængderne x_1 , x_2 og x_3 af tre tilsætningsstoffer. Følgende forsøg blev foretaget (kodede tal):

(x_1, x_2, x_3)	(1,1,1)	(2,9,4)	(3,3,9)	(4,7,5)	(5,5,7)	(6,3,3)	(7,6,2)	(8,9,6)
Y	30	85	55	75	76	56	85	106

Det forudsættes, at regressionsforudsætningerne er opfyldt.

1) Vurder ud fra forklaringsgraden, og ud fra en vurdering af om der er outliers, om en lineær model i de tre variable, dvs. af formen $Y = \alpha_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ er rimelig.

Det antages i det følgende, at ovenstående model gælder.

2) Undersøg om modellen kan reduceres, dvs. om nogle af koefficienterne kan antages at være 0.

3) Angiv regressionsligningen i den endelige model.

3) Bestem et estimat for Y i tilfældet $x_1 = 4$, $x_2 = 5$ og $x_3 = 6$, og angiv et 95% konfidensinterval for denne værdi.

Opgave 9.11

Det formodes, at den producerede mængde Y af et stof ved en given produktion er en lineær funktion af de anvendte mængder x_1 , x_2 , og x_3 af tre råvarer.

Følgende ikke særligt systematiske observationer foreligger:

x_2		1			2			4			5		
x_3		1	3	5	1	3	5	1	3	5	1	3	5
x_1	1	637				6.70							
	2			8.02							7.54		
	3		8.70				9.08						
	5				10.30				10.40				
	10							15.34					16.12

1) Vurder på basis af disse observationer, om en lineær model i x_1 , x_2 og x_3 er rimelig.

2) Foretag så vidt mulig en reduktion af modellen, og angiv til sidst regressionsligningen for den endelige model.

3) Beregn et 95% konfidensinterval for regressionskoefficienterne i den endelige model.

4) Beregn et 95% konfidensinterval for middelværdien af Y hvis $x_1 = 0.3$, $x_2 = 0.4$ og $x_3 = 0.1$.

Opgave 9.12

En fabrik fremstiller salpetersyre ved oxidering af ammoniak med luft. I løbet af processen ledes kvælstofoxider under afkøling ind i en absorptionskolonne, idet absorptionen i gennemstrømmende salpetersyre afhænger af kølevandstemperaturen x_1 ($^{\circ}\text{C}$), lufttemperaturen x_2 ($^{\circ}\text{C}$) og salpetersyrekoncentrationen x_3 .

Man ønsker at teste, om sammenhængen mellem mængden Y af ikke-absorberede kvælstofoxider i et givet tidsrum og x_1 , x_2 og x_3 (aproksimativt) var lineær, og ønskede i bekræftende fald at estimere denne sammenhæng.

Følgende observationer af Y (kodede tal) fandtes:

			x_2				
			10	20	30	40	
x_3	-5	x_1	5	32	50	62	88
			10	48	60	73	96
			15	54	70	90	102
	5	x_1	5	32	49	64	86
			10	50	64	82	92
			15	57	74	88	108

- 1) Vurder på basis af disse observationer, om en lineær model i x_1 , x_2 og x_3 er rimelig.
- 2) Undersøg, om modellen kan reduceres, dvs. om nogle af regressionskoefficienterne kunne være 0.
- 3) Opskriv regressionsligningen.
- 4) Angiv et 95% konfidensinterval for β_1 .
- 5) Angiv et estimat for Y i tilfældet $x_1 = 8$, $x_2 = 20$ og $x_3 = 4$, og angiv et 95% konfidensinterval for denne værdi.

Opgave 9.13

Følgende sammenhørende data er 25 målinger mellem den jævnstrøm (y) en vindmølle udvikler og vindhastigheden (x).

x	5.00	6.00	3.40	2.70	10.0	9.70	9.55	3.05	8.15	6.20	2.90	6.35	4.60
y	1.582	1.822	1.057	0.500	2.236	2.386	2.294	0.588	2.166	1.866	0.653	1.930	1.562
x	5.80	7.40	3.60	7.85	8.80	7.00	5.45	9.10	10.20	4.10	3.95	2.45	
y	1.737	2.088	1.137	2.179	2.112	1.800	1.501	2.303	2.310	1.194	1.144	0.123	

- 1) Vurder på basis af ovennævnte observationer hvilket polynomium $y = \alpha_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \dots + \beta_p x^p$ af lavest mulig grad p , der indenfor måleområdet $[0; 10]$ giver en tilfredsstillende beskrivelse af Y 's variation.
- 2) Angiv regressionsligningen for den model, man i spørgsmål 1 har fundet frem til.
- 3) Skitser grafen indenfor intervallet $[0; 10]$, og beregn værdien af y svarende til en vindhastighed på $x = 7$, og et 95% konfidensinterval for $x = 7$

Opgave 9.14

Ved et fuldstændigt randomiseret forsøg foretoges følgende observationer mellem den ikke-statistiske variabel x og den statistiske variabel Y :

x	10	20	30	40	50	60	70	80
Y	72.5	78.3	79.6	78.9	76.9	76.6	68.9	66.4

- 1) Bestem ved en polynomial regressionsanalyse det polynomium i x af lavest mulig grad, der giver en tilfredsstillende beskrivelse af Y 's variation.
Betragt såvel forklaringsgrader som residualer
- 2) Opskriv regressionsligningen for den i spørgsmål 1 fundne model
- 3) Find middelværdien af Y , når $x = 45$.
- 4) Find den værdi x_m som giver den største y - værdi . Angiv endvidere den til x_m svarende estimerede middelværdi \hat{Y}_m

Opgave 9.15

Ved nogle forsøg med målinger af det tryk, som udgår fra jetmotorer, målt for udvalgte værdier af ændringen i udstødningsdysens vinkel x værdier af ændringen i trykket Y . Resultaterne var:

x	4	4.5	5	6.5	7	7.3	7.5
$Y(i\%)$	24.5	26.3	27.2	67.3	77.8	80.6	83.4

1. Bestem ved en polynomial regressionsanalyse det polynomium af lavest grad, der giver en tilfredsstillende beskrivelse af Y 's variation.
2. Benyt den fundne model til at bestem middelværdien af Y , når $x = 6$ og beregn det dertil svarende 95% konfidensinterval .

10. RANGTEST (Fordeling ukendt)

10.1 INDLEDNING

De testprocedurer vi har benyttet i kapitlerne 3,4 og 5 har alle været baseret på, at i det mindste approksimativt kendte fordelingen (normal-, binomial- eller Poisson-fordelt) og testen vedrørte parametre i fordelingen såsom μ , σ eller p . Denne form for statistik kunne kaldes "parametrisk statistik". Kendes fordelingen ikke, og kan man heller ikke approksimere den til en kendt fordeling, så må man benytte de såkaldte "ikke-parametriske test". Disse forudsætter ikke, at fordelingen er kendt, og kunne derfor også kaldes "fordelingsfri test".

Da det ligger udenfor denne bogs centrale emner, er beskrivelsen i det følgende kun oversigtsmæssigt og skrevet med petit.

Rangtest


Disse test har som forudsætning at observationerne afspejler mindst en ordning, samt at de foreliggende observationer er baseret på (eventuelt underliggende) kontinuerte statistiske variable, jævnfør det følgende eksempel 10.1

Eksempel 10.1. Smagsprøveeksperiment.

En levnedsmiddelproducent overvejede at introducere et nyt konservesprodukt i stedet for et hidtil produceret. I disse overvejelser indgik resultatet af en smagsprøvning foretaget af et panel bestående af 15 smagsdommere. De smagte hver på det nye produkt og gav dette en "karakter" ved afsættelse af et kryds på et standardliniestykke, hvis ene endepunkt svarer til værst mulig smag, det andet endepunkt til bedst mulig smag.

Med henblik på den statistiske analyse af resultaterne transformerades disse til tal, idet hvert af de 15 standardliniestykker inddeltes lineært efter en skala fra 0 (værst mulig smag) til 100 (bedst mulig smag), og krydsenes placering aflæstes.

De transformerede forsøgsresultater var: 75, 90, 66, 82, 75, 88, 55, 80, 83, 75, 70, 80, 68, 86, 84.

Ved et stort antal tilsvarende smagsprøvninger af det hidtil producerede konservesprodukt var medianen 70. Producenten ønsker at få at vide, om forsøgsresultaterne giver et eksperimentelt bevis for, at det nye produkt smager bedre end det hidtidige. 

Tallene i eksempel 10.1 er ikke sædvanlige måletal. Betragtes eksempel vis de 3 første observationer 75,90 og 66, så kan vi kun slutte, at smagsdommer 2 synes produktet smager bedre end nr. 1 som igen synes det smager bedre end nr. 3. Differensen mellem de to første er 15, mellem nr. 1 og 3 er 9. Vi kan derfor muligvis også slutte, at der vil være større smagsforskel i det første tilfælde end i det sidste. Derimod kan vi næppe slutte, at smagsforskellen i det første tilfælde er $\frac{15-9}{15} \cdot 100 = 40\%$ større end i det sidste. Det har ingen fysisk realitet at foretage procentiske

sammenligninger mellem smagsforskelle. Sådanne sammenligninger ville derimod have mening, hvis der forelå sædvanlige måletal, jævnfør eksempel 1.1, hvor, hvor procentiske forskelle mellem hårheden umiddelbart kan fortolkes.

En forsigtig fortolkning af de transformerede resultater af smagsprøvningen er derfor, at de kun afspejler en ordning med hensyn til smag.

Da man afsatte punkter på en standardliniestykke hvorved i princippet ethvert tal mellem 0 og 100 kunne tænkes, er kravet om kontinuitet opfyldt.

Forudsætningerne for at kunne udføre en rangtest er dermed opfyldt.

En "diskret" skala må altså i princippet ikke anvendes. Man kunne dog godt have anvendt en skala fra 0 til 100 da den ville være tilstrækkelig "finmasket" i forhold til antallet af smagsdommere. En skala fra 0 til 3 ville derimod ikke kunne bruges (der ville blive alt for mange ens "målinger").

Man ser undertiden nogle som "for en sikkerheds skyld" anvender en Rangtest, selv om de faktisk med rimelighed kunne have udført en parametrisk test. Hvis man derved får en forkastelse af nulhypotesen er konklusionen korrekt. Hvis derimod man får en accept, kunne det være at en parametrisk test ville give en forkastelse (da de parametriske test har størst styrke). Derfor skal parametriske test naturligvis foretrækkes, hvis man som tidligere beskrevet har en rimelig formodning om fordelingen. Har man således en stor stikprøve, så kan man i følge den centrale grænseværdisætning være sikker på, at gennemsnittet \bar{x} er normalfordelt, selv om fordelingen ikke er det.

De forskellige typer rangtest blive forklaret i forbindelse med passende eksempler.

I den forbindelse får vi brug for følgende oplysninger:

Definition af rangtal: Lad os antage vi har givet n forskellige tal.
 Det mindste af disse tal gives rangtallet 1, det næstmindste tal gives rangtallet 2 osv.

Eksempel:

Observationer	6	17	7	13	5	2
Tilhørende Rangtal	3	6	4	5	2	1

Sum af Rangtal: Sum af tal fra 1 til n kan vises at være $S = 1 + 2 + 3 + \dots + n = \frac{n \cdot (n+1)}{2}$.

Bevis: $S = 1 + 2 + 3 + \dots + (n-2) + (n-1) + n$

$$S = n + (n-1) + (n-2) + (n-3) \dots + 3 + 2 + 1$$

$$\text{Heraf fås } 2S = (n+1) + (n+1) + (n+1) + \dots + (n+1) = n \cdot (n+1)$$

Eksempel $S = 1 + 2 + 3 + \dots + 6 = \frac{6 \cdot (6+1)}{2} = 15$

Median : Observationerne ordnes i rækkefølge:
 Ulige antal observationer: median = midterste tal,
 Lige antal observationer: median = gennemsnit af de to midterste tal

Eksempel: Observationer 6, 17, 7, 13, 5, 2.

Ordnet i rækkefølge: 2, 5, 6, 7, 13, 17. Median 6,5

10.2. WILCOXONS RANGTEST FOR 1 STIKPRØVE

Eksempel 10.2 (fortsættelse). Wilcoxon's rangtest for 1 variabel

Det hidtidige produkt havde en median på 70. Producenten ønsker at undersøge på et signifikansniveau på 5% om det nye produkt smager bedre end det gamle.

Løsning:

Lad m betegne fordelings median.

Nulhypotese: $H_0: m = 70$ Alternativ hypotese $H: m > 70$

Differenserne i forhold til medianen på 70 beregnes.

x_i	75	90	66	82	75	88	55	80	83	75	70	80	68	86	84
$x_i - 70$	5	20	-4	12	5	18	-15	10	13	5	0	10	-2	16	14

Den numeriske værdi af differenserne tilordnes rangtal, således at den mindste værdi får rangtallet 1, den næstmindste rangtallet 2, osv. Hvis flere værdier er ens tilordnes de alle gennemsnittet af de rangtal de skulle have haft. Er en differens 0 tilordnes intet rangtal.

Rangtallene forsynes derefter med fortegn for de tilsvarende differenser.

Dette giver følgende tabel

x_i	75	90	66	82	75	88	55	80	83	75	70	80	68	86	84
$x_i - 70$	5	20	-4	12	5	18	-15	10	13	5	0	10	-2	16	14
Rang af $ x_i - 70 $	4	14	2	8	4	13	11	6,5	9	4	-	6,5	1	12	10
Fortegn			-2				-11						-1		

Lad W_+ være summen af rangtallene med positivt fortegn og W_- summen af rangtallene med negativt fortegn. Da

vi udkød en værdi, her vi, at summen af rangtallene er $S = \frac{14 \cdot 15}{2} = 105$.

Da $w_- = 2+11+1=14$ følger heraf, at $w_+ = 105 - 14 = 91$.

Hvis nulhypotesen er sand ville vi forvente at der lå nogenlunde lige mange tal over og under medianen på 70, og at summen af rangtallene for de med positivt fortegn ville være nogenlunde lig med summen af dem med negativt fortegn dvs. omkring $\frac{105}{2} \approx 52,5$.

Man kan vise, at hvis antallet n af tal i stikprøven er større end 10, så vil fordelingen af summen af rangtallene med negativ fortegn W_- og dem med positivt fortegn W_+ begge være approksimativt normalfordelt med middelværdi $\mu = \frac{n \cdot (n+1)}{4}$ og spredning $\sigma = \sqrt{\frac{n \cdot (n+1) \cdot (2n+1)}{24}}$.

Metode 1: Da antallet i stikprøven er større end 10 kan man approksimere med en normalfordeling med middelværdi $\mu = \frac{n \cdot (n+1)}{4} = \frac{14 \cdot 15}{4} = 52,5$ og spredning $\sigma = \sqrt{\frac{n \cdot (n+1) \cdot (2n+1)}{24}} = \sqrt{\frac{14 \cdot 15 \cdot 29}{24}} = 15,93$.

Vi har P -værdi $= P(w_- \leq 14) = \text{normCdf}(-\infty, 14, 52,5, 15,93) = 0,008$,

(eller P -værdi $= P(w_+ \geq 91) = \text{normCdf}(91, \infty, 52,5, 15,93) = 0,008$)

Da P -værdi $< 0,01$ forkastes H_0 (2 stjernet), dvs. vi har vist, at det nye produkt smager bedre end det gamle.

Metode 2: Hvis $n \leq 10$ er det nødvendigt at benytte nedenstående tabel. Benyttes denne i dette tilfælde findes, at for $n = 14$ er $w_{0,05} = 25$ og $w_{0,01} = 15$ (da testen er ensidet)

Idet $w = \min\{w_-, w_+\} = \min\{14, 91\} = 14$ ses, at $w < w_{0,01}$.

H_0 forkastes (2 stjernet), dvs. vi har igen vist, at det nye produkt smager bedre end det gamle. ◆

TABEL 10.1. Rang test vedrørende 1 statistisk variabel (Wilcoxon's teststørrelse).

Tabel over kritiske værdier af Wilcoxon's teststørrelse W_α

$\alpha \backslash n$	0.10	0.05	0.02	0.01	to-sidet test
	0.05	0.025	0.01	0.005	én-sidet test
5	0				
6	2	0			
7	3	2	0		
8	5	3	1	0	
9	8	5	3	1	
10	10	8	5	3	
11	13	10	7	5	
12	17	13	9	7	
13	21	17	12	9	
14	25	21	15	12	
15	30	25	19	15	
16	35	29	23	19	
17	41	34	27	23	
18	47	40	32	27	
19	53	46	37	32	
20	60	52	43	37	
21	67	58	49	42	
22	75	65	55	48	
23	83	73	62	54	
24	91	81	69	61	
25	100	89	76	68	

Idet $w = \min\{w_+, w_-\}$ forkastes nulhypotesen H_0 hvis $w < w_\alpha$ (jævnfør eksempel 15.2).

10.3 WILCOXONS RANGTEST FOR 2 UAFHÆNGIGE STIKPRØVER

Hvis de i de første afsnit angivne metoder (t -test m.m.) ikke kan anvendes, fordi forudsætningerne ikke er opfyldte, så kan man sædvanligvis anvende denne metode.

Nulhypotesen er, at de to stikprøver er taget fra den samme population med samme medianer. Hvis vi nu danner en rækkefølge for de samlede observationer under et, så burde rangtallene fra de to stikprøver fordele sig tilfældigt i forhold til hinanden. Summen for hver stikprøve burde derfor være nogenlunde den samme. Hvis de afviger meget fra hinanden må nulhypotesen kunne forkastes.

Lad antallet i den ene stikprøve være n_1 og summen af rangtallene være w_1 .

Lad tilsvarende antallet i den anden stikprøve være n_2 og summen af rangtallene være w_2 .

Er $n_1 > 5$, $n_2 > 5$ og $n_1 \leq n_2$ gælder, at fordelingen af summerne W_1 og W_2 er approksimativt normalfordelt med middelværdi $\mu = \frac{n_1 \cdot (n_1 + n_2 + 1)}{2}$ og spredning $\sigma = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}$

Metoden demonstreres i følgende eksempel 10.3.

Eksempel 10.3 . Wilcoxon's rangtest for 2 uafhængige stikprøver

Man ønsker at sammenligne reaktionstiden for mænd, som har taget én type medicin "A", med reaktionstiden for de mænd, der har taget en anden type medicin "B". Forsøg har imidlertid vist, at disse reaktionstider ikke er normalfordelte, men fordelingen er skæv mod højre.

Man kan derfor ikke benytte en t -test, men måtte benytte en rangtest.

Blandt 14 mænd udvalgte man tilfældigt 7 som indtog medicin A mens de øvrige 7 indtog medicin B. Efter en passende tid målt reaktionstiden for de 14 personer. En person måtte desværre udgå på grund af specielle omstændigheder.

Resultaterne var:

Reaktionstid	Medicin A	1.96	2.24	1.71	2.41	1.62	1.93	
	Medicin B	2.11	2.43	2.07	2.71	2.50	2.84	2.88

Undersøg om der er signifikant forskel på reaktionstiden.

Løsning:

Lad m_A og m_B betegne de to fordelingers medianer.

Nulhypotese: $H_0: m_A = m_B$ Alternativ hypotese $H: m_A \neq m_B$

Vi tildeler de samlede observationer rangtal.

Reaktionstid	Medicin A	1.96	2.24	1.71	2.41	1.62	1.93	
	Medicin B	2.11	2.43	2.07	2.71	2.50	2.84	2.88
Rangtal	A	4	7	2	8	1	3	
	B	6	9	5	11	10	12	13

Vi beregner nu summen w_A og w_B af rangtallene.

$$w_A = 4 + 7 + 2 + 8 + 1 + 3 = 25$$

$$\text{Idet } w_A + w_B = \frac{(n_A + n_B)(n_A + n_B + 1)}{2} = \frac{(7 + 6)(7 + 6 + 1)}{2} = 91 \text{ er } w_B = 91 - 25 = 66$$

Da summen af w_A og w_B er et fast tal, vil en lille værdi af den ene delsum betyde en stor værdi af den anden, og dermed en stor forskel på de to summer. Der gælder derfor, at jo mindre en af delsummerne er, jo større er muligheden for at stikprøverne er udtaget af forskellige fordelinger.

Tabelmetode

I tabel 10.2 er for forskellige værdier af antal observationer n_1 og n_2 angivet et interval.

I vort tilfælde er $n_A = 6 < n_B = 7$, og teststørrelsen altså $w_A = 25$. Ved opslag i tabellen under den tosidet test for signifikansniveau 5% fås, at intervalgrænserne er 28 og 56. Da $w_A \leq 28$ forkastes nulhypotesen, dvs. der er signifikant forskel på reaktionstiderne.

Aproksimation med normalfordeling

Da de to stikprøver er større end 5 er det tilladeligt at approksimere med en normalfordeling med middelværdi

$$\mu = \frac{n_1 \cdot (n_1 + n_2 + 1)}{2} = \frac{6 \cdot (6 + 7 + 1)}{2} = 42 \text{ og spredning } \sigma = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}} = \sqrt{\frac{6 \cdot 7 \cdot (6 + 7 + 1)}{12}} = 7$$

Vi har P -værdi = $P(W_A \leq 25) = \text{normCdf}(-\infty, 25, 42, 7) = 0.0076$.

Da P -værdi < 0.025 forkastes H_0 , dvs. vi har vist, at der er en signifikant forskel på reaktionstiderne. ◆

TABEL 10.2. Rang test vedrørende 2 uafhængige statistiske variable (Wilcoxon's teststørrelse).

I: Tosidet test: $\alpha = 0.05$ Ensided test: $\alpha = 0.025$

$n_2 \backslash n_1$	3		4		5		6		7		8		9		10	
	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ
3	5	16	6	18	6	21	7	23	7	26	8	28	8	31	9	33
4	6	18	11	25	12	28	12	32	13	35	14	38	15	41	16	44
5	6	21	12	28	18	37	19	41	20	45	21	49	22	53	24	56
6	7	23	12	32	19	41	26	52	28	56	29	61	31	65	32	70
7	7	26	13	35	20	45	28	56	37	68	39	73	41	78	43	83
8	8	28	14	38	21	49	29	61	39	73	49	87	51	93	54	98
9	8	31	15	41	22	53	31	65	41	78	51	93	63	108	66	114
10	9	33	16	44	24	56	32	70	43	83	54	98	66	114	79	131

II: Tosidet test: $\alpha = 0.10$ Ensided test: $\alpha = 0.05$

$n_2 \backslash n_1$	3		4		5		6		7		8		9		10	
	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ	W_n	W_ϕ
3	6	15	7	17	7	20	8	22	9	24	9	27	10	29	11	31
4	7	17	12	24	13	27	14	30	15	33	16	36	17	39	18	42
5	7	20	13	27	19	36	20	40	22	43	24	46	25	50	26	54
6	8	22	14	30	20	40	28	50	30	54	32	58	33	63	35	67
7	9	24	15	33	22	43	30	54	39	66	41	71	43	76	46	80
8	9	27	16	36	24	46	32	58	41	71	52	84	54	90	57	95
9	10	29	17	39	25	50	33	63	43	76	54	90	66	105	69	111
10	11	31	18	42	26	54	35	67	46	80	57	95	69	111	83	127

$n_1 \leq n_2$: H_0 forkastes hvis $w_{n_1} < w_n$ eller $w_{n_1} > w_\phi$ (ligger udenfor intervallet $[w_n ; w_\phi]$)

Hvis $n_1 \geq 10$ og $n_2 \geq 10$ er teststørrelsen approksimativt normalfordelt med middelværdi $\mu = \frac{n_1 \cdot (n_1 + n_2 + 1)}{2}$ og

spredning $\sigma = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}$

Da det kan være ret besværligt at rangordne tallene, kan man få TI 89 til at gøre det på følgende måde:

APPS ► STAT/LIST, slet indholdet i alle gamle lister.

► Indtast i list 1 først de 6 tal fra medicin A og derefter de 7 tal fra medicin B

► Indtast i list 2 tallet 1 6 gange og tallet 2 7 gange (se skemaet nedenfor)

list1	list2
1.96	1
2.24	1
1.71	1
2.41	1
1.62	1
1.93	1
2.11	2
2.43	2
2.07	2
2.71	2
2.50	2
2.84	2
2.88	2

Cursor i list 1 ► F3 ► 2 -Ops ► 2: Sort list, Adjust all ► Enter ► ENTER

Man ser nu at LIST 1 er sorteret i stigende rækkefølge, og man kan nu udfylde skemaet.

10.4. KRUSKAL WALLIS TEST . Erstatte ensidet variansanalyse

Er forudsætningerne for en ensidet variansanalyse ikke opfyldte, så kan vi benytte nedenstående rangtest (kaldet Kruskal-Wallis test), der som tidligere nævnt kun kræver at observationerne mindst afspejler en ordening, og at den bagved liggende fordeling er kontinuert.

Eksempel 10.5 . Kruskal-Wallis test for mere end 2 variable.

Et levnedsmiddels smag kan tænkes at afhænge af hvilken af 3 produktionsmetoder der anvendes. For at undersøge om det er tilfældet planlægges følgende forsøg:

Med hver af de 3 metoder fremstilles i en forsøgsproduktion 6 prøver.

En eksperternager vurderer de i alt 18 smagsprøver enkeltvis og i tilfældig rækkefølge uden kendskab til, hvilken metode der er anvendt i det enkelte tilfælde. Efter hver smagning markeres resultatet ved afsætning af et kryds på et standardliniestykke, hvis ene endepunkt svarer til værst mulig smag, det andet endepunkt til bedst mulig smag. Ved den statistiske analyse af resultaterne transformeres disse til tal, idet hvert af de 18 standardliniestykker indeles lineært efter en skala fra 0 (værst mulig smag) til 100 (bedst mulig smag). De transformerede resultater er de tal, som angiver krydsernes placering, og kan betragtes som stikprøveværdier af $q (=3)$ kontinuerte statistisk uafhængige variable med ukendte fordelingstyper.

De transformerede forsøgsresultater blev:

Metode M_1	61	69	79	61	59	
Metode M_2	62	58	47	59	63	48
Metode M_3	57	45	60	54	57	

Det bemærkes, at der ved forsøget kun fremkom 5 observationer for metoderne M_1 og M_3 på grund af tekniske fejl ved fremstillingen af 2 prøver.

Idet m_1 , m_2 og m_3 betegner de 3 fordelingers medianer, ønsker vi på grundlag af stikprøveværdierne at teste nulhypotesen

H_0 : De 3 fordelinger er ens (hvilket indebærer, at $m_1 = m_2 = m_3$) imod den alternative hypotese H : De 3 fordelinger er ikke ens.

Løsning: TI89

Observationerne tilordnes rangtal som om de kom fra samme population (som ved Wilcoxon's test), Da det kan være ret besværligt at rangorden tallene, kan man få TI 89 til at gøre det på følgende måde:

APPS ► STAT/LIST , slet indholdet i alle gamle lister.

► Indtast i list 1 først de 5 tal fra metode M1 og derefter de 6 tal fra metode M2 og tilsidst de 5 tal fra metode M3

► Indtast i list 2 tallet 1 5 gange og tallet 2 6 gange og tallet 3 5 gange (se skemaet nedenfor)

list1	list2
61	1
69	1
79	1
61	1
59	1
62	2
58	2
osv.	-
57	3

Cursor i list 1 ► F3 ► 2 -Ops ► 2: Sort list, Adjust all ► Enter ► ENTER

Man ser nu at LIST 1 er sorteret i stigende rækkefølge, og man kan nu udfylde skemaet.

Metode M ₁	61	69	79	61	59		Sum af rangtal R_i
Rangtal	11.5	15	16	11.5	8.5		62.50
Metode M ₂	62	58	47	59	63	48	
	13	7	2	8.5	14	3	47.5
Metode M ₃	57	45	60	54	57		
	5.5	1	10	4	5.5		26.0

Da alle stikprøvestørrelser er større end eller lig 5 gælder det , at teststørrelsen

$$\chi^2 = \frac{12 \cdot \sum_{i=1}^3 n_i (\bar{R}_i - \bar{R})^2}{n(n+1)} \text{ er } \chi^2 \text{-fordelt med frihedsgradstal } f = 3 - 1 = 2$$

Som det ses af formlen vil χ^2 være tæt ved nul hvis nulhypotesen er sand, mens den får en stor værdi hvis de gennemsnitlige rangtal afviger meget fra hinanden, så testen er ensidet.

Beregningsteknisk er det lettere at anvende den i appendix 10.1 angivne omskrivning:

$$\chi^2 = \frac{12 \cdot \sum_{i=1}^k \frac{R_i^2}{n_i}}{n \cdot (n+1)} - 3 \cdot (n+1) = \frac{12 \cdot \left(\frac{62.5^2}{5} + \frac{47.5^2}{6} + \frac{26.0^2}{5} \right)}{16 \cdot (16+1)} - 3 \cdot (16+1) = 6.02$$

Da P -værdi = $P(\chi^2 > 6.02) = \text{chi2cdf}(6.02, \infty, 2) = 0.0493 < 0.05$ forkastes nulhypotesen (tæt ved accept), og vi må derfor konkludere, at de 3 fordelinger ikke er identiske.

Ud fra de fundne summer må man kunne slutte, at metode 1 giver en bedre smag end metode 3.

OPGAVER

Opgave 10.1

Et mellemprodukt ved en kemisk proces bliver med mellemrum undersøgt for urenheder. Urenhedsniveauet (i ppm) blev ved sidste måling målt til

2.4	2.5	1.7	2.1	2.2	2.6	1.3	1.9	2.0	2.5	2.6	2.3	2.0	1.8	1.3	1.7	2.0	1.9	2.3	1.9	2.4	1.6
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Kan du ved en rangtest vise, at urenhedsniveauet er mindre end 2.5 ppm ($\alpha = 0.05$).

Opgave 10.2

Ved en psykologisk undersøgelse ønskede man at undersøge om den førstefødte af to tvillinger har en tendens til at blive mere aggressiv end den sidstfødte. Man gav 12 tvillingepar den samme psykologiske test. Resultaterne af testen ses nedenfor. Et større pointtal viser tegn på en større aggressivitet.

Tvillingepar nr	1	2	3	4	5	6	7	8	9	10	11	12
førstfødte	86	71	77	68	91	72	77	91	70	71	88	87
sidstfødte	88	77	76	64	96	72	65	90	65	80	81	72

- 1) Hvorfor er en rangtest rimelig at anvende i dette tilfælde.
- 2) Foretag en testning af om man på grundlag heraf konkludere at den førstfødte er mere aggressiv end den sidstfødte.

Opgave 10.3

Ved en udgravning på Cypern fandt man bl.a. Byzantinske mønter fra to perioder af Manuel I's regering (1143-1180). Der var 9 mønter fra den 1. udmøntning og 7 mønter fra den 4. udmøntning. Mønterne blev analyserede for indholdet af sølv, og man fandt følgende resultater (i %):

1. udmøntning	59	68	64	70	66	77	72	69	62
2. udmøntning	53	56	55	51	62	58	58		

Test, om der er forskel på sølvindholdet for de to udmøntninger

- 1) ved en rangtest, idet man er usikker på om forudsætningerne for en parametriske test er opfyldt.
- 2) ved et parametriske test.

Opgave 10.4

Et støberifirma har købt 2 nye maskiner som producerer stempler af en anden type end de hidtidige maskiner. Efter nogen tids produktion begyndte der at komme klager over, at de nye stempler ikke har samme holdbarhed som de hidtidige. For at teste dette udtager man tilfældigt stempler fra såvel den nye som den gamle type og måler på en testmaskine, hvor længe det varer inden de går i stykker.

Man fandt (i timer):

Gamle stempler	947	1358	821	1032		
Nye stempler	765	834	699	957	1021	756

Undersøg ved en ikke-parametriske test på et signifikansniveau på 5% om de gamle stempler holder længere end de nye stempler.

Opgave 10.5

Ved 2 forskellige metoder til syntetisering af et stof fandtes følgende udbytter (%)

Metode 1	578	562	619	544	536	564	532
Metode 2	642	587	631	625	598	592	571

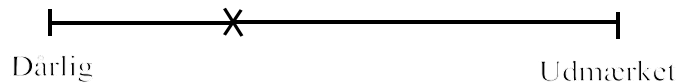
Test, om middeludbyttet er forskelligt ved de 2 metoder.

Da man ikke er sikker på, at forudsætningerne for en parametriske test er opfyldt, så vælges at udføre en rangtest.

Opgave 10.6

Med henblik på markedsføring af dansk Feta-ost i et mellemøstligt land gennemførtes en prøvesmagningsundersøgelse i det pågældende land. Som et led i undersøgelsen medvirkede 32 smagsdommere ved en bedømmelse af 4 ostetyper, hvoraf 2 var bestemte ostefabrikater fra det pågældende land (Feta-typerne 1 og 2), medens de 2 andre var danske, henholdsvis UF-Feta-ost og traditionel dansk Feta-ost. De 32 dommere inddeltes tilfældigt i 4 grupper a 8. Hver af dommerne i samme gruppe fik en prøve af samme ostetype, jfr. nedenstående skema, idet osteprøverne præsenteredes, så smagerne ikke havde mulighed for at identificere de enkelte fabrikater og heller ikke fik oplyst deres oprindelsesland.

Ved hver prøvesmagning markeredes vurderingen af osteprøven ved afsætning af et kryds på et standardliniestykke



hvorefter krydset på sædvanlig måde "oversattes" til en 0 - 100 - skala. Herved fremkom følgende forsøgsresultater:

Udenlandsk Feta-type 1 (gruppe 1)	45	68	33	61	48	22	60	48
Udenlandsk Feta-type 2 (gruppe 2)	86	95	59	75	69	80	86	59
Dansk UF-Feta (gruppe 3)	81	54	69	63	65	88	61	87
Traditionel dansk Feta (gruppe 4)	35	52	70	45	42	60	43	48

Foretag en statistisk analyse af forsøgsresultaterne.

Opgave 10.7

En å er blevet kraftigt forurenet. For at få et indtryk af, om forureningen påvirker de tre vigtigste fiskearter i samme grad, fangedes nogle fisk, som anbragtes i passende bassiner. En ekspert iagttog fiskene og gav en vurdering af deres helbredstilstand ved for hver fisk at afsætte et kryds på et standardliniestykke, hvis ene endepunkt svarer til meget sygeligt udseende og unormal opførsel, og hvis andet endepunkt svarer til rask udseende og normal opførsel. Derefter opdeltes hvert standardliniestykke efter en skala fra 0 til 25. Idet de tre arter kaldes A, B og C fremkom følgende tabel:

A	13	11	16	13	2	7	11	13		
B	24	15	11	19	14	16	18	19	15	7
C	22	14	20	11	17	10	15	11	16	21

Test på grundlag af ovenstående, om det kan antages, at de tre fiskearter lider lige meget under forureningen.

APPENDIX

Appendix 2.1. Begrundelse for grænserne for kontrolkort

Vi danner k undergrupper hver med n observationer.

Gruppe	Observationer	Gennemsnit	Variationsbredde	Spredning
1	$x_{11}, x_{12}, \dots, x_{1n}$	\bar{x}_1	R_1	s_1
2	$x_{21}, x_{22}, \dots, x_{2n}$	\bar{x}_2	R_2	s_2
.
.
k	$x_{k1}, x_{k2}, \dots, x_{kn}$	\bar{x}_k	R_k	s_k
Total		$\bar{\bar{x}}$	\bar{R}	\bar{s}

Grænser for R - kort. Det kan bevises, at $E(R) = d_2 \sigma$ og $\sigma(R) = d_3 \cdot \sigma$, hvor d_2 og d_3 kun afhænger af stikprøvestørrelsen n . Disse konstanter samt alle de i det følgende nævnte konstanter kan aflæses i tabel 2.2.

Et R - kort med 3σ - grænser er derfor bestemt ved :

$$NKG_R = (d_2 - 3 \cdot d_3) \cdot \sigma \text{ og } \emptyset KG_R = (d_2 + 3 \cdot d_3) \cdot \sigma.$$

Idet et estimat for $E(R) \approx \bar{R}$, fås $\sigma \approx \frac{\bar{R}}{d_2}$.

Indsættes dette, fås $NKG_R = (d_2 - 3 \cdot d_3) \cdot \frac{\bar{R}}{d_2} = D_3 \cdot \bar{R}$ og $\emptyset KG_R = (d_2 + 3 \cdot d_3) \cdot \frac{\bar{R}}{d_2} = D_4 \bar{R}$.

Grænser for s - kort. Selv om der gælder, at $E(s^2) = \sigma^2$, gælder det ikke, at $E(s) = \sigma$.

Man kan imidlertid vise, at $E(s) = c_4 \cdot \sigma$ og $\sigma(s) = \sqrt{1 - c_4^2} \cdot \sigma$.

Følgelig er et s - kort med 3σ - grænser bestemt ved :

$$NKG_s = (c_4 - 3 \cdot \sqrt{1 - c_4^2}) \cdot \sigma = B_5 \cdot \sigma \text{ og } \emptyset KG_s = (c_4 + 3 \cdot \sqrt{1 - c_4^2}) \cdot \sigma = B_6 \cdot \sigma.$$

Idet et estimat for $E(s) \approx \bar{s}$, fås $\sigma \approx \frac{\bar{s}}{c_4}$.

Indsættes dette, fås $NKG_s = \left(1 - \frac{3}{c_4} \sqrt{1 - c_4^2}\right) \cdot \bar{s} = B_3 \cdot \bar{s}$ og $\emptyset KG_s = \left(1 + \frac{3}{c_4} \sqrt{1 - c_4^2}\right) \cdot \bar{s} = B_5 \cdot \bar{s}$.

Grænser for \bar{x} - kort. Idet $E(\bar{x}) = \mu$ og $\sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}}$, er et \bar{x} - kort med 3σ - grænser:

$$NKG_{\bar{x}} = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}} = \mu - A_1 \cdot \sigma \text{ og } \emptyset KG_{\bar{x}} = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}} = \mu + A_1 \cdot \sigma$$

Benyttes et R - kort er $\sigma \approx \frac{\bar{R}}{d_2}$, og dermed

$$NKG_{\bar{x}} = \bar{x} - 3 \cdot \frac{\bar{R}}{d_2 \cdot \sqrt{n}} = \mu - A_2 \cdot \bar{R} \text{ og } \emptyset KG_{\bar{x}} = \bar{x} + 3 \cdot \frac{\bar{R}}{d_2 \cdot \sqrt{n}} = \mu + A_2 \cdot \bar{R}$$

Benyttes et s kort er $\sigma \approx \frac{\bar{s}}{c_4}$, og dermed

$$NKG_{\bar{x}} = \bar{x} - 3 \cdot \frac{\bar{s}}{c_4 \cdot \sqrt{n}} = \mu - A_3 \cdot \bar{s} \text{ og } \emptyset KG_{\bar{x}} = \bar{x} + 3 \cdot \frac{\bar{s}}{c_4 \cdot \sqrt{n}} = \mu + A_3 \cdot \bar{s}.$$

APPENDIX 6.1. 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 observationsæt udregnes gennemsnit og spredning.

Faktor	Observationer	Gennemsnit	Spredning
R_1	$x_{11}, x_{12}, x_{13}, \dots, x_{1n_1}$	\bar{x}_1	S_1
R_2	$x_{21}, x_{22}, x_{23}, \dots, x_{2n_2}$	\bar{x}_2	S_2
R_3	$x_{31}, x_{32}, x_{33}, \dots, x_{3n_3}$	\bar{x}_3	S_3
R_r	$x_{r1}, x_{r2}, x_{r3}, \dots, x_{rn_r}$	\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 .

Beregninger:

Lad $N = n_1 + n_2 + n_3 + \dots + n_r$

Der beregnes et "vægtet" gennemsnit af varianserne (vægtet efter frihedsgraderne)

$$s_e^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2 + (n_3 - 1) \cdot s_3^2 + \dots + (n_r - 1) \cdot s_r^2}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1) + \dots + (n_r - 1)} = \frac{SAK_1 + SAK_2 + \dots + SAK_r}{N - r}$$

Sidste omskrivning følger af, at $s^2 = \frac{SAK}{n - 1}$

s_e^2 er variansen for "forsøgsfejlen" eller på engelsk "error". s_e^2 har $N - r$ frihedsgrader.

Vi har nu $SAK_e = (N - r) \cdot s_e^2$

Man indtaster alle N målinger og beregner variansen s_{total}^2

Heraf fås $SAK_{total} = (N - 1) \cdot s_{total}^2$

Man kan vise, at $SAK_R + SAK_e = SAK_{total}$ og at det tilsvarende gælder for frihedsgraderne
Tallene indsættes nu traditionelt i en såkaldt variansanalysetabel, og P-værdien beregnes.

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	$N - r$	$s_e^2 = \frac{SAK_e}{N - r}$		
Total	SAK_{total}	$N - 1$			

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 : $[\bar{x}_i - r_{kon}; \bar{x}_i + r_{kon}]$

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

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 $\chi^2 = \frac{\text{ovenstående}}{1 - \frac{1}{N - k} - \sum_{i=1}^k \frac{1}{n_i - 1}}$

Lad Y være χ^2 -fordelt med frihedsgrade $k - 1$.

H_0 forkastes, hvis P -værdi = $P(Y > \chi^2) < \alpha$.

c) 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:

		\hat{y}_i	d_{ij}	
T_1	108, 110, 112	110	2, 0, 2	En ensidet variansanalyse på d_{ij} giver $F = 0.15$, og dermed P -værdi = 0.9285, dvs. en accept af nulhypotesen.
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:

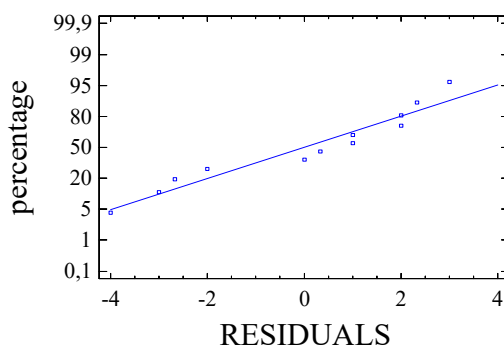
Lad residualerne (fra eksempel 4.1) -2, 0, 2, -3, 2, 1, 2.7, -2.3, -0.3, 1, 3, -4

De ordnes i rækkefølge og man beregner deres kumulative frekvens

i	Residualer x	$\frac{i-0.5}{12} \cdot 100 \%$
1	-4	4.1
2	-3	12.5
3	-2.3	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	2.7	87.5
12	3	95.8

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

Normal Probability Plot for RESIDUALS



APPENDIX 7.1. 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 7.2.

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		Rækkegennemsnit
		K ₁	K ₂	
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:

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

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

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

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

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

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

(alternativt: $SAK_{\text{rækker}} = n \cdot q \cdot SAK_{\bar{x} \text{ rækker}}$, 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 forsøgsfejls 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

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

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

Appendix

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

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

$\tilde{\mu}_{ij} = \bar{x}_{i.} + \bar{x}_{.j} - \bar{x}_{..}$ (jævnfør betragtningerne i afsnit 2.3.2.2 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:

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

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

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

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

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

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

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

Appendix 7.2. Transformation af Binomialfordelte eller Poissonfordelt variable

To binomialfordelte eller Poissonfordelte 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. Transformeret data som angivet nedenfor er det imidlertid tilladeligt at bruge variansanalyseteknikken på de transformerede data

Endvidere får man så den fordel, at man får en eksakt værdi for forsøgsfejls varians (støj), 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.

Tabel 1

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 $p(\mu)$	$Y = \sqrt{X}$	$\frac{1}{4}$

Eksempel : 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 TI89:

Da antallet af dræbte larver anses for at være binomialfordelt, foretages den i ovennævnte tabel 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.

De transformerede tal for klædefabrikat 1 gemmes i list1, klædefabrikat 2 gemmes i list 2 osv.

F6 ► ANOVA2-Way ► ENTER ► DESIGN=Block, Levels o Col Factor =4 ► ENTER

Næste skema udfyldes med list1, list2, list3, list4 og list5 ► 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		

Appendix

Block: Metoder	0.080445	3	0.026815		
Error	0.008881	9	0.000987		

Her svarer "Error" en sum af "støj"+"vekselvirkning".

Da vi fra ovennævnte tabel 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 = $P(F > 1.92) = FCdf(1.92, \infty, 3, 1000) = 0.12$

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

Skal vi finde ud af hvilken virkning der er størst, kan vi udregne konfidensintervaller for de transponerede tal.

		Klædefabrikat				Gennemsnit
		1	2	3	4	
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.

APPENDIX 9.1. Formler til beregning af enkelt regressionsanalyse uden gentagelser.

I dette appendix vises hvorledes man kan beregne en enkelt regressionsanalyse uden gentagelser, blot man har en lommeregner med regressionsprogram. I eksemplet er formlerne anvendt på et konkret eksempel.

Forudsætning:

Data :

x	x_1	x_2	x_3		x_N
y	y_1	y_2	y_3	...		y_N

De N -værdier er uafhængige observationer af stokastisk uafhængig normalfordelte variable Y_i med samme varians σ^2 .

Det antages endvidere at man har fundet, at data kan beskrives ved en lineær model.

Vi har derfor at **middelværdien** af den statistiske variable Y er en lineær funktion af x af formen

$$E(Y|x) = \beta_0 + \beta_1 x$$

Beregninger:

1) De N punktpar indtastes i lommeregner.

Regressionsprogram aktiveres, og blandt beregnede størrelser findes estimater for

regressionskoefficienter: $\hat{\beta}_0$ og $\hat{\beta}_1$, korrelationskoefficient r , gennemsnit \bar{x} , spredning s_y .

2) Udfylder variansanalysetabel:

Udregner $SAK_{total} = (N - 1) \cdot s_y^2$, $SAK_{model} = r^2 \cdot SAK_{total}$ og $SAK_{residual} = SAK_{total} - SAK_{model}$.

Variation (Source)	SAK (SS)	f (df)	$s^2 = \frac{SAK}{f}$	F
Model	SAK_{model}	1	$s_{model}^2 = \frac{SAK_{model}}{1}$	$F_{model} = \frac{s_{model}^2}{s_{residual}^2}$
Residual	$SAK_{residual}$	$N - 2$	$s_{residual}^2 = \frac{SAK_{residual}}{N - 2}$	
Total	SAK_{total}	$N - 1$		

Test:

Lad α være signifikansniveau.

1) H_0 : Regressionslinien er vandret $\Leftrightarrow H_0: y$ er uafhængig af $x \Leftrightarrow H_0: \text{Model} = 0 \Leftrightarrow H_0: \beta_1 = 0$

Metode 1.

Hvis modellen gælder så burde punkterne (uanset om H_0 er sand eller ej) ligge eksakt på en ret linie (og dermed $s_{residual}^2 = 0$), hvis ikke forsøgsresultaterne havde været påvirket af "støjen".

Et estimat for forsøgsfejls (støjens) varians σ^2 er derfor $s_{residual}^2$.

Er H_0 sand, så burde (jævnfør definitionen af SAK_{model}) s_{model}^2 være nul. Når det ikke er tilfældet skyldes det, at forsøgsresultaterne har været påvirket af "støjen". Af samme grund som før må derfor også s_{model}^2 være et estimat for σ^2 .

Vi har følgelig, at hvis H_0 er sand, så er $F_{model} = \frac{s_{model}^2}{s_{residual}^2} \approx 1$.

Det kan vises, at hvis nulhypotesen ikke er sand, så vil $F_{model} > 1$, og at F_{model} er F -fordelt med en tællerfrihedsgrad på 1 og en nævnerfrihedsgrad på $N - 2$.

Testen bliver følgelig en ensidet F -test, dvs.

H_0 forkastes, hvis P -værdi = $P(Z < F_{model}) < \alpha$, hvor Z er F -fordelt $(f_T, f_N) = (1, N - 2)$.

Metode 2. Lad $t = \frac{\hat{\beta}_1}{s_{\beta_1}}$, hvor $s_{\beta_1} = \frac{|\hat{\beta}_1| \cdot s_{\text{residual}}}{s_{\text{model}}}$ er et estimat for spredningen på β_1 .

Det kan vises, at t er t -fordelt med $N - 2$ frihedsgrader.

Lad T være t -fordelt med $N - 2$ frihedsgrader

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

En fordel ved denne metode er, at man også kan teste $H_0: \beta_1 > 0$ og $H_0: \beta_1 < 0$ ved ensidede test.

Hvis begge variable X og Y er statistiske variable kan man tilsvarende teste korrelationen ρ ved ovennævnte t -test.

2) $H_0: \beta_1 = a$, hvor a er en given konstant.

Lad $t = \frac{\hat{\beta}_1 - a}{s_{\beta_1}}$, hvor $s_{\beta_1} = \frac{|\hat{\beta}_1| \cdot s_{\text{residual}}}{s_{\text{model}}}$

H_0 forkastes, hvis $|t| > t_{1-\frac{\alpha}{2}}(N-2)$ (for $a = 0$ svarer det til ovennævnte metode 2).

Konfidensinterval for β_1 : $\left[\hat{\beta}_1 - t_{1-\frac{\alpha}{2}}(N-2) \cdot s_{\hat{\beta}_1}; \hat{\beta}_1 + t_{1-\frac{\alpha}{2}}(N-2) \cdot s_{\hat{\beta}_1} \right]$ hvor $s_{\beta_1} = \frac{|\hat{\beta}_1| \cdot s_{\text{residual}}}{s_{\text{model}}}$.

Lad $\hat{\mu} = E(Y|x = x_0)$ være et estimat for middelværdien μ for Y for en given værdi $x = x_0$.

Konfidensinterval for μ $\left[\hat{\mu} - t_{1-\frac{\alpha}{2}}(N-2) \cdot \sqrt{\hat{V}(\mu)}; \hat{\mu} + t_{1-\frac{\alpha}{2}}(N-2) \cdot \sqrt{\hat{V}(\mu)} \right]$

hvor $\hat{\mu} = \hat{\beta}_0 + \hat{\beta}_1 x_0$, $\hat{V}(\mu) = s_{\text{residual}}^2 \cdot \left(\frac{1}{N} + \frac{(x_0 - \bar{x})^2 (\hat{\beta}_1)^2}{SAK_{\text{model}}} \right)$

Prædistinationsinterval: (Konfidensinterval) for 1 ny observation for en given x -værdi:

$\left[\hat{\mu} - t_{1-\frac{\alpha}{2}}(N-2) \cdot \sqrt{\hat{Q}}; \hat{\mu} + t_{1-\frac{\alpha}{2}}(N-2) \cdot \sqrt{\hat{Q}} \right]$, hvor $\hat{Q} = s_{\text{residual}}^2 \cdot \left(1 + \frac{1}{N} + \frac{(x - \bar{x})^2 \cdot (\hat{\beta}_1)^2}{SAK_{\text{model}}} \right)$

APPENDIX 9.2. Beregning af enkelt regressionsanalyse med lige mange gentagelser

Forudsætning: Data :

x	x_1	x_2	x_3	...		x_k
y	y_{11}	y_{21}	y_{31}	...		y_{k1}
	y_{12}	y_{22}	y_{32}	...		y_{k2}

	y_{1n}	y_{2n}	y_{3n}	...		y_{kn}

y_{ij} - værdierne er uafhængige observationer af statistisk uafhængig normalfordelte variable Y_i .

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

Der antages, at der er varianshomogenitet (ønskes dette testet se under punkt b)

Lad α være signifikansniveau.

Beregninger:

a) "Lack og fit test:

H_0 : Lineær model gælder $\Leftrightarrow H_0: (x_i, \mu_i)$ ligger på en ret linie

$\Leftrightarrow H_0$: Residual for gennemsnitspunkter = 0.

- 1) For hver x - værdi x_i indtastes de n y -værdier, og man beregner spredningen s_i .

Der beregnes et estimat for den fælles varians $s_e^2 = \frac{s_1^2 + s_2^2 + \dots + s_k^2}{k}$.

s_e^2 har $k(n-1) = N-k$ frihedsgrader.

- 2) De $N=n \cdot k$ punktpar (x_i, y_{ij}) indtastes i lommeregner.

Regressionsprogram aktiveres, og blandt beregnede størrelser findes estimater for:

regressionskoefficienter: $\hat{\beta}_0$ og $\hat{\beta}_1$, korrelationskoefficient r , gennemsnit \bar{x} , spredning s_y .

- 3) Man udregner $SAK_{\text{total}} = (N-1) \cdot s_y^2$, $SAK_{\text{model}} = r^2 \cdot SAK_{\text{total}}$ og $SAK_e = (N-k) \cdot s_e^2$

$SAK_{\text{lack of fit}} = SAK_{\text{total}} - SAK_{\text{model}} - SAK_e$

- 4) Udfylder variansanalysetabel:

Variation (Source)	SAK (SS)	f (df)	$s^2 = \frac{SAK}{f}$	F
Model	SAK_{model}	1	$s_{\text{model}}^2 = \frac{SAK_{\text{model}}}{1}$	$\left(F_{\text{model}} = \frac{s_{\text{model}}^2}{s_{\text{residual}}^2} \right)$
Lack of fit	$SAK_{\text{lack of fit}}$	$k-2$	$s_{\text{lack of fit}}^2 = \frac{SAK_{\text{lack of fit}}}{k-2}$	$F_{\text{lack of fit}} = \frac{s_{\text{lack of fit}}^2}{s_0^2}$
Gentagelser (error)	SAK_e	$N-k$	$s_e^2 = \frac{SAK_e}{N-k}$	
Total	SAK_{total}	$N-1$		

- 5) H_0 forkastes, hvis P - værdi = $P(Z > F_{\text{lack of fit}}) < \alpha$, hvor Z er F - fordelt $(f_T, f_N) = (k-2, N-k)$.

Såfremt H_0 accepteres (og et residualplot også virker rimelig) fortsætter testningen:

Da såvel s_e^2 som $s_{\text{lack of fit}}^2$ nu er et udtryk for forsøgsfejls varians σ^2 , foretages en pooling:

$$s_{\text{residual}}^2 = \frac{(N-k) \cdot s_e^2 + (k-2) \cdot s_{\text{lack of fit}}^2}{N-2} = \frac{SAK_e + SAK_{\text{lack of fit}}}{N-2},$$

og F_{model} beregnes (se variansanalysetabel).

- 6) Formlerne for de forskellige test svarer nu fuldstændig til formlerne i afsnit 11.1.

b) Varianshomogenitet (se appendix 4.1)

Appendix 9.3. Transformation til lineær model .

Nr	Model	Kommentar
1	Linear model: $Y = a + b \cdot X$	
2	Exponential model: $Y = \exp(a + b \cdot X)$	$Y = e^{a+bX} = e^a \cdot e^{bX} \Leftrightarrow \ln(Y) = a + bX$ Sættes $Z = \ln(Y)$ fås $Z = a + b \cdot X$
3	Reciprocal-Y model: $Y = 1/(a + b \cdot X)$	$Y = \frac{1}{a + b \cdot X} \Leftrightarrow \frac{1}{Y} = a + b \cdot X$ Sættes $Z = \frac{1}{Y}$ fås $Z = a + b \cdot X$
4	Reciprocal-X model: $Y = a + b/X$	$Y = a + \frac{b}{X}$. Sættes $W = \frac{1}{X}$ fås $Y = a + b \cdot W$
5	Double reciprocal model: $Y = 1/(a + b/X)$	$Y = \frac{1}{a + \frac{b}{X}} \Leftrightarrow \frac{1}{Y} = a + b \frac{1}{X}$ Sættes $Z = \frac{1}{Y}$ og $W = \frac{1}{X}$ fås $Z = a + b \cdot W$
6	Logarithmic-X model: $Y = a + b \cdot \ln(X)$	Sættes $W = \ln(X)$ fås $Y = a + b \cdot W$
7	Multiplicative model: $Y = a \cdot X^b$	$Y = a \cdot X^b \Leftrightarrow \ln(Y) = \ln(a) + b \cdot \ln(X)$ Sættes $Z = \ln(Y)$ og $W = \ln(X)$ fås $Z = a + b \cdot W$
8	Square root-X model: $Y = a + b \cdot \sqrt{X}$	$Y = a + b \cdot \sqrt{X}$. Sættes $W = \sqrt{X}$ fås $Y = a + b \cdot W$
9	Square root-Y model: $Y = (a + b \cdot X)^2$	$Y = (a + b \cdot X)^2 \Leftrightarrow \sqrt{Y} = a + b \cdot X$ Sættes $Z = \sqrt{Y}$ fås $Z = a + b \cdot X$
10	S-curve model: $Y = \exp(a + b/X)$	$Y = e^{a + \frac{b}{X}} \Leftrightarrow \ln(Y) = a + \frac{b}{X}$ Sættes $Z = \ln(Y)$ og $W = \frac{1}{X}$ fås $Z = a + b \cdot W$
11	Logistic model: $Y = \exp(a + b \cdot X)/(1 + \exp(a + b \cdot X))$	$Y = \frac{e^{a+bX}}{1 + e^{a+bX}} \Leftrightarrow \ln\left(\frac{1}{Y} + 1\right) = a + bX$ Sættes $Z = \ln\left(\frac{1}{Y} + 1\right)$ fås $Z = a + b \cdot X$
12	Log probit model: $Y = \text{normal}(a + b \cdot \ln(X))$	$Y = \Phi(a + b \cdot \ln(X)) \Leftrightarrow \Phi^{-1}(Y) = a + b \cdot \ln(X)$ Sættes $Z = \Phi^{-1}(Y)$ og $W = \ln(X)$ fås $Z = a + b \cdot W$

APENDIX 9.4. Formler til beregning af multipel regressionsanalyse.

I dette appendix vises hvorledes man kan beregne en multipel regressionsanalyse, blot man har en "matematiklommeregner" med et matrixprogram.

Lad der være givet k uafhængige variable og N observationer

$(x_{i1}, x_{i2}, \dots, x_{ik}, y_i)$, $i = 1, 2, \dots, N$ og $N > k$.

x_1	x_2	\cdot	\cdot	\cdot	x_k	y
x_{11}	x_{12}	\cdot	\cdot	\cdot	x_{1k}	y_1
x_{21}	x_{22}	\cdot	\cdot	\cdot	x_{2k}	y_2
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
x_{N1}	x_{N2}	\cdot	\cdot	\cdot	x_{Nk}	y_N

Lad regressionsligningen være $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$, (1)

hvor $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ er regressionskoefficienterne

Bestemmelse af estimater for regressionskoefficienterne

Modellen kan i matrixnotation skrives $\bar{y} = X \cdot \bar{\beta}$

$$\text{hvor } \bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdot & \cdot & \cdot & x_{1k} \\ 1 & x_{21} & x_{22} & \cdot & \cdot & \cdot & x_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{N1} & x_{N2} & \cdot & \cdot & \cdot & x_{Nk} \end{bmatrix}, \quad \bar{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_k \end{bmatrix}.$$

Vi ønsker ved mindste kvadraters metode, at finde en vektor $\hat{\beta}$, der er et estimat for vektoren $\bar{\beta}$.

Løsningen til et sådant overbestemt ligningssystem $X \cdot \bar{\beta} = \bar{y}$ er (se eventuelt M. Oddershede Larsen: Matricer og lineære ligninger) bestemt ved

$$X^T \cdot X \cdot \hat{\beta} = X^T \cdot \bar{y} \quad (\text{kaldet **normalligningssystemet**}). \quad (3)$$

Matricen $X^T \cdot X$ er en kvadratisk symmetrisk matrix, som sædvanligvis ved regressionsproblemer ikke er singular. Der eksisterer derfor en invers matrix $(X^T \cdot X)^{-1}$, hvorved løsningen til normalligningssystemet (3) bliver

$$\hat{\beta} = (X^T \cdot X)^{-1} \cdot X^T \cdot \bar{y} \quad (4)$$

Herved er regressionskoefficienterne bestemt..

Beregning af variansanalysetabel.

Variation (Source)	SAK (SS)	f (df)	$s^2 = \frac{SAK}{f}$	F
Model	SAK_{model}	k	$s_{\text{model}}^2 = \frac{SAK_{\text{model}}}{k}$	$F_{\text{model}} = \frac{s_{\text{model}}^2}{s_{\text{residual}}^2}$
Residual	SAK_{residual}	$N - k - 1$	$s_{\text{residual}}^2 = \frac{SAK_{\text{residual}}}{n - k - 1}$	
Total	SAK_{total}	$N - 1$		

Som ved den ensidede regressionsanalyse, er residualerne forskellen mellem en observeret værdi y_i og den tilsvarende værdi \hat{y}_i beregnet ud fra modellen, dvs. $r_i = y_i - \hat{y}_i$.

$$SAK_{\text{residual}} = \sum_{i=1}^n r_i^2.$$

Sættes $\vec{r} = \begin{bmatrix} r_1 \\ r_2 \\ \cdot \\ \cdot \\ r_N \end{bmatrix}$ kan vi foretage følgende omskrivning

$$SAK_{\text{residual}} = \sum_{i=1}^n r_i^2 = \vec{r}^T \cdot \vec{r} = (\vec{y} - X\hat{\beta})^T \cdot (\vec{y} - X\hat{\beta}).$$

$SAK_{\text{total}} = (n-1) \cdot s_y^2$ hvor s_y^2 er spredningen på y -værdierne

$$SAK_{\text{model}} = SAk_{\text{total}} - SAk_{\text{residual}}$$

Vurdering af model

Har man ikke gentagelser kan man beregne forklaringsgraden $r^2 = \frac{SAK_{\text{model}}}{SAK_{\text{total}}}$

(se vurdering i eksempel 7.4)

Har man gentagelser kan man foretage en "lack of fit test" (se hvordan i eksempel 7.7)

Undersøgelse af om modellen kan reduceres.

1) $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$ mod H: Mindst en af koefficienterne er forskellig fra 0.

Teststørrelse F_{model} er F - fordelt med tællerfrihedsgrad k og nævnerfrihedsgrad $N - k - 1$.

Hvis P -værdi = $P(F > F_{\text{model}}) < \alpha$ forkastes H_0 , dvs. y er ikke uafhængig af x -værdierne.

2) Forkastes H_0 vil man dernæst undersøge om nogle af koefficienterne kunne være 0, dvs. teste nullhypoteserne

$H_0: \beta_i = 0$ mod den alternative hypotese $H: \beta_i \neq 0$

Teststørrelsen er $t_i = \frac{\hat{\beta}_i}{s(\hat{\beta}_i)}$, som kan vises at være t -fordelt med $n - p$ frihedsgrader, hvor p er antal regressionskoefficienter

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

Beregningen af $s(\hat{\beta}_i)$ (kaldet "standard error" for $\hat{\beta}_i$) beregnes på følgende måde.

For den såkaldte kvadratiske symmetriske $p \times p$ kovariansmatrix $\sigma^2 \cdot (X^T \cdot X)^{-1}$

(p er antal regressionskoefficienter) gælder, at

a) diagonalelementerne er varianserne for regressionskoefficienterne og

b) elementerne udenfor diagonalen C_{ij} angiver kovariansen mellem β_i og β_j

Vi beregner derfor $\sigma_{\text{residual}}^2 \cdot (X^T \cdot X)^{-1}$, idet vi erstatter σ^2 med sit estimat s_{residual}^2

Af diagonalelementerne fås $V(\hat{\beta}_0)$, $V(\hat{\beta}_1)$, $V(\hat{\beta}_2)$, $V(\hat{\beta}_3)$ osv.

Konfidensintervaller.

$$100(1 - \alpha)\% \text{ konfidensinterval for } \beta_i. \quad \underline{\underline{\hat{\beta}_i - t_{1-\frac{\alpha}{2}}(N-p) \cdot s(\hat{\beta}_i) \leq \beta_i \leq \hat{\beta}_i + t_{1-\frac{\alpha}{2}}(N-p) \cdot s(\hat{\beta}_i)}}}$$

Konfidensinterval for et til punktet \bar{x}_0 svarende værdi \hat{y}_0 :

$$\text{Lad } \bar{x}_0 = \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \\ \cdot \\ \cdot \\ x_{0k} \end{bmatrix}$$

$$\underline{\underline{\left[\hat{y}_0 - t_{1-\frac{\alpha}{2}}(N-p) \cdot \sqrt{s_{\text{residual}}^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0}; \hat{y}_0 + t_{1-\frac{\alpha}{2}}(N-p) \cdot \sqrt{s_{\text{residual}}^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0} \right]}}$$

Forklaring på formlen

Forklaringen bygger for simpelhedens skyld på det enkle regressionspolynomium $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$.

Er $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ de estimerede værdier, og indsættes punktet (x_{01}, x_{02}) i ligningen, fås den dertil svarende estimerede y -værdi $\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_{01} + \hat{\beta}_2 x_{02}$.

Ifølge reglerne for varians af en linearkombination fås

$$V(\hat{\beta}_0 + \hat{\beta}_1 x_{01} + \hat{\beta}_2 x_{02}) = V(\hat{\beta}_0) + x_{01}^2 \cdot V(\hat{\beta}_1) + x_{02}^2 \cdot V(\hat{\beta}_2) + 2x_{01} \cdot V(\hat{\beta}_0, \hat{\beta}_1) + 2x_{02} \cdot V(\hat{\beta}_0, \hat{\beta}_2) + 2x_{01} \cdot x_{02} \cdot V(\hat{\beta}_1, \hat{\beta}_2).$$

$$\text{Sættes } \bar{x}_0 = \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \end{bmatrix} \text{ og idet kovariansmatricen er } \sigma^2 \cdot (X^T \cdot X)^{-1} = \begin{bmatrix} V(\hat{\beta}_0) & V(\hat{\beta}_0, \hat{\beta}_1) & V(\hat{\beta}_0, \hat{\beta}_2) \\ V(\hat{\beta}_1, \hat{\beta}_0) & V(\hat{\beta}_1) & V(\hat{\beta}_1, \hat{\beta}_2) \\ V(\hat{\beta}_2, \hat{\beta}_0) & V(\hat{\beta}_2, \hat{\beta}_1) & V(\hat{\beta}_2) \end{bmatrix} \text{ ses, at}$$

$$V(\hat{\beta}_0 + \hat{\beta}_1 x_{01} + \hat{\beta}_2 x_{02}) = \sigma^2 \cdot \begin{bmatrix} 1 & x_{01} & x_{02} \end{bmatrix} \cdot \begin{bmatrix} V(\hat{\beta}_0) & V(\hat{\beta}_0, \hat{\beta}_1) & V(\hat{\beta}_0, \hat{\beta}_2) \\ V(\hat{\beta}_1, \hat{\beta}_0) & V(\hat{\beta}_1) & V(\hat{\beta}_1, \hat{\beta}_2) \\ V(\hat{\beta}_2, \hat{\beta}_0) & V(\hat{\beta}_2, \hat{\beta}_1) & V(\hat{\beta}_2) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \end{bmatrix}$$

$$\Leftrightarrow V(\hat{\beta}_0 + \hat{\beta}_1 x_{01} + \hat{\beta}_2 x_{02}) = \sigma^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0.$$

Konfidensintervallet bliver følgelig

$$\underline{\underline{\left[\hat{y}_0 - t_{1-\frac{\alpha}{2}}(n-p) \cdot \sqrt{s_{\text{residual}}^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0}; \hat{y}_0 + t_{1-\frac{\alpha}{2}}(n-p) \cdot \sqrt{s_{\text{residual}}^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0} \right]}}$$

Appendix

APPENDIX 10.1. Kruskal-Wallis rangtest for to eller flere statistiske variable

Lad der være givet k stikprøver, og lad den samlede population være rangordnet (se eventuelt nedenstående skema hentet fra eksempel 10.3)

							Antal	Gennemsnit \bar{R}_i	Sum af rangtal
Metode M_1	61	69	79	61	59				
Rangtal	11.5	15	16	11.5	8.5		$n_1 = 5$	$\bar{R}_1 = 12.50$	$R_1 = 62.5$
Metode M_2	62	58	47	59	63	48			
Rangtal	13	7	2	8.5	14	3	$n_2 = 6$	$\bar{R}_2 = 7.91666$	$R_2 = 47.5$
Metode M_3	57	45	60	54	57				
Rangtal	5.5	1	10	4	5.5		$n_3 = 5$	$\bar{R}_3 = 5.20$	$R_3 = 26.0$
...								...	
...								...	
Metode M_k									
Rangtal							n_k	\bar{R}_k	R_k
Total							n	\bar{R}	

$$\text{Teststørrelse: } \chi^2 = \frac{12 \cdot \sum_{i=1}^k n_i (\bar{R}_i - \bar{R})^2}{n(n+1)} = \frac{12 \cdot \sum_{i=1}^k \frac{R_i^2}{n_i}}{n \cdot (n+1)} - 3 \cdot (n+1)$$

Hvis stikprøvestørrelserne alle er større eller lig med 5 gælder, at teststørrelsen er χ^2 -fordelt med $k - 1$ frihedsgrader.

Hypotesen H_0 : De k fordelinger er ens.

H_0 forkastes, hvis $\chi^2 > \chi_{1-\alpha}^2(k - 1)$

eller P -værdi = $P(Y > \chi^2) < \alpha$, hvor Y er χ^2 -fordelt med $k - 1$ frihedsgrader.

Grundlæggende operationer på lommeregnerne TI-89 og TI-Nspire samt på PC-programmet TI-Nspire.

TI 89

Oprette en "Folder": VAR-Link\F1\ 5: Create Folder\ Skriv navn på folder.

Vælg en mappe som den aktuelle mappe: MODE\ Current Folder\navn

Formål: Det kan være praktisk ikke at gemme alle sine resultater i MAIN.

Resultater som eksakte tal eller decimaltal med et bestemt antal decimaler, eller cifre:

"Mode" vælg "Auto" hvorved man får eksakte tal,

Ønskes decimaltal: vælg gul tast, ENTER (eller skriv et tal i A som decimaltal).

Ønskes m decimaler: Mode ► Display Digits ► Fix m

Ønskes m cifre : Mode ► Display Digits ► Float m

Beregning af sandsynlighedsfordelinger:

Metode 1 Vælg HOME\ CATALOG,, F3\ vælg den ønskede fordeling\ENTER

(tryk evt på "forbogstav" for hurtigt at komme til det ønskede navn).

Fordel: Hurtig ved beregning af sandsynligheder, såsom $P(X < 0.87)$ da resultatet straks indsættes på HOME-linien.

Ulempe: Man skal huske parametrene rækkefølge (de kan dog ses nederst på skærmen)

Metode 2: Vælg APPS\ Stats/List\F5\ vælg den ønskede fordeling\ENTER

Fordel: Der fremkommer nu en menu, som er næsten selvforklarende.

Ulempe: Skal resultatet ned på HOME-linien (man vil regne videre), bliver det lidt besværligt: HOME, Var-Link\ StatsVar mappen markeres den ønskede størrelse, ENTER

Tal indlagt på liste Vælg APPS\ Stats/List\ indtast data i eksempelvis "list1"

Fordele og ulemper som under metode 2 ovenfor

Beregning af gennemsnit , spredning , middelværdi osv.

1) Hvis tal indlagt på liste 1

F4\ 1: 1-Var Stats , I menu sættes "List" til "List1" (Benyt evt. Var-Link til at finde List1)

Udskriften består af en række statistiske størrelser.

2) Anvendes med få tal og kun ønsker beregning af en enkelt størrelse.

HOME\ MATH\6.Statistics\

Gennemsnit: Mean ({liste}), Varians: Variance({liste}), Spredning: stdDev({liste})

1 faktor på 1 eller 2 niveauer

Beregning af test

a) **normalfordelte variable**

1) Hvis tal indlagt på liste(r)

F6\i menu vælg relevant test\ENTER\Data\ENTER\udfyld menu\ENTER

2) Oprindelige data ikke kendt, men kun middelværdi osv.

Som ovenfor, men nu vælges Stats fremfor Data

b) **Binomialfordeling.**

1 variabel $b(n,p)$: $H_0: p = p_0$ P-værdi = binomCdf(n, p_0, low, up)

2 variable: F6: 2-Prop-ZTest (Kræver der kan approksimeres til normalfordeling, se oversigt 3.5.3)

Poissonfordeling:

1 variabel $P(\mu)$: $H_0: \mu = \mu_0$ P-værdi = PoissonCdf($n \cdot \mu_0, low, up$)

2 variable: $H_0: \mu_1 = \mu_2$ P-værdi = NormCdf(...) (Kræver approksimere normalfordeling, se oversigt 3.5.3)

Beregning af konfidensintervaller

Som under test, idet der nu skal vælges F7 i stedet for F6 (se evt. oversigt 3.5)

1 faktor på mere end 2 niveauer

a) normalfordelte variable

Data indlagt i Stats/List

Test: F6\i menu vælg ANOVA\udfyld menu\ENTER

Konfidensinterval: Findes som ekstra søjler efter de øvrige lister.

b) Binomialfordeling og Poissonfordeling: se oversigt 6.6.2 og 6.6.3

2 faktorer på mere end 2 niveauer

normalfordelte variable

Data indlagt i Stats/List

Test: F6\i menu vælg ANOVA2Way \udfyld menu\ENTER

Finder P-værdi for vekselvirkning. Fortsættelse: Se oversigt 7.7

Konfidensinterval: Se oversigt 7.7

Regressionsanalyse

Model med 2 parametre uden gentagelser

Vurdering af model: F4: Regressions, vælg model, udfyld menu,Enter

Man kan nu aflæse koefficienter, forklaringsgrad og se liste med residualer

Test og konfidensintervaller

Hvis model ikke lineær, så skal den gøres lineær ved indføre nye liste (eksempelvis $\log(x)$)

Test af hældning: F6: LinReg T_test, udfyld menu

Konfidensinterval for hældning: F7: LinReg T-int, udfyld menu bl.a Interval=Slope

Konfidensinterval for y svarende til $x = t$, F7: LinReg T-int, udfyld menu bl.a Interval=Response,x-value =t

Prædinationsinterval ses nederst i udskrift

TI-Nspire-PC og TI-nspire- håndholdt

De to versioner er grundlæggende ens.

Forskellene er stort set,

- 1) Vælg nyt dokument m.m.
TI-Nspire: Vælg "Filer" osv.
TI-nspire: Tryk "doc", osv.
- 2) Få adgang til grundlæggende valg:
TI-Nspire: Vælg Indsæt osv.
TI-nspire: Tryk "Menu", osv.

Oprette en "Folder", Mappe

TI-Nspire: Oprette mappe inden man går ind i programmet

TI-nspire: Tryk på CTRL Menu og vælg

Gemme en matrix A eller en funktion

Ti-nspire : skriv a ► tryk :=

TI-Nspire:Skriv A:=

Oprette en matrix A:

TI-nspire: Vælg: Beregninger ► skriv a ► tryk := ► vælg menu ► Matricer og Vektorer ► osv

TI-Nspire: Vælg: Beregninger ► Skriv a:= ► Vælg Matricer og Vektorer ► osv.

Grundlæggende operationer på TI89 og TI-nspire

Resultater som eksakte tal eller decimaltal med et bestemt antal decimaler, eller cifre:

TI-nspire/TI-Nspire : Ønskes decimaltal: Vælg tryk ctrl ,Enter (eller skriv et tal som decimaltal)

TI-nspire:

Ønskes m cifre:tryk"doc" ► Indstillinger og status ► Dokumentindstillinger ► vis cifre ► Flydende m

Ønskes m decimaler:tryk"doc" ► Indstillinger og status ► Dokumentindstillinger ► vis cifre ► Fast m

TI-Nspire :

Ønskes m cifre: Vælg Filer ► Indstillinger ► Dokumentindstillinger ► vis cifre ► Flydende m

Ønskes m decimaler: Vælg Filer ► Indstillinger ► Dokumentindstillinger ► vis cifre ► Fast m

Oprette et "Dokument": Filer\ Nyt TI.Nspire dokument

Formål: Hvis man eksempelvis vil aflevere en opgave til andre skal man gemme i et dokument

Beregning af sandsynlighedsfordelinger:

Vælg Beregninger\Statistik\Fordelinger\ vælg den ønskede fordeling\udfyld menu\ENTER

Husk fordelingens navn og parametrenes rækkefølge kan man skrive direkte

Tal indlagt på liste Lister og regneark ► giv en liste et navn og indtal tal i listen

Beregning af gennemsnit , spredning , middelværdi osv.

- 1) Hvis tal indlagt på liste
Lister og regneark ► giv en liste et navn og indtal tal i listen ► vælg statistik ► statistiske beregninger ► statistik med 1 variabel ► udfyld menuer ► Enter.
Blandt mange tal findes det ønskede
- 2) Anvendes med få tal og kun ønsker beregning af en enkelt størrelse.
Beregninger\statistik\listematematik\vælg
Middel: $\text{mean}(\{\text{liste}\})$, Stikprøvevarians: $\text{varSamp}(\{\text{liste}\})$, Standardafvigelse for stikprøve: $\text{stDevSamp}(\{\text{liste}\})$

1 faktor på 1 eller 2 niveauer

Beregning af test

a) normal, 1 variabel

- 1) Lister og regneark ► udfyld liste (husk overskrift) ► Statistik ► t-test for 1 variabel ► menu:data ► udfyld menu ► ENTER
- 2) Oprindelige data ikke kendt, men kun middelværdi osv.
Som ovenfor, men nu vælges Statistik fremfor Data

b) normal, 2 variable

- 1) Lister og regneark ► Udfyld lister med overskrift m1 og m2 ► Statistik ► t-test for 2 variable ► menu:data ► ok ► menu: List1: skriv m1 ► List 2: Skriv m2 ► "alternative Hyp" ► samlet: nej (hvis parvise observationer så ja ► vælg variabelreference ► ok
- 2) Oprindelige data ikke kendt, men kun middelværdi osv.
Som ovenfor, men nu vælges Statistik fremfor Data

c) Binomialfordeling.

- 1 variabel $b(n,p)$: $H_0: p = p_0$ P-værdi = $\text{binomCdf}(n, p_0, \text{low}, \text{up})$
- 2 variable: beregninger, statistik, statistiske test, z-test for 2 andele (Kræver approksimere normalfordeling,

Poissonfordeling:

- 1 variabel $P(\mu)$: $H_0: \mu = \mu_0$ P-værdi = $\text{PoissonCdf}(n \cdot \mu_0, \text{low}, \text{up})$
- 2 variable: $H_0: \mu_1 = \mu_2$ P-værdi = $\text{NormCdf}(\dots)$ (Kræver approksimere normalfordeling, se oversigt 3.5.3)

Beregning af konfidensintervaller

Som under test blot vælges nu konfidensintervaller fremfor test

1 faktor på mere end 2 niveauer

a) normalfordelte variable

Data indlagt i Lister og regneark

Test: Statistik\statistiske test\vælg ensidet variansanalyse ANOVA\udfyld menu\ENTER

Konfidensinterval: Trykkes på xbar vises de 4 gennemsnit, tilsvarende for Lower- og Upper list

- b) **Binomialfordeling og Poissonfordeling:** se oversigt 6.6.2 + 6.6.3

2 faktorer på mere end 2 niveauer

normalfordelte variable

Data indlagt i Statistik og lister

Test: Vælg Statistik, Statistiske test, 2 sidet Anova, udfyld menuer herunder vælg "Variabelreference"

Finder P-værdi for vekselvirkning. Fortsættelse: Se oversigt 7.7

Konfidensinterval: Se oversigt 7.7

Regressionsanalyse

Model med 2 parametre uden gentagelser

Vurdering af model: Data indtastes i "lister og regneark" ► Dokumentværktøjer ► Statistik ► Statistiske beregninger ► Lineær regression ► Udfyld menu ► ok

Man kan nu aflæse koefficienter, forklaringsgrad og nederst i "resid" er liste med residualer

Test og konfidensintervaller: Hvis model ikke lineær, så skal den gøres lineær ved indføre nye lister (som $\log(x)$)

Test af hældning: Vælg statistik ► Statistiske tests ► lineær regressionstest ► udfyld menu ► ok

Konfidensinterval for hældning: Vælg statistik ► Konfidensintervaller ► Lin.Reg t-intervaller ► vælg Hældning ► Udfyld menu ► ok

Konfidensinterval for y svarende til $x = t$, Som under konfidensinterval for hældning, men vælg svar

Prædiktionsinterval ses nederst i udskrift.

FACITLISTE

KAPITEL 1

- 1.1 69.94 15.91
 1.2 (1) P -værdi = 0.48% (2) 1.315 [1.186; 1.445]
 1.3 (1) P -værdi = 0.043 (2) 32.5 [18%; 47%]
 1.4 P -værdi = 0,063179

KAPITEL 2

- 2.1 (1) $NKG_R=0$, $\emptyset KG_R=10.14$, $NKG_{\bar{x}}=451.4$, $\emptyset KG_{\bar{x}}=457.8$, tvivlsomt, (2) 1.02 %
 2.2 $NKG_s=0$, $\emptyset KG_s=1.482$, $NKG_{\bar{x}}=11.36$, $\emptyset KG_{\bar{x}}=13.62$
 2.3 (1) - (2) - (3) $\bar{x}=2928.9$ $\tilde{\sigma}=131.4$
 2.4 (1) $NKG_R=0$, $\emptyset KG_R=1.0$, $NKG_{\bar{x}}=85.28$, $\emptyset KG_{\bar{x}}=86.44$ (2) Alarm 5
 2.5 (1) $C_p=1.33$ (2) 8.88% (3) 18.5 0.42%
 2.6 $NKG=0.82$, $\emptyset KG=19.84$
 2.7 $\emptyset KG=16.3$ $NKG=0$ eller $\emptyset KG=16.43$ $NKG=0$
 2.8 -
 2.9 (1) $\emptyset KG=9.29$ $NKG=0$

KAPITEL 3

- 3.1 (1) $NKG=0$, $\emptyset KG=3.47$ (2) $c=4$, $n=300$, (3) 5.24 (4a) 0.70% (4b) 61%
 3.2 (1) $n=230$, $c=11$, (2) 78.7%, 51.8%, 27.0%, 11.3%, 0.3% (3) -
 3.3 (1) $n=180$, $c=8$, (2) - (3) - (4) ca 650.3
 3.4 (1) (82,2) (2) (60, 120, 1, 3, 3) (3) - (4) -
 3.5 (1) (78, 78, 3, 6, 6) (2) - (3) - (4) AOQL= ca 2.6%

KAPITEL 5

- 5.1 P -værdi = 0.10
 5.2 (1) P -værdi = 0.001 (2) [0.69 ; 3.13]
 5.3 P -værdi = 0.0214
 5.4 P -værdi = 0.047, [0.015 ; 1.87]
 5.5 a) P -værdi = 0.00017, b) 0.2258 c) [11.93 ; 14.96]
 5.6 (1) 87 (2) 87 (3) - (4) -
 5.7 36
 5.8 P -værdi = 0.589
 5.9 P -værdi = 0.1398
 5.10 P -værdi = 0.022
 5.11 (1) P -værdi = 0,0479 (2) P -værdi = 0.00336
 5.12 P -værdi = 0.233
 5.13 (1) P -værdi = 0.2563 (2) P -værdi = 0.0177
 5.14 (1) P -værdi = 0.0135 (2) P -værdi = 0.00023
 5.15 P -værdi = 0.00028

KAPITEL 6

- 6.1 P - værdi = 0.0004, B2
 6.2 P - værdi = 0.0218
 6.3 (1) P - værdi = 0.0064 (2) S2, S4
 6.4 (1) P - værdi = 0.0006, R forskellig fra de øvrige, ja
 6.5 (1) P - værdi = 0.0002 (2) T må foretrækkes.
 6.6 (1) - (2) -
 6.7 $\chi^2 = 5.33$, P - værdi = 0.0696
 6.8 $\chi^2 = 12.64$, P - værdi = 0.027

KAPITEL 7

- 7.1 (1) nej, P - værdi = 0.0108, (2a) K1 (2b) frit valg (2c) ikke A1 K2
 7.2 (1) ja (2) glas 2 og enten fosfor 1 eller 3.
 7.3 (1) limtyper (2) lim II eller lim III: II: [23.76 ; 28.11], III: [22.69 ; 27.04]
 7.4 (1) Kun syrer har virkning, Svovlsyre størst. (2) 26.45, [25.17 ; 27.73]
 7.5 (1) Nej, P - værdi = 0.046 (2) ja, P - værdi = 0.078
 7.6 Begge har signifikant virkning, vælg vævemethode 2, og ikke matrialtypen 3.
 7.7 Begge, Vitaminbehandling 1 eller 3.

KAPITEL 9

- 9.1 (1) nej (2) 0.7205 (3) $y = 607.1 - 21.4025x$ (4) [-28.5 ; -14.3] (5) [249.1 ; 258.8]
 9.2 (1) 70.5% (2) $y = 40.78 + 0.766x$ (3) P - værdi = 0.024 (4) [0.362 ; 1.169]
 (5) [57.95 ; 100.17] (6) [72.51 ; 85.61]
 9.3 (1) $r^2 = 0.99$ (2) $y = 40.678x^{0.53597}$ (3) 54.92 [53.36 ; 56.49]
 9.4 (a) $r^2 = 0.9606$ (b) $r^2 = 0.9947$, Punkter ligger mere tilfældigt om kurve, Ja
 (c) $y = 57.866 - 0.9707x$ (d) 13 (e) [12.65 ; 13.52]
 9.5 (1) - (2) $r^2 = 0.9805$ (3) $\tilde{\alpha}_0 = 171.8$ $\tilde{\beta}_1 = -0.1585$ (4) 92.786 [91.7 ; 93.7]
 9.6 (1) $t = 35.435 - 3.6393 \cdot \ln(\text{dosis})$ (2) [-7.55 ; -0.2726] (3) [0.265 ; 6.929]
 9.7 (1) - (2) P - værdi = 0.9746 (3) $P(\text{lack of fit}) = 0.5453$, $y = 6.683 + 0.3327x$
 (4) P - værdi = $2.9 \cdot 10^{-7}$ (5) [25.77 ; 40.76] (6) [56.9 ; 62.69]
 9.8 (1) $Y = -0.1438 + 108.959x$ (2) P - værdi = 0.144 (3) ja
 9.9 (1) $r^2 = 0.722$ (2) - (3) $Y = 1131.8 + 9.5846 \cdot x_i$ (3) 2329.96 [2267 ; 2393]
 9.10 (1) ja (2) ja (3) $Y = 21.625 + 3.4615x_1 + 6.2881x_2$ (4) 66.91, [62.74 ; 71.09]
 9.11 (1) - (2) $y = 5.199 + 0.9964x_1 + 0.1689x_3$ (3) [0.964 ; 1.029], [1.63; 1.75] (4) 5.515, [5.285 ; 5.745]
 9.12 (1) - (2) - (3) $y = 5.9167 + 2.251x_1 + 1.648x_2$ (4) [1.93 ; 2.57] (5) 30.51, [27.68 ; 33.34]
 9.13 (1) - (2) $Y = -1.144 + 0.7197x - 0.0379x^2$, $Y = 2.036$, [1.96 ; 2.114],
 9.14 (1) - (2) $Y = 68.720 + 0.5664x - 0.007625x^2$, (3) $y = 78.77$ (4) (37.14, 79.24)
 9.15 (1) $Y = 613.56 - 355.11x + 60.4679x^2 - 3.3623x^3$, (2) $y = 53.45$, [47.85; 59.05]

KAPITEL 10

10.1 $w_+ = 5$, forkastelse

10.2 $w_- = 24.5$, accept

10.3 (1) $w=29.5$, forkastelse (2) P-værdi = 0.00021

10.4 $w = 28$, accept

10.5 $w = 33$, forkastelse

10.6 $\chi^2 = 15.60$, P - værdi = 0.0013

10.7 $\chi^2 = 5.61$, P - værdi = 0.0604

STIKORDSREGISTER

A

acceptsandsynlighed 28
Acceptable Quality Level AQL 29
additiv model 83, 86, 88
adjusted R^2 121
aftagerens risiko 29
alarmkriterier 9, 15
Appendix 144
Average Outgoing Quality, AOQ 31
Average Outgoing Quality Limit, AOQL 31
Average Sample Number, ASN 33
Average Total Inspection, ATI 31

B

bagatelgrænse 44
Bartlett's test 65, 146
behandlinger 7, 80
binomialfordeling
 kontrolkort 20
 test
 een variabel 3
 to variable 51
 mere end 2 variable 73
 faktorforsøg 90
blokforsøg, randomiseret 49, 71, 89
boxplot 1

C

c - kontrolkort 22
consumers risk 29

D

determinationskoefficient 104
delforsøg 41
dimensionering 44, 49
dobbelstikprøveplan 32
 tabel 37

E

ekstrapolation 107
én faktor ad gangen 80
enkelt regressionsanalyse 101
 med gentagelser 114
 formler 155
 uden gentagelser 100
 formler 153

eksponentiel model 111
enkelt stikprøveplan 28
 tabel 36
ensidet variansanalyse 63
 beregninger-formler 64, 145

F

facitliste 165
faktorer 41
 1 faktor på 2 niveauer
 normalfordelte 45, 49
 binomialfordelte 51, 57
 Poissonfordelte 52, 57
 1 faktor på mere end 2 niveauer 62
fejl af type II 47
F- fordeling 56
F - test 49, 109
 simplificeret 65, 146
forklaringsgrad 104
 adjusted 121
formler til beregning af
 test af differens mellem
 μ_1 og μ_2 i normalfordelte prøver 54
 p_1 og p_2 i binomialfordelte prøver 57
 μ_1 og μ_2 i Poissonfordelte prøver 57
ensidet variansanalyse 63, 145
tosidet variansanalyse 80, 148
enkelt regressionsanalyse 153, 155
multipel regressionsanalyse 157
forudsætninger for
 variansanalyse 64
 regressionsanalyse 109
fuldstændig faktorstruktur 81
fuldstændig randomiseret blokforsøg 71, 89

G

gentagelser 41
godkendelseskontrol 27
godkendelsestal c 28
grænser for kontrolkort 11

H

histogram 2
hovedvirkning 86
hypotesetest

middelværdi :1 normalfordelt stikprøve 1
 middelværdi:2 normalford. stikprøver 45
 varians: 2 normalfordelte variable 48, 56

I,J

ikke-parametrisk test 53, 135
 ISO-9000 7

K

kapabilitet 18
 kapabilitetsindeks 18, 19
 konfidensinterval
 1 normalfordelt variabel 3
 1 binomialfordelt variabel 3
 differens, 2 normalfordelte variable 45, 54
 differens, 2 binomialfordelte variable 51, 57
 differens, 2 Poissonfordelte variable 52, 57
 i ensidet variansanalyse 67, 146
 i tosidet variansanalyse 83, 92, 149
 LSD 67
 regressionskoefficient 107
 formler 154
 for den til x svarende værdi af Y 110
 formler 110
 kontrolkortanalyse 9
 kontrol
 af binomialfordelt variabel 20
 af fejlprocent 20
 af normalfordelt variabel 10, 15
 af Poissonfordelt variabel 22
 løbende 15
 rektificerende 31
 kontrolkort 11, 14
 kontrolgrænser 9
 kvalitetskontrol 7
 kvalitetsstyring 7
 kvalitet
 tilfredsstillende 29
 utilfredsstillende 29
 korrelationskoefficient 107
 Kruskal- Wallis test 140, 160
 kvalitativ faktor 41, 62
 kvantitativ faktor 41, 62
 kvartilplot 2

L

lack of fit test 155
 leverandørens risiko 29

Levines test 65
 Limiting Quality LQ 29
 lineær model 101
 logaritme model 111
 LSD (Least Signifikant Difference) 67

M

median 2
 mindste kvadraters metode 102
 MR=Moving Range 14
 multipel regression 117, 157

N

nedre kontrolkortgrænse 9
 niveau for faktor 45
 NKG 15
 normalfordeling
 boxplot 2
 histogram 2
 plot 2, 147
 test, 1 variabel 1
 2 variable
 af middelværdi 45, 54
 af varians 48, 56
 np - kontrolkort 20

O

OC-Kurve 28, 33
 opgaver
 kapitel 1 5
 kapitel 2 23
 kapitel 3 38
 kapitel 5 58
 kapitel 6 77
 kapitel 7 95
 kapitel 9 128
 kapitel 10 142
 outliers 104, 108
 oversigter over
 centrale formler i kapitel 5 54
 centrale formler i kapitel 6 75
 fremgangsmåde: tosidet variansanalyse 92
 fremgangsmåde: regressionsanalyse 126

P

parvise observationer 49, 55
 planlægning af forsøg 40, 80
 Poissonfordeling

Stikord

- kontrollkort 22
- test
 - 1 variable 4
 - 2 variable 52, 57
 - mere end 2 variable 74
- polynomial regressionsanalyse 120
- poolet estimat for varians 47
- potens model 111
- proces i statistisk kontrol 8
- proceskontrol 7, 22
- proces ude af statistisk kontrol 8
- procesvariablen 8, 10
- prædestinationsinterval 110

Q

- Quality Control 7
- Quality management 27

R

- randomisering 42
- randomiseret forsøg 44, 64
- randomiseret blokforsøg 71, 89
- rangtal 156
- rangtest 53, 135
 - 1 variabel 136
 - 2 variable 53, 138
 - mere end 2 variable 140, 160
- regressionsanalyse
 - enkelt
 - med gentagelser 114
 - formler 155
 - uden gentagelser 100
 - formler 153
 - transformation 111
 - forudsætninger 109
 - multipel 117
 - polynomial 120
- regressionskoefficienter 101
- regressionsligning 101
- regressionslinie 101
- rektificerende kontrol 31
- repræsentativ 41
- risikopunkter 29
- R - kontrollkort 10
- residual 102
 - studentized 118
- rækkefaktor 82

S

- SAK = SS 87
- Satterthwaites test 45, 54
- screeningsforsøg 98
- seks sigma proces 18
- Shewart-kontrollkort 9
- s - kontrollkort 11, 13
- specifikationsgrænser 18
- SPC 7
- SS = Sum of squares 87
- statistisk gyldigt forsøg 41
- statistisk uafhængige 41
- stikprøveplan
 - enkelt 28
 - dobbelt 32
 - ækvivalente 33
- statistisk godkendelseskontrol 27
- statistisk proceskontrol 7
- stikord 168
- stikprøvestørrelse n 9, 28
- studentized residualer 117
- søjlefaktor 82

T

- Tabeller
 - Bestemmelse af enkelt stikprøveplan 36
 - Bestemmelse af dobbelt stikprøveplan 37
 - kontrollkort 11
 - Wilcoxons rangtest for 1 variabel 137
 - Wilcoxons rangtest for 2 variable 139
- tilfredsstillende kvalitet 99
- TI89 : Grundlæggende operationer 161
- TI-Nspire:Grundlæggende operationer 163
- tolerancegrænser 18
- tosidet variansanalyse 80, 84
- transformation 105, 111

U

- uafhængige statistiske variable 41
- utilfredsstillende kvalitet 29

V

- variabeltransformation 151
- variansanalyse
 - ensidet 63
 - beregninger, formler 64
 - tabel 87

tosidet 80, 84
 beregninger, formler 82
 med vekselvirkning 84
 additiv 83, 86, 88
varianshomogenitet 65 ,115
vekselvirkning 82

W

Wilcoxons rangtest for 1 variabel 136
Wilcoxons rangtest for 2 variable 138

X

x - streg kontrollkort 7

Æ

ækvivalente stikprøveplaner 33

Ø

ØKG 15
øvre kontrolgrænse 9