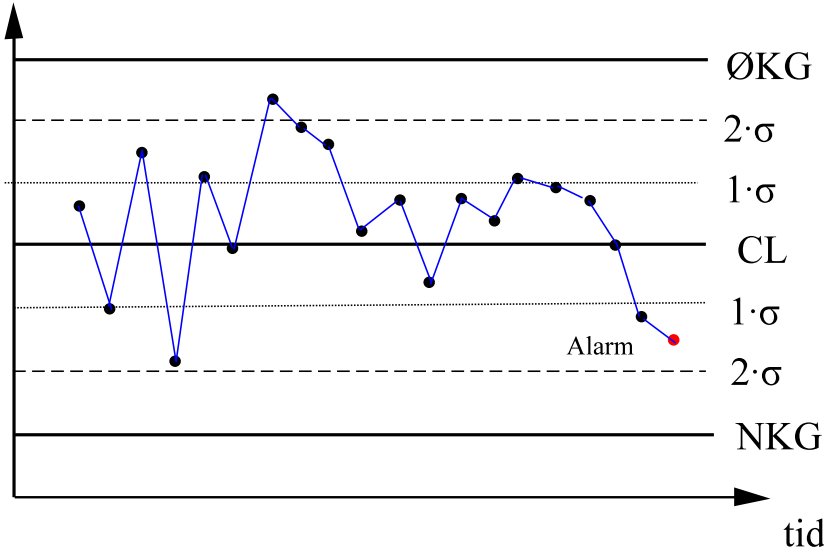


VIDEREGÅENDE STATISTIK

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

"Kontrolteori": kapitel 2 og 3

"Faktorer på 2 eller flere niveauer": kapitel 5 , 6, 7 og 8

Regressionsanalyse" kapitel 9

kan læses stort set uafhængigt af hinanden.

Kan man af tidsmæssige grunde ikke nå alt, så kan man overspringe en eller flere af delemner som

Screeningsforsøg: kapitel 8,

Statistisk godkendelseskontrol: kapitel 3.

Regnemidler

I eksemplerne er beregningerne er i videst mulig omfang foretaget ved anvendelse af statistikprogrammet SAS.JMP.

Ønskes i stedet at anvende TI-Nspire, kan man på hjemmesiden www.larsen-net.dk under "statistik 3" finde bogen Videregående statistik , version 12, hvor præces samme pensum er beregnet ved anvendelse af PC-programmet TI-Nspire.

Man kan også på samme adresse finde bøger hvor der er anvendt statistikprogrammet STATGRAPHICS og regnearket EXCEL.

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

23. juli 2017

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 af kontrolkort	9
2.4	Kontrolkortanalyse	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	16
2.7	Tolerancegrænser og kapabilitet	17
2.8	Procesvariablen X er diskret	21
	4.8.1 X er binomialfordelt	21
	4.8.2 X er Poissonfordelt	22
	Opgaver	24

3 STATISTISK GODKENDELSESKONTROL

3.1	Indledning	28
3.2	Enkelt stikprøveplan	29
3.3	Rektificerende kontrol	32
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

5.1	Normalfordelte variable	45
5.1.1	Test af differens mellem middelværdier	45
5.1.2	Parvise observationer (blokforsøg)	50
5.2	Binomialfordelte variable	51
5.3	Poissonfordelte variable	53
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 varianser 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	74
6.5	Poissonfordelte variable	75
6.6	Oversigt over centrale formler i kapitel 4.	76
6.6.1	Oversigt over fremgangsmåde ved ensidet variansanalyse	76
6.6.2	Test af parametre p_1, p_2, \dots, p_k for binomialfordelte variable	77
6.6.3	Test af parametre $\mu_1, \mu_2, \dots, \mu_k$ for Poissonfordelte variable	77
	Opgaver	78

7 2 FAKTORER PÅ 2 ELLER FLERE NIVEAUER.

TOSIDET VARIANSANALYSE

7.1	Indledning	81
7.2	Planlægning af forsøg.	81
7.2.1	Een faktor ad gangen.	81
7.2.2	Fuldstændig faktorstruktur	82

7.3	Formler og metode	83
7.4	Beregning af tosidet variansanalyse	85
7.5	Fuldstændigt randomiseret blokforsøg	92
7.6	Andre varianter af variansanalyse.	93
	7.6.1 Manglende observationer (ubalancerede forsøg)	93
	7.6.2 Trinvis variansanalyse	93
7.7	To binomialfordelte eller Poissonfordelte faktorer i et fuldstændigt faktorforsøg.	96
7.8	Oversigt over fremgangsmåde ved tosidet variansanalyse	97
	Opgaver	98

8 FLERE END TO FAKTORER, SCREENINGSFORSØG

8.1	Indledning	102
8.2	Flere end 2 faktorer med begrænset antal forsøg.	102
8.3	Screeningsforsøg.	104
8.4	Nomenklatur	107
8.5	Definitionsrelationer og aliasrelationer	109
8.6	Planlægning af et partielt 2^k faktorforsøg	111
8.7	Konfundering af blokke	114
8.8	Beregning af et partielt 2^k faktorforsøg	116
8.9	Sekventiel forsøgsstrategi.	123
8.10	Oversigt over fremgangsmåde ved partielt 2^k faktorforsøg.	124
	Opgaver	125

9 ENKELT REGRESSIONSANALYSE

9.1	Indledning.	128
9.2	Bestemmelse af regressionsligning	129
9.3	Vurdering af om regressionsligning beskriver data godt	131
9.4	Test og konfidensintervaller.	137
9.5	Transformation af data	140
9.6	Enkelt regressionsanalyse med flere y - observationer for hver x - værdi	143
9.7	Multipel regressionsanalyse	146
	9.7.1 Indledning	146
	9.7.2 Analyse med én y - observation for hver x - værdi	146
9.8	Polynomial regressionsanalyse	150
	9.8.1 Indledning	150
	9.8.2 Beregning af polynomial regressionsanalyse	151
9.9	Oversigt over fremgangsmåde ved regressionsanalyse.	156

Opgaver	158
10 RANGTEST (FORDELING UKENDT)	
10.1 Indledning	165
10.2 Wilcoxon's rangtest for 1 stikprøve	166
10.3 Wilcoxon's rangtest for 2 uafhængig stikprøver	168
10.4 Kruskal Wallis test	170
Opgaver	172
APPENDIX	
4.1 Begrundelse for grænserne for kontrolkort	174
6.1 Formler til beregning af ensidet variansanalyse	177
7.1 Formler til beregning af tosidet variansanalyse	178
7.2. Transformation af binomial-og Poissonfordelte variable til tosidet variansanalyse	183
8.1. Formler til beregning af enkelt regressionsanalyse uden gentagelser	185
8.2 Formler til beregning af enkelt regressionsanalyse med lige mange gentagelser ..	187
8.3 Transformation til lineær model	188
8.4 Formler til beregning af multipel regressionsanalyse	189
10.1 Kruskal-Wallis rangtest for to eller flere variable	192
FACITLISTE	193
STIKORD	196

1 REPETITION AF GRUNDLÆGGENDE BEGREBER 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.

Vi vil i dette afsnit kort repetere nogle af disse beregninger.

1.2. NORMALFORDELT VARIABEL

Beregning af normalfordelte sandsynligheder med SAS.JMP

SAS.JMP vælger altid at foretage disse beregninger i en "Data Table"

Beregninger starter altid på følgende måde

File ► New ► Data Table ► Giv en søjle en overskrift ► Skriv et tilfældigt tal f.eks 1 i kolonnen nedenunder ► Cursor på overskriften, højre musetast ► Vælg i menu "Formula" ► i rubrikken "no formula" kan man indsætte det ønskede (skal man eksempelvis dividere så klogest at bruge jmp-tastatur der står øverst).

Lad den statistiske variabel X være normalfordelt med middelværdi μ og spredning σ .

Der gælder da $p = P(X < x) = \text{NormalDistribution}(x, \mu, \sigma)$

Kender man omvendt sandsynligheden p og skal finde x fås den ved

$P(X < x) = p \Leftrightarrow x = \text{Normal Quantile}(p, \mu, \sigma)$

I visse situationer har man brug for den såkaldte z-funktion, som er en normalfordeling med middelværdi 0 og spredning 1.

Hvis man har en normalfordeling med middelværdi μ og spredning σ gælder følgende relation

$$x_p = z_p \cdot \sigma + \mu \text{ eller } z = \frac{x - \mu}{\sigma}$$

Man vil se relationerne anvendt i det følgende eksempel.

Eksempel 1.1 Beregning af normalfordelte sandsynligheder.

Lad X være normalfordelt med middelværdien $\mu = 7.29$ og spredningen $\sigma = 0.134$

1) Find $P(X \leq 7.2)$

2) Find $P(X > 7.6)$

3) Find x for hvilket det gælder, at $P(X < x) = 0.9$

4) Man kan ikke ændre spredningen, men ønsker at ændre middelværdien så $P(X < 7.1) = 0.025$

Løsning:

Formula ► vælg i menu "Probability" ► Normal Distribution, ► indsætter tal ► ok

$$1) P(X \leq 7.2) = \text{Normal Distribution}(7.2, 7.29, 0.134) = 0.2509$$

$$2) P(X > 7.6) = 1 - P(X \leq 7.6) = 1 - \text{Normal Distribution}(7.6, 7.29, 0.134) = 0.0103$$

$$3) x = \text{Normal Quantile}(0.9, 7.29, 0.134) = 7.46$$

$$4) 7.1 = z_{0.025} \cdot 0.134 + \mu \Leftrightarrow \mu = 7.1 - \text{Normal Quantile}(0.025) \cdot 0.134 = 7.36$$

1. Repetition af centrale begreber

Eksempel 1.2 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						

a) Vurder ud fra et histogram og et normalfordelingsplot om resultaterne er rimelig normalfordelt.

I det følgende forudsættes det at være tilfældet.

b) Hårheden bør være over 260 (brinell enheder).

Test på et signifikansniveau på $\alpha = 5\%$ om dette er tilfældet.

c) Forudsat hårheden er signifikant over 260 brinell, skal angives et estimat for hårheden, samt et 95% konfidensinterval for denne.

d) Beregn sandsynligheden for, at hårheden i middel er over 270 enheder.

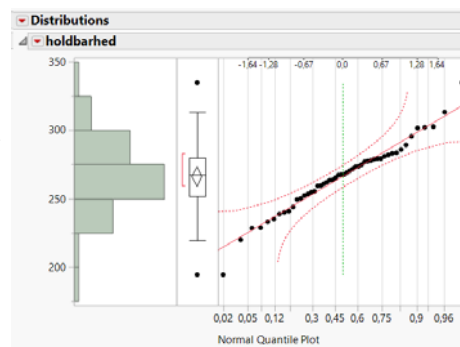
Løsning:

Data indtastes:

- a) vælg Analyze ► Distribution, ► indsætter hårhed ► ok
 Der fremkommer et histogram
 rød pil ved "hårhed" ► Normal Quantile Plot, ► ok
 Der fremkommer et normalfordelingsplot

hårhed
283,5
273,3
278,8
238,7
334,9
302,6
239,9

Da histogrammet er pænt symmetrisk med et toppunkt, og punkterne ligger tæt på linien må fordelingen være rimelig normalfordelt.



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.

Klik på Analyze ► Distribution ► marker "holdbarhed" og tryk på Y, Columns ► ok

Der fremkommer et histogram og nogle tabeller.

tryk på rød pil ved holdbarhed ► Test Mean. I den fremkomne menu skriv 260 ► ok

Test Mean	
Hypothesized Value	260
Actual Estimate	266,218
DF	49
Std Dev	25,0931
t Test	
Test Statistic	1,7522
Prob > t	0,0860
Prob > t	0,0430*
Prob < t	0,9570

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

I t Test er Prob > |t| P-værdi for den to-sidede test $H: \mu \neq 260$

Prob > t er P-værdi for den ensidede test $H: \mu > 260$, og Prob < t svarer til $H: \mu < 260$

Da P-værdi = 4.30 % < 5%, forkastes H_0 (svagt).

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

c) I en anden tabel på samme side findes "Summary Statistics"

Af udskriften ses, at

Hårheden i middel er 266.2 og

et 95% konfidensinterval er [259.09; 273.35]

Summary Statistics	
Mean	266,218
Std Dev	25,09313
Std Err Mean	3,5487045
Upper 95% Mean	273,34939
Lower 95% Mean	259,08661
N	50

$$d) P(X > 270) = 1 - P(X < 270) = 1 - \text{Normal Distribution}(270, 266,218, 25,093) = 0.44 = 44\%$$

1.3. BINOMIALFORDELT VARIABEL

Eksempel 1.3. 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 rigtig. For at teste påstanden købes 1000 af dem. Det viser sig, at 33 ud af de 1000 er defekte.

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

Løsning:

X = antal defekte chips af 1000

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

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

$$P\text{-værdi} = P(X \geq 33) = 1 - P(X \leq 32)$$

Kald en søjle for p, og indtast et tilfældigt tal i første række.

Placer cursor i p's hoved ► højre musetast ► Formula ► skriv 1-(- vælges fra jmp tastatur) ► vælg Discrete Probability ► Binomial Distribution ► Udfyld pladserne ► p=0,02, n=1000, k = 32 ► Apply

$$1 - \text{Binomial Distribution}(0,02, 1000, 32) \quad P\text{-værdi} = 0.00433$$

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



1.4. POISSONFORDELT VARIABEL

Eksempel 1.5: Poisson-test

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 $H: \mu < 147.6$.

$$P\text{-værdi} = P(X \leq 94) = \text{Poisson Distribution}(147.6, 94) = 1.52 \cdot 10^{-6}$$

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

Maksimumstemperaturen, der opnås ved en bestemt opvarmningsproces, har en variation der er tilfældig og kan beskrives ved en normalfordeling med en middelværdi på 113.3° og en spredning på 5.6°C .

- 1) Find procenten af maksimumstemperaturer, der er mindre end 116.1°C .
- 2) Find procenten af maksimumstemperaturer, der ligger mellem 115°C og 116.7°C .
- 3) Find den værdi, som overskrides af 57.8% af maksimumstemperaturerne.

Man overvejer at gå over til en anden opvarmningsproces. Man udfører derfor 16 gange i løbet af en periode forsøg, hvor man måler maksimumstemperaturen, der opnås ved denne nye proces.

Resultaterne var

116.6 , 116.6 , 117.0 , 124.5 , 122.2 , 128.6 , 109.9 , 114.8 , 106.4 , 110.7 , 110.7 , 113.7 , 128.1 , 118.8 , 115.4 , 123.1

- 4) Giv et estimat for middelværdien og spredningen.

Opgave 1.2

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 SAS-JMP, at tegne et histogram, et boxplot og et normalfordelingsplot. Angiv endvidere gennemsnit og spredning.

Opgave 1.3

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.
- c) Beregn sandsynligheden for, at indholdet af A i middel ligger mellem 1.2 og 1.3.

1. Repetition af centrale begreber

Opgave 1.4

Det forventes, at lovgivningen bliver strammet omkring mængden af skadelige partikler i bilers udstødningsgas. 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.

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.

Opgave 1.5

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	Vin	Alkoholprocent
6	Flasker	Rumindhold
7	Film	0: defekt, 1: ikke defekt
8	Tekstil	Antal fejl pr. m ²
9	Aksler	Diameter

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

2.2 STATISTISK KONTROL.

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**.

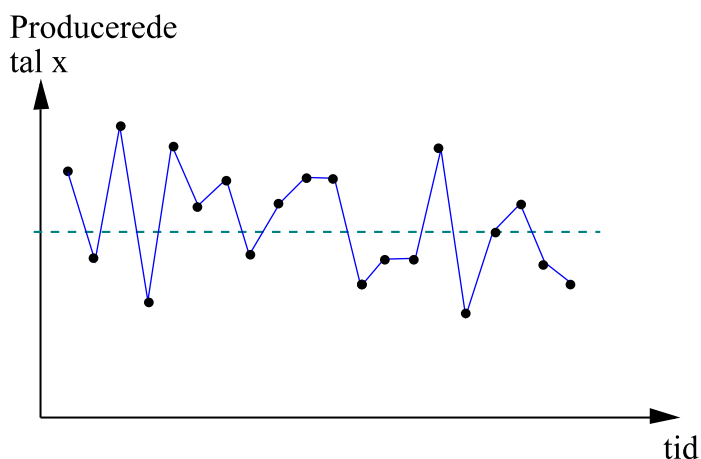


Fig 2.1. *Proces i statistisk kontrol*

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.

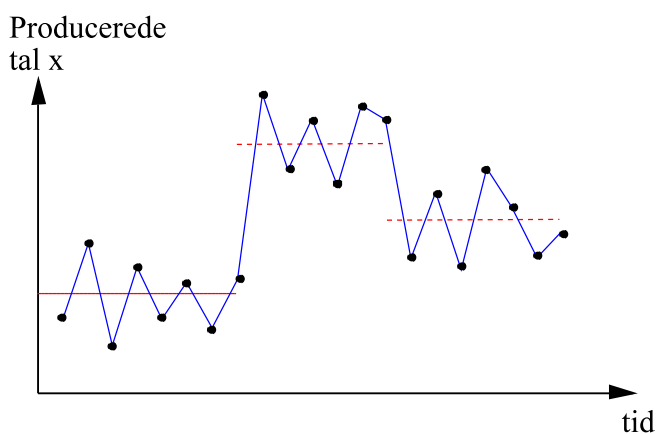


Fig.2.2. *Proces ude af statistisk kontrol*

2.3 OPBYGNING AF KONTROLKORT .

Kontrolkort (også kaldet Shewhart 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.

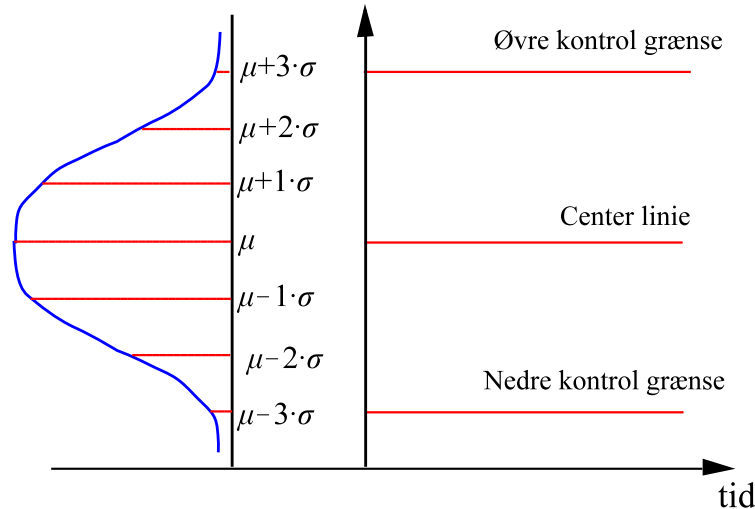


Fig 2.3. Kontrolkort med øvre og nedre kontrolgrænser

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 parametrene 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, arbejds hold 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.

Når kontrolkortet er lavet anvendes den jo i "den løbende kontrol", dvs. man foretager målinger eksempelvis hveranden time, og håber, at man så hurtigt opdager hvis processen er ved at komme ud af kontrol. I afsnit 2.6 er dette nærmere beskrevet.

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

I afsnit 2.5 behandles tilfældet hvor X approksimativt er en kontinuert normalfordelt variabel.

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 i appendix 2.1 er vist de tabeller og formler hvormed man kunne beregne dem ved "håndkraft".

Eksempel 2.1. Kontrol af stof i levnedsmiddelprodukt.

En levnedsmiddelvirkosomhed har problemer med at holde koncentrationen af et bestemt stof A i et konservesprodukt mellem nogle tolerancegrænser. Man vælger derfor at få foretaget en kontrolkortanalyse. På basis af tidligere erfaringer inddeles målingerne i 30 undergrupper, som hver har deres karakteristika:(råvarecharge, apparatur, tidspunkt på dagen osv.).

Hver undergruppe er på 5 målinger.

Gruppe	Målinger	Gruppe	Målinger
1	13 8 2 5 8	16	16 11 14 8 17
2	0 6 1 9 15	17	9 4 4 8 9
3	4 2 4 3 4	18	6 1 1 3 13
4	3 15 8 3 5	19	7 0 5 7 2
5	5 10 5 4 0	20	10 0 10 12 7
6	9 5 13 7 7	21	3 7 5 10 12
7	0 4 4 3 9	22	3 0 10 5 4
8	9 3 0 6 0	23	3 3 0 6 9
9	14 0 0 5 3	24	0 2 3 6 7
10	3 9 5 0 2	25	2 3 5 4 10
11	5 8 0 7 8	26	3 1 4 2 4
12	3 2 2 7 4	27	2 4 5 13 4
13	5 11 14 8 3	28	0 22 7 2 11
14	13 5 5 12 7	29	3 5 9 8 6
15	7 0 1 0 6	30	9 7 10 13 0

- 1) Beregn uden anvendelse af regnemidler og \bar{x} og R - værdier for gruppe 1.
- 2) Foretag ved hjælp af \bar{x} - og R - kort en kontrolkortanalyse og opstil kontrolkort, der kan benyttes til en løbende kontrol af indholdet af det undersøgte stof.

Løsning .

$$1) \bar{x} = \frac{13+8+2+5+8}{5} = \frac{36}{5} = 7.2 \quad R = 13 - 2 = 11$$

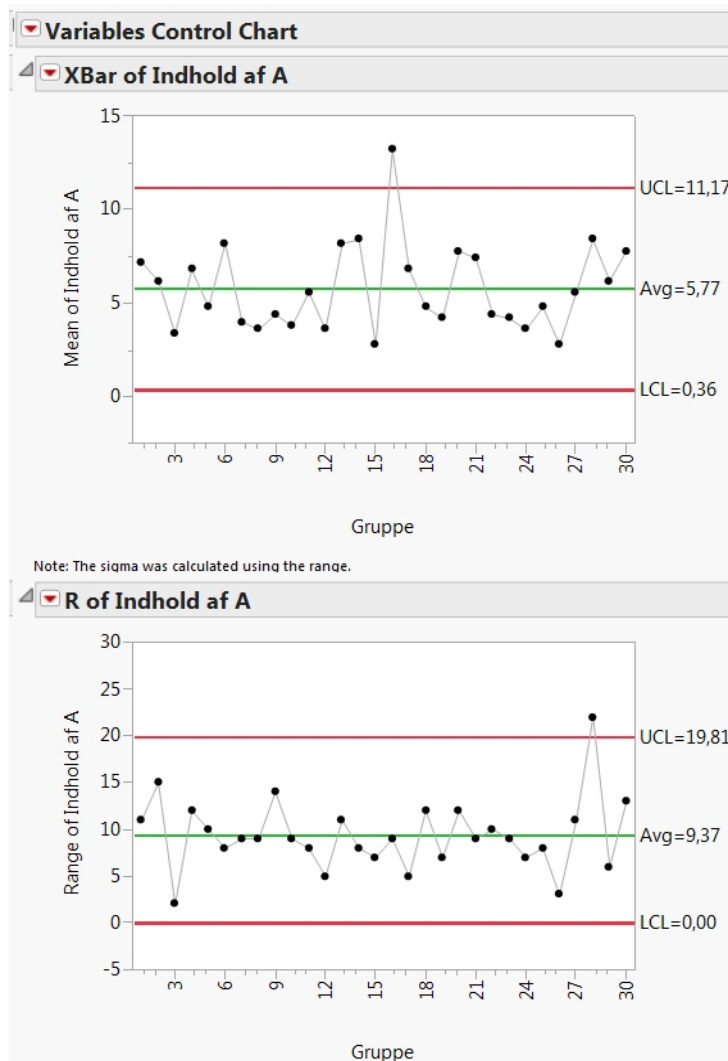
- 2) Data indtastes på sædvanlig måde:

Gruppe	Indhold af A
1	13
1	8
1	2
1	5
1	8
2	0
2	6
2	1
2	9
2	15
3	4
3	2
3	4

osv.

2. Statistisk Proceskontrol

Vælg Analyze ► Quality and Proces ► Control Chart ► X Bar ► I menu vælg Process = Indhold af A
► Sample Label = gruppe ► Marker "Xbar, R, kSigma" ► OK
Vi får følgende udskrift



Først foretages en R - kort analyse.

Det ses, at gruppe 28 er udenfor kontrolgrænserne på R -kortet.

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.

Det kan somme tider være svært umiddelbart at se om et punkt falder indenfor eller udenfor kontrolgrænsen

Det er ikke tilfældet her, men ellers kan man gøre følgende :Rød pil ved R -kort ► Test beyond limits
Nu bliver alle punkter udenfor markeret

Synes man figuren er for lille og uoverskuelig, så

Højre musetast på figur ► Size/Scale ► Y-axis ► Angiv Min , Max og Increment.

Gruppe 28 udskydes.

Cursor placeres på gruppe 28 på R-kort ► venstre musetast, I datatabel markeres nu gruppe 28 med blå ► Cursor på gruppen ► Højre musetast ► Exclude

133	27	5
134	27	13
135	27	4
136	28	0
137	28	22
138	28	7
139	28	2
140	28	11
141	29	3
142	29	5

Gentag med det nye datasæt

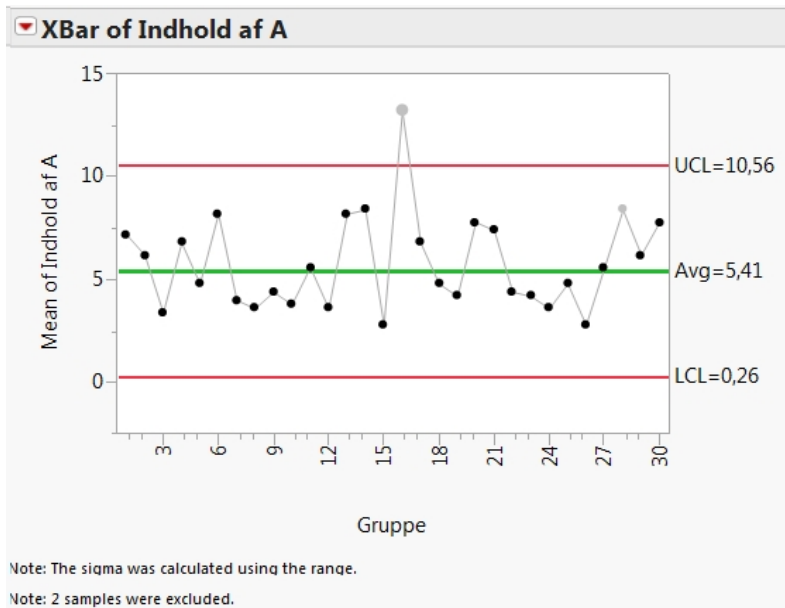
Rød pil ved "Variable Control Chart ► Redo ► Redo Analysis

Vi får nye kontrolkort med nye grænser

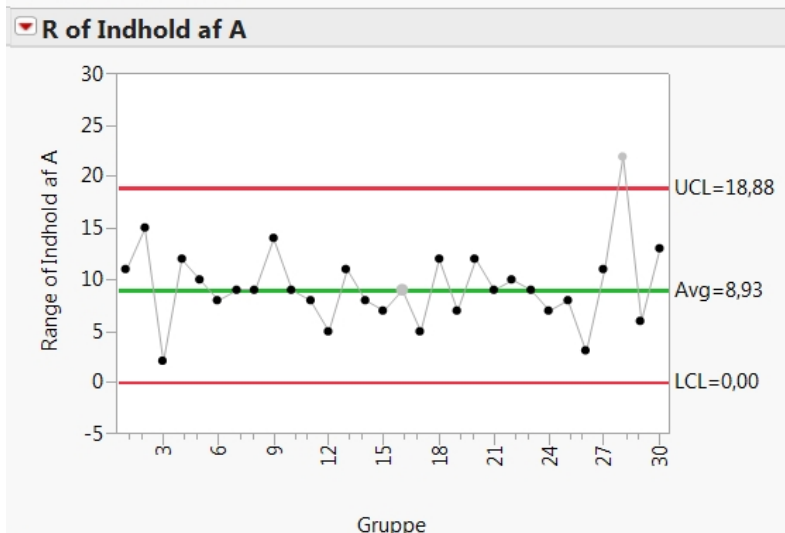
For R-kortet er UCL = 18.88.

Vi ser, at nu er der ingen udenfor R-kortet, men stadig en gruppe (gruppe 16) udenfor kontrolgrænserne på \bar{x} - kortet.

Vi udskyder nu dette punkt efter samme metode som før
Processen gentages .



Derefter er der ingen punkter på hverken \bar{x} - kortet eller R- kortet, der er udenfor grænserne
De to kort kan nu benyttes til den løbende proceskontrol.



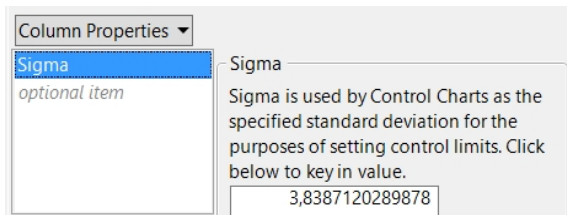
Kontrolgrænserne for middelværdien er [0.26 ; 10.56]
De to kort med de reviderede grænser kan nu benyttes til den løbende proceskontrol.

Ønskes et estimat for sprednin-

gen for en enkelt måling vælges

Vælg Rød pil ved "Variable Control Chart" ► Save Sigma

Placer cursor på overskriften i data "Indhold af A" ► højre musetast , Column Info



Et estimat for spredningen er 3.8387



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:

Vælg Analyze ► Quality and Proces

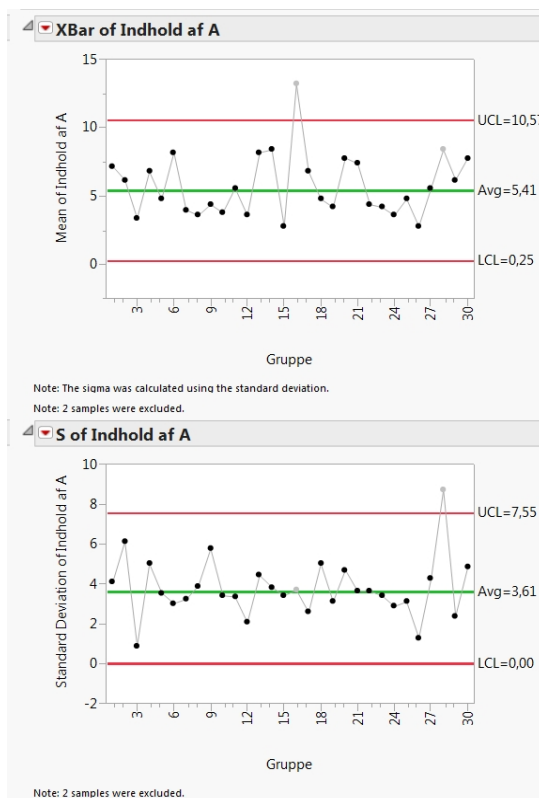
► Control Chart ► X Bar ► I menu vælg Process = Indhold af A ► Sample Label = gruppe

► Marker "Xbar, S, kSigma" ► OK

Der fremkommer som i eksempel 2.1 kontrolkort, hvor 2 punkter falder udenfor kontrolgrænserne.

Som før fjernes disse, og man får følgende kontrolkort.

Vi antager, at nu er processen i kontrol.



Det ses, at grænserne for middelværdien er $NKG = 0.25$ og $ØKG = 10.57$

Hvis vi sammenligner med R-kortet var grænserne der $NKG = 0.26$ og $ØKG = 10.56$

Forskellen skyldes, at spredningen jo bliver beregnet på en lidt forskellig måde, men forskellen er uden praktisk betydning.

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 kontrolkortet for spredningen for et "Moving range control chart"

Kontrolkortet 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) Data indtastes i en søjle benævnt x .

På værktøjslinien vælg Analyze ► Distribution

I den fremkomne menu dobbeltklikkes på " x " (under "Selected Columns") ► ok

Der fremkommer et "histogram", et "boxplot" og en række statistiske oplysninger.

For at kunne sammenligne med en normalfordeling tegnes en normalfordelingskurve:

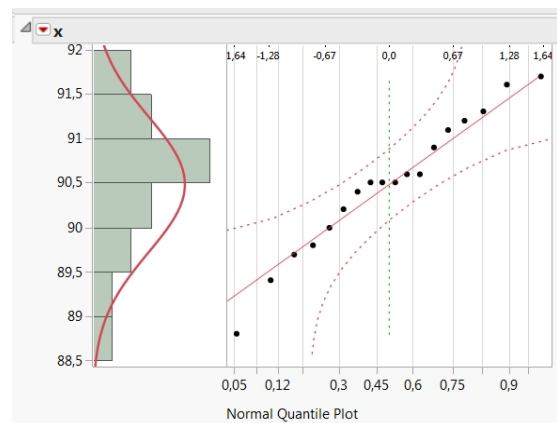
Placer cursor på "rød pil ved x " ► tryk på højre musetast og vælg "Continuous Fit ► normal

Der tegnes nu en normalfordelingskurve med samme middelværdi og spredning som fra data.

Cursor placeres på "rød pil ved x " ► tryk på højre musetast og vælger "Normal Quantile Plot"

Der er nu bl.a. fremkommet følgende figur (har slettet markering ved Box plot)

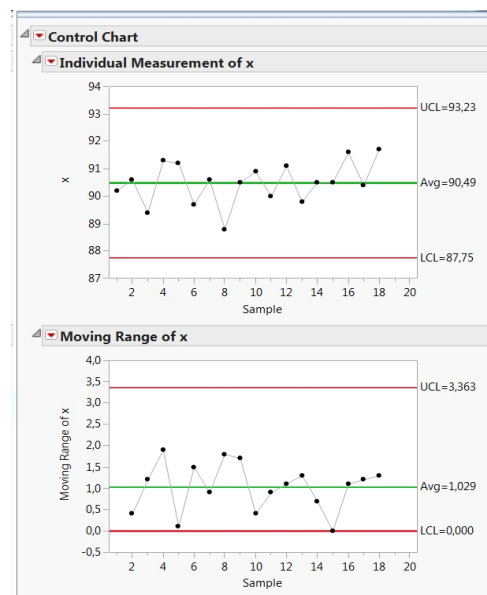
Da punkterne ligger nogenlunde på en ret linie antages fordelingen at være rimelig normalfordelt



2) På værktøjslinien vælg Analyze ► Quality and Process ► Control Chart ► IR ► indsæt x ► OK

Der fremkommer følgende to kontrollkort

Da alle punkter falder indenfor kontrolgrænserne er processen i kontrol.



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 (findes i SAS.JMP).

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.

Det følgende eksempel viser hvordan.

Eksempel 2.4. Løbende kontrol.

Lad der være oprettet et kontrolkort med $n = 3$, $\bar{x} = 5.41$ og $r = 3.84$

Man kalder en søjle for “Målinger” og indtaster 3 tal eksempelvis 7 8 5.

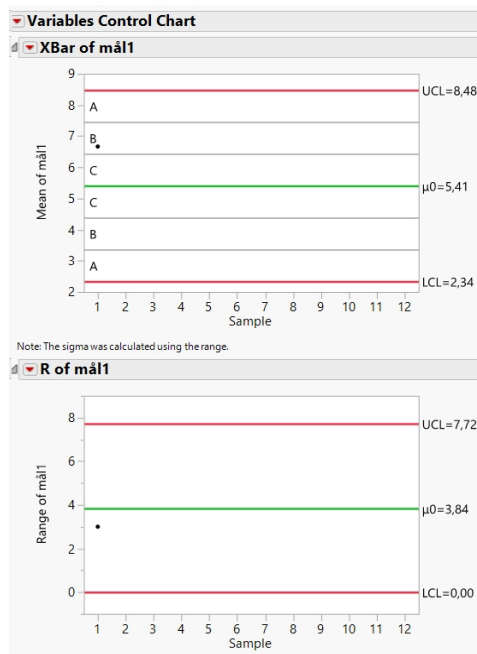
Analyze ► Quality and Proces ► Control Chart ► X Bar ► I menu vælg Process = målinger ► Sample Size Constant ► Vælg stikprøvestørrelse (3) ► Specify Stats ► indsæt mean(range) = 3,84 og mean(measure)= 5.41 ► OK

Der fremkommer nu nogle kontrolkort for xBar og R

Vælg “rød pil ved overskrift xBar

marker ► tests ► alle test og Show Center lines, Show Control lines, Show Zones, Test beyond limits.

Vi har nu følgende kort



Vi kan nu løbende indtaste vore målinger.

Trykker man på Tests (i menuen fra før), ser man, at der er 8 test .

Markerer man eksempelvis Test 1 ses hvad den tester.

Test 1: 1 point beyond Zone A.

Test 2: 9 points in a row in a single (upper or lower) side of Zone C.

Test 3: 6 points in a row steadily increasing or decreasing.

Test 4: 14 points in a row alternating up and down.

Test 5: 2 out of 3 points in a row in Zone A or beyond.

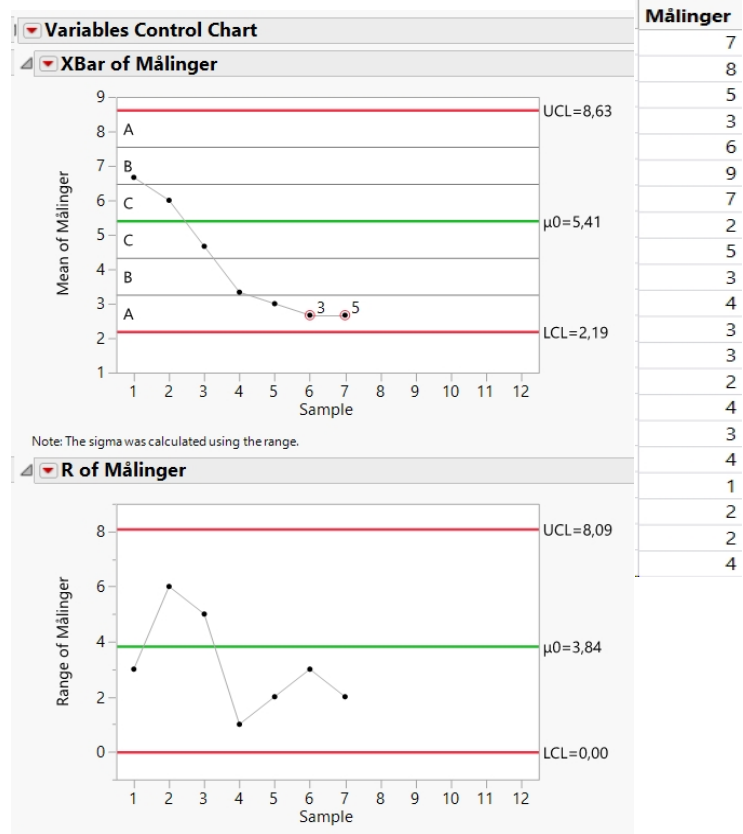
Test 6: 4 out of 5 points in a row in Zone B or beyond.

Test 7: 15 points in a row in Zone C above and below the center line.

Test 8: 8 points in a row on both sides of the center line with none in Zones C.

På basis af nogle indtastede tal fremkommer følgende kontrolkort med indtegnede punkter, og en tydelig markering af hvilke alarmgrænser der er overtrådt.

Som det ses, er der slået alarm nr 3 og 5.



2.7 TOLERANCEGRÆNSER OG KAPABILITET

I forbindelse med en fabrikation er der ofte fastsat “tolerancegrænser” eller “specifikationsgrænser.

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** $C_p = \frac{\text{ØTG}_x - \text{NTG}_x}{6\sigma}$ (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. $\text{ØTG}_x - \text{NTG}_x \approx 6\sigma$ 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 tolerancegrænse.

Endvidere er fordelingen muligvis ikke helt normalfordelt, og endvidere kendes spredningen jo ikke eksakt men man udregner på basis af nogle målinger er 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 $C_p \geq 2 \Leftrightarrow \text{ØTG}_x - \text{NTG}_x \geq 12\sigma$

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

I dette tilfælde kan fejlproduktion (næsten) helt undgås, forudsat middelværdien ligger rimelig tæt ved midtpunktet af tolerancegrænserne $\frac{\text{ØTG}_x - \text{NTG}_x}{2}$.

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 $\frac{\text{ØTG}_x - \text{NTG}_x}{2}$.

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.

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

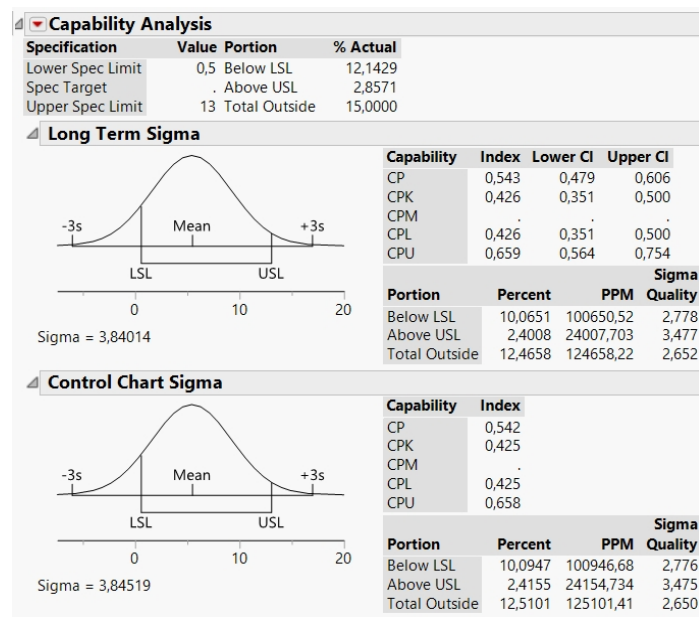
I eksempel 2.1 var ønsket at koncentrationen blev holdt indenfor tolerancegrænserne NTG = 0,5 og ØTG = 13.

Beregn kapabilitetsindeks C_p og vurder ud fra den og en beregning af fejlprocenten om processen har en rimelig lille fejlproduktion.

Løsning

Data er allerede indlagt fra eksempel 2.1 med de 2 punkter udelukket, og vi har fundet de 2 kontrolkort. Vi vælger s-kortet (forskellen mellem beregning af s er dog meget lille, så vi kunne også godt vælge R-kort)

Vælg rød pil ved "Variable Control Chart" ► Capability ► Udfyld menu: "Lower spec. Limit" : 0,5 og "Upper Spec Limit" = 13 ► OK



Man får $C_p = 0.543$ hvilket må siges at være meget lavt og da samtidig fejlprocenten på 12.5% er ret høj, giver processen alt for mange fejl, så den må ændres kraftigt.

Da C_{pk} er lidt lavere end C_p kunne man få en lidt lavere fejlprocent, hvis man kunne hæve processens gennemsnit til $\frac{13+5}{2} = 6.75$, men det vil næppe hjælpe meget.

Forklaring på forskellen mellem Long Term Sigma og Control Chart Sigma

Når man beregner kontrolkort beregnes spredningen ved at man for hver gruppe (på 5) beregner en spredning, og derefter tager gennemsnittet af tallene (30 værdier af s).

Ved "Long Term Sigma" beregner man spredningen af alle 30 · 5 = 150 tal på en gang.

Som man ser giver det en i dette tilfælde en meget lille forskel.

Imidlertid kan det for mindre stikprøver hvor målingerne måske ligger meget tæt sammen indenfor grupperne give en større forskel.

Man vælger derfor normalt at vælge udskriverne ved "Long Term Sigma"

Eksempel 2.6. Beregning af kapabilitetsindeks (data ikke givet)

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 ØTG = 24 , NTG = 2

1) Beregn kapabilitetindexet 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) X = koncentrationen af stoffet ved en enkelt måling

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

$P(X > 24) + P(X < 2) = 1 - P(X < 24) + P(X < 2) =$ 1 - Normal Distribution (24, 7.407, 2.84) + Normal Distribution (2, 7.407, 2.84)

$= 0.0285 = 2.85\%$

3) Bør indstilles på en middelværdi på $T = 13$. ◆

Det kan være nyttigt at beregne et konfidensinterval for C_p

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

Eksempel 2.7. 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

n	øvt	nedt	s	cp	ned-kon	øvre-kon
20	62	38	1,75	2,2857142857	1,5649450045	3,0055794297

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

Når man laver et sådan opstilling har man for fremtiden et "program" hvor men bare kan indsætte n, øvre og nedre tolerancegrænse

Formler:

$C_p = \frac{\text{øvt} - \text{nedt}}{6 \cdot s}$ ned-kon = $cp \cdot \sqrt{\frac{\text{ChiSquare Quantile}(0,025, 19)}{n-1}}$ øvre-kon = $cp \cdot \sqrt{\frac{\text{ChiSquare Quantile}(0,975, 19)}{n-1}}$

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 $\sqrt{n \cdot p \cdot (1-p)}$

er spredningen $\sqrt{\bar{y} \cdot \left(1 - \frac{\bar{y}}{n}\right)}$

Eksempelvis er $C_p = \frac{\text{ØTG} - \text{NTG}}{6 \cdot \sqrt{\bar{y} \cdot \left(1 - \frac{\bar{y}}{n}\right)}}$, ligesom det forklarer formlerne i oversigt 2.3.

Oversigt 2.3

np - kontrolkort. Procesvariablen X er binomialfordelt b(n,p)				
Forudsætning	Centerlinie	Nedre kontrolgrænse	Øvre kontrolgrænse	Estimater
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 2.8 (np - kort)

En fabrikant af nogle specielle typer keramikfliser som er beregnet til at kunne klare høje temperaturer ønsker udarbejdet et kontrolkort.

a) 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

Lav et kontrolkort.

b) Til brug ved den løbende kontrol sættes ØTG = 13 og NTG = 0.

Beregn kapabilitetindex C_p , og vurder om fejlprocenten er rimelig.

Løsning

a) Procesvariablen X er bestemt ved :

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

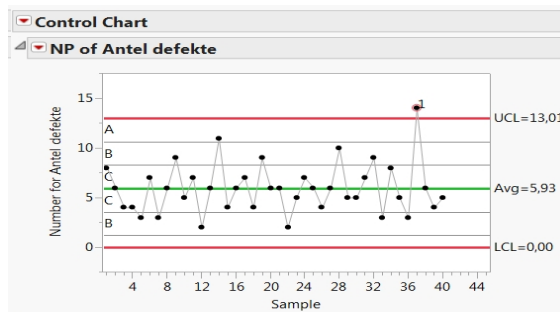
X er binomialfordelt $b(100, p)$

Data indtastes i en kolonne "antal defekte"

Vælg *Analyze* ► *Quality and Progress* ► *Control Chart* ► *NP* ► I menu vælg *Process = antal defekte* ► *Constant Size = 100* ► *OK*

Vælg rød pil på øverste overskrift ► *Tests = ALL tests* ► *Show Zones*

Der fremkommer følgende kort



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 $\hat{p} = \frac{\bar{y}}{100} = 0.0593$.

b) $C_p = \frac{OTG - NTG}{6 \cdot \sqrt{n \cdot p \cdot (1 - p)}} = \frac{13 - 0}{6 \cdot \sqrt{100 \cdot 0.0593 \cdot (1 - 0.0593)}} = 0.912$

Da værdien er mindre end 1.33, er processen ikke helt god. ◆

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 $\hat{\mu} = \frac{\bar{c}}{n}$, 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 2.9.(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 kontrollkortanalyse skal konstrueres et c-kort for processen

Løsning

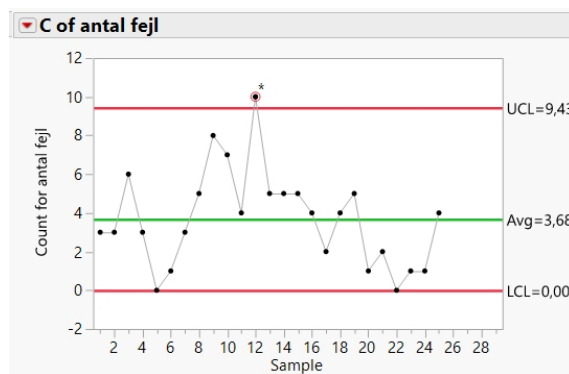
Idet det antages, at den variable er Poissonfordelt dannes et c-kort på samme måde som np-kortet.

Data indtastes i en kolonne “antal fejl”

Vælg Analyze ► Quality and Progress ► Control Chart ► c ► I menu vælg Process = antal fejl

► Constant Size =1 ► OK

Vælg rød pil ved overskrift ► Test Beyond Limits

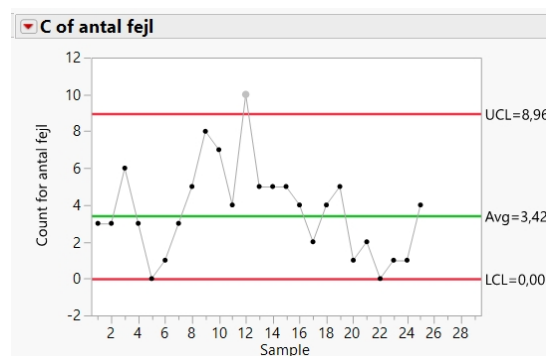


Punkt nr 12 falder uden for kontrolgrænserne, dvs. processen formodes at have været ude af statistisk kontrol på det pågældende tidspunkt.

Disse problemer bør undersøges nærmere.

Går man ud fra, at man vil se bort fra de målingen fjernes dette.

Fjernes den pågældende undergruppe, fås følgende reviderede kontrolgrænser ud fra den reviderede estimerede processtilstand:



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

$$\bar{c}_1 = 3.42, \quad \underline{NKG} = 0, \quad \underline{\text{ØKG}} = 8.96$$



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
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

1) Lav et $\bar{x} - R$ - kontrolkort.

2) Estimér, hvor stor en del af produktionen der ville være fejlproduktion (underfyldte glas under 450 g), 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	13.0	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 kontrolgrænserne 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 angiv et estimat for processens middelværdi og spredning.

Opgave 2.4

På en papirfabrik fabrikeredes en bestemt papirtype ved en af virksomhedens maskiner. Under fabrikationen målt løbende værdien af en bestemt egenskab G (vægt/arealenhed).

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 kontrolkortene til en løbende kontrol.

Opstil et kontrolkort til den løbende kontrol (med alarmgrænser), og indsæt følgende 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	86.9
	86.1	86.7	86.5	86.6	86.3	85.3	85.1	85.1	85.0	85.3

Hvis en alarm forekommer, skal man angive 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

Ved indførelse af en proceskontrol observeret følgende data:

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

a) Foretag en kontrolkortanalyse med opstilling af \bar{x} - kort og R - kort.

b) Der er fastlagt følgende tolerancegrænser for produktionen: $NTG = 11$ $\text{Ø}TG = 15$

Beregn kapabilitetsindeks C_p , og vurder om procesen holder sig rimeligt indenfor grænserne.

Opgave 2.7

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

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

Opgave 2.8

På en samlefabrik indgik en aksemontering. Som et led i en proceskontrol ønskede man specielt at undersøge risikoen for defekte samlinger. Der kan højst findes 1 defekt samling på en aksemontering.

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

Lav kontrolkort med henblik på kontrol af den fremtidige produktion .

Opgave 2.9

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	14	6	9	11	15	8	3	6	7	4	9	10

a) Opstil et kontrolkort på basis af disse tal.

Hvis der er punkter der falder udenfor kontrolkortet fjernes disse.

Angiv kontrolgrænserne og gennemsnittet af fejl på basis af dette kort.

b) Idet tolerancegrænserne er $\text{ØTG} = 18$ og $\text{NTG} = 0$ skal man vurdere om antallet af mindre fejl er forsvindende.

Vurderingen skal ske ved bl.a. at se på kapabilitetsindeks C_p .

Opgave 2.10

For antal defekte på kvadratiske ståloverflader skal foretages en løbende kontrol.

Antallet er poissonfordelt med en middelværdi på $c = 4$.

a) Opret et kontrolkort med alarmgrænser (bemærk, at der kun testes på 4 alarmer)

b) Indtast nedenstående tal på kortet, og angiv hvornår og hvilke alarmgrænser der er overtrådt.

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	7	8	1	7	8	9	9	7	7	8	7	8	3	1	11	2	4	5

Opgave 2.11

Der foretages en løbende kontrol af plastikbægre på den måde, at man udtager 200 bægre og optæller antallet af defekte bægre. Det oplyses, at $\text{ØTG} = 14$ og $\text{NTG} = 0$, samt at $np=4$.

a) Beregn kapabilitetsindeks C_p , og vurder om fejlprocenten er rimelig.

b) Lav et kontrolkort med passende alarmgrænser.

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 aftageren 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. Tegning af OC - kurve.

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

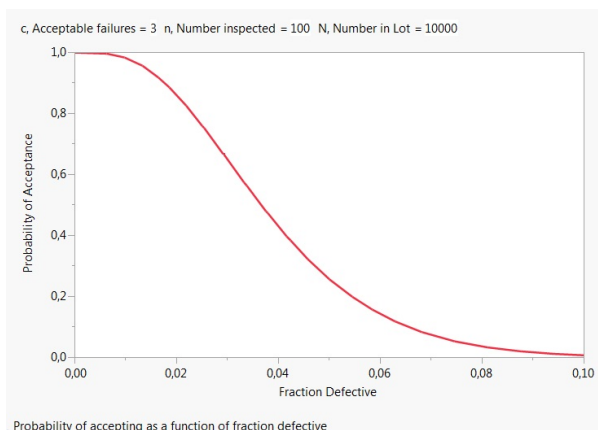
a) Tegn stikprøveplanens OC - kurve.

b) Aflæs på kurven acceptsandsynligheden for fejlprocent 5%.

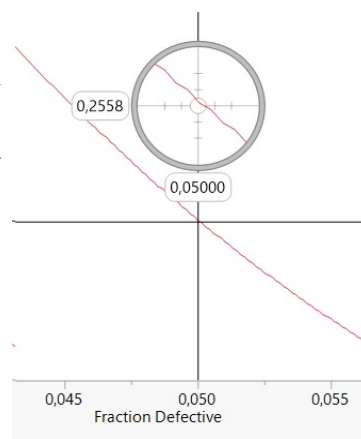
c) Beregn acceptsandsynligheden P_a for fejlprocenten 5 %

Løsning:

a) Vælg View ► JMP-starter ► Control ► OC-curves ► Acceptance-single ► ok ► indsæt n = 100, c = 3, N = 10000



b) Vælg den smalle ramme øverst på figur .Der fremkommer en menu. Vælg Tools og crosshair (evt.magnifier)
Følgende figur fremkom(som man kan stabilisere ved at trykke på højremusetast), hvor man kan aflæse at for p = 5% skønnes acceptsandsynligheden at være 25.6%



c) 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 $h(N, N \cdot p, n) =$ Hypergeometric Distribution $(N, N \cdot p, n, c)$

Indtastes tallene i et skema
(hvorved det er hurtigt at beregne mange punkter), fås

N	n	c	p	accept
10000	100	3	0,05	0,2564...
10000	100	3	0,02	0,8598...
10000	100	3	0,06	0,1417...

$$P_a(0.05) = P(X \leq 3) = \underline{\underline{0.256}}$$

Ved at man udtager stikprøven fra en meget stor mængde, kunne man i stedet approksimere med binomialfordeling ◆

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.

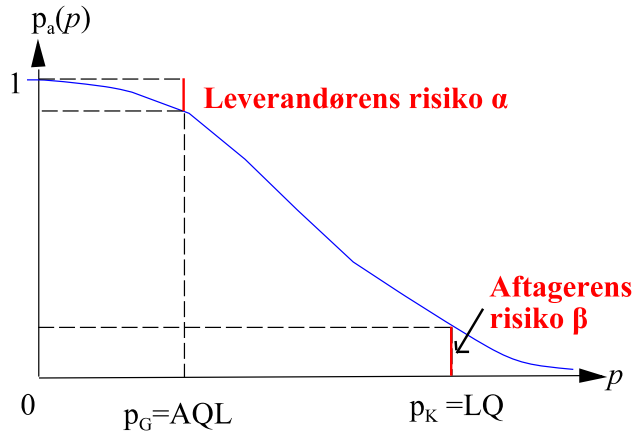
1) den tilfredsstillende kvalitet AQL (Acceptable Quality Level), og den tilsvarende acceptsandsynlighed $P_a(AQL) = 1 - \alpha$, hvor α kaldes “leverandørens risiko”.

(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 $\beta = 5\%$.



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

Idet $AQL = 0,6\%$ og $LQ = 3,1\%$ er $R = \frac{LO}{AQL} = \frac{p_K}{p_G} = \frac{31}{0,6} = 51,667$.

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 $n = 229$

Kontrol: Tallene indtastes i skemaet fra eksempel 3.3.

$p = 0,6\%$ $P(X < 0,006) = 0,9518 = 95,18\%$

$p = 3,1\%$ $P(X < 0,031) = 0,07134 = 7,13\%$

N	n	c	p	accept
10000	229	3	0,006	0,95175...
10000	229	3	0,031	0,07134...

Som det ses, passer det fint, hvis vi er leverandør, mens aftagerens risiko er noget mindre end 10%. Da vi jo kun kan udtage en stikprøve med et helt antal elementer, er det umuligt at få begge krav til at passe præcist.

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) storaftager 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 formelen $AOQ = p \cdot P_a(p)$.

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 formelen $ATI = n + (N - n) \cdot (1 - P_a(p))$

Bevis:

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

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

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 $b(229, 0.02)$

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

og dermed og $ATI = 229 + (10000 - 229) \cdot (1 - 0.326) = 68112$ ◆

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

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

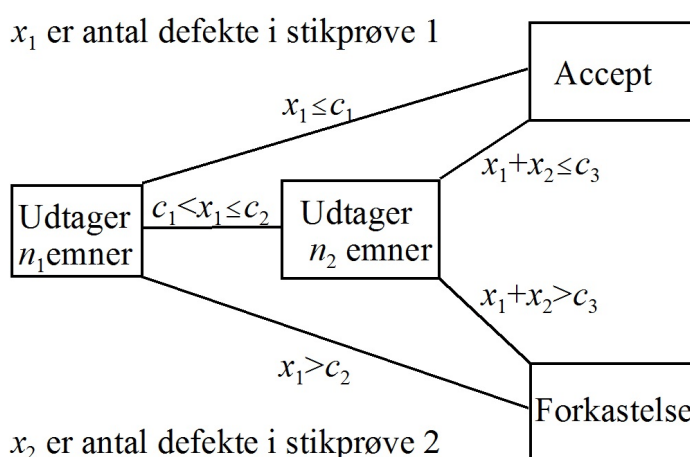
Ved en dobbelt-stikprøveplan $(n_1, n_2, c_1, c_2, c_3)$ udtages tilfældigt en stikprøve af størrelsen n_1 og antallet x_1 af defekte i stikprøven optælles.

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

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) Tegn OC-kurven
- 3) Find acceptandsynligheden for $p = 2\%$ ved aflæsning på kurven.
- 4) Find ASN for de to risikopunkter, samt for $p = 2\%$ og 4% .
- 5) Tegn punkterne for den enkelte stikprøveplan og for den dobbelte stikprøveplan for de i punkt 4) fundne værdier, samt forbind punkterne med rette linier.

LØSNING:

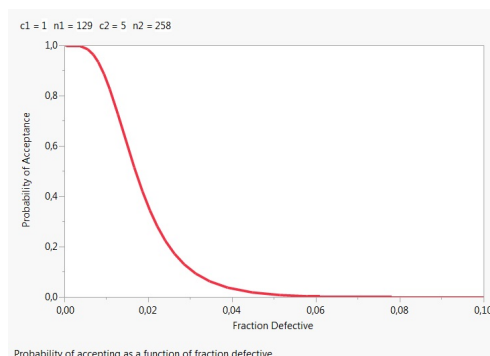
1) Idet $AQL = 0,6\%$ og $LQ = 3,1\%$ er $\frac{LO}{AQL} = \frac{p_k}{p_G} = \frac{31}{0,6} = 51,667$.

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

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

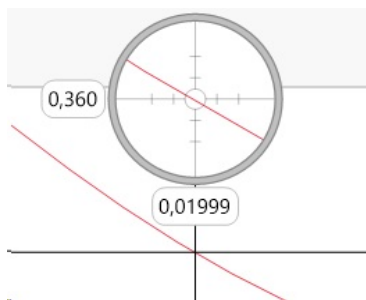
Vi finder derfor $(n_1, n_2, c_1, c_2, c_3) = (129, 258, 1, 4, 4)$.

- 2) Vælg View ► JMP-starter ► Control ► OC-curves ► Acceptance-Double ► ok ► indsæt $c_1=1, n_1 = 129, c_2 = 5, n_2 = 258$



- 3) Vælg den smalle ramme øverst på figur og klik på mus to gange. Der fremkommer en menu. Vælg Tools og crosshair (evt. magnifier)

Følgende figur fremkom (som man kan stabilisere ved at trykke på højremusetast), hvor man kan aflæse at for $p = 2\%$ skønnes acceptandsynligheden at være 36%.



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

Der laves to kolonner p og ASN

I kolonne p skrives 0,6

I ASN vælges Formula og man indtaster

Der fremkommer nu ASN-værdien

Man indtaster nu 2 osv.

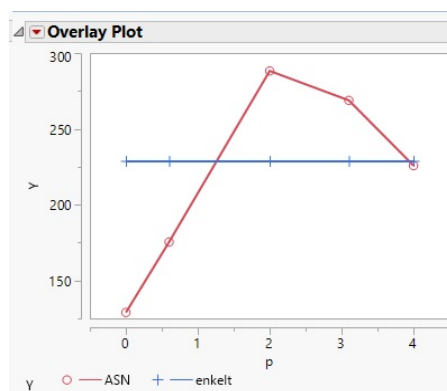
Resultatet ses i skemaet

p	ASN
0,6	175,46046877
2	288,28831486
3,1	268,90253039
4	225,86299013
0	129

5) Der indlægges nu en yderligere søjle

p	ASN	enkelt
0,6	175,46046877	229
2	288,28831486	229
3,1	268,90253039	229
4	225,86299013	229
0	129	229

Vælg Graph ► Overlay Plot ► Indlæg ASN og enkelt i y og p i x ► ok
Rød pil ved Overlay Plot ► Connect Thru Missing



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 enkelte stikprøveplan.

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)	10% (p_K)
$n_1 \cdot p$						
1	14.50	$2 \cdot n_1$	0	1	16	232
2	11.90	n_1	0	1	21	250
3	8.07	$2 \cdot n_1$	0	2	30	242
4	7.54	n_1	1	2	52	392
5	6.79	n_1	0	2	43	296
6	6.48	$2 \cdot n_1$	1	3	60	389
7	5.39	n_1	1	3	76	411
8	5.39	$2 \cdot n_1$	0	3	49	264
9	5.09	$2 \cdot n_1$	1	4	77	392
10	4.65	n_1	2	4	116	539
11	4.31	$2 \cdot n_1$	0	4	68	293
12	4.25	n_1	1	4	104	442
13	4.19	$2 \cdot n_1$	1	5	96	402
14	3.88	n_1	2	5	143	555
15	3.63	n_1	3	6	187	678
16	3.60	$2 \cdot n_1$	1	6	116	417
17	3.38	n_1	2	6	172	582
18	3.26	$2 \cdot n_1$	2	8	168	547
19	3.21	n_1	3	7	215	691
20	3.09	n_1	4	8	262	810
21	2.96	$2 \cdot n_1$	3	10	227	672
22	2.85	n_1	4	9	290	826
23	2.77	$2 \cdot n_1$	3	11	246	682
24	2.62	$2 \cdot n_1$	4	13	307	805
25	2.60	n_1	5	11	368	956
26	2.46	$2 \cdot n_1$	4	14	329	811
27	2.44	n_1	5	12	400	977
28	2.32	n_1	5	13	435	1008
29	2.22	n_1	5	14	470	1045
30	2.21	$2 \cdot n_1$	3	15	341	755
31	2.12	n_1	5	16	539	1141
32	1.97	$2 \cdot n_1$	4	20	475	935

Opgaver

Opgave 3.1

Et firma modtager leverancer på 1000 stk kuglepenne. Leverancerne kontrolleres ved at man udtager 50 stk, og godkender partiet, hvis der er 2 eller færre defekte.

- Tegn stikprøveplanens OC-kurve
- Aflæs på kurven acceptsandsynligheden for en fejlprocent på 3%
- Beregn acceptsandsynligheden p_a for fejlprocenten 3%.

Opgave 3.2

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.

- Angiv øvre og nedre kontrolgrænse for et kontrolkort, 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%.

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

- Beregn aftagerens risiko ved en fejlprocent på 3.

- 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,

- i 2 forretninger ud af de 5,
- i alle 5 forretninger.

Opgave 3.3

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%.

- Bestem stikprøvestørrelsen n (afrund til nærmeste med 10 delelige tal) og godkendelsestallet c .

- Tegn stikprøveplanens OC-kurve.

Aflæs på tegningen acceptsandsynligheden for $p = 4\%$

- Beregn for den valgte stikprøveplan acceptsandsynligheden for $p = 4\%$.

Opgave 3.4

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 .

2) Tegn OC-kurve

3) Aflæs på tegningen acceptsandsynlighed for $p = 10\%$

4) Beregn acceptsandsynligheden for $p = 10\%$

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.

5) 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
------	------	------	------	------	------	------	------	------

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

7) 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 3.5

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) Tegn OC - kurven

4) Aflæs acceptsandsynligheden for $p = 4\%$.

5) Beregn den gennemsnitlige stikprøvestørrelse ASN for de to risikopunkter, samt for $p = 4\%$ og for $p = 10\%$ og afbild resultatet grafisk.

6) Giv på basis heraf en vurdering af om det ville være en fordel at gå over til den dobbelte stikprøveplan.

Opgave 3.6

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 den dobbelte stikprøveplans parametre.

2) Tegn OC - kurven, og aflæs acceptsandsynligheden for $p = 5\%$.

3) Bestem den gennemsnitlige stikprøvestørrelse for $p = 9\%$

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 4.1 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 4.1 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

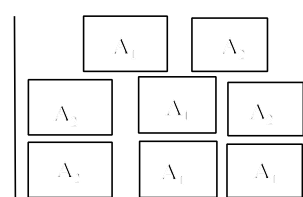
Ét delforsøg hvor apparatur nr 3 og laborant C benyttes.

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

4.3.Krav til statistisk gyldigt forsøg

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 side 54

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 SAS.JMP.

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, beregnes 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 M₁ er mindst 2 minutter mindre end ved metode M₂.

Løsning:

- a) Lad X_1 = tidsforbrug ved anvendelse af metode M₁ og X_2 = tidsforbrug ved anvendelse af metode M₂.

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.

Der anvendes 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).

Man vil i et sådant tilfælde ofte være interesseret i hvor differensen ligger med 95% sikkerhed. Man vil derfor se på et 95% konfidensinterval.

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

5. 1 Faktor på 2 niveauer

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.

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 .

Nullhypotese $H_0: \mu_1 - \mu_2 = d$ (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

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

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

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

Data indtastes i 2 søjler. Derefter lægges tallet 2 til alle tal fra metode 1

Dette er her sket ved at man har lavet en ny søjle med ændringerne, og så ved hjælp af formula har skabt en ny søjle "ny tid"

Metode	Tidsforbrug	Metode	Tidsforbrug	x	ny tid
m1	87,8	m1	87,8	2	89,8
m1	91,9	m1	91,9	2	93,9
m1	89,8	m1	89,8	2	91,8
m1	89	m1	89	2	91
m1	92,6	m1	92,6	2	94,6
m1	89,4	m1	89,4	2	91,4
m1	91,4	m1	91,4	2	93,4
m1	88,7	m1	88,7	2	90,7
m1	90,1	m1	90,1	2	92,1
m1	92,4	m1	92,4	2	94,4
m2	92,4	m2	92,4	0	92,4
m2	94,6	m2	94,6	0	94,6
m2	93	m2	93	0	93
m2	94	m2	94	0	94
m2	92,4	m2	92,4	0	92,4
m2	92,9	m2	92,9	0	92,9
m2	96,4	m2	96,4	0	96,4
m2	92,1	m2	92,1	0	92,1
m2	92,8	m2	92,8	0	92,8
m2	93,4	m2	93,4	0	93,4

Vi har nu:

$$H_0: m_1 = m_2 \quad H: m_1 < m_2$$

Analyze ► Fit by x ► Y-Response: ny tid ,x-Factor: Metode ► ok ► rød pil øverst på figur, t-test

t Test

m2-m1

Assuming unequal variances

Difference	1,0900	t Ratio	1,628317
Std Err Dif	0,6694	DF	16,99463
Upper CL Dif	2,5024	Prob > t	0,1219
Lower CL Dif	-0,3224	Prob > t	0,0609
Confidence	0,95	Prob < t	0,9391

Da vi skal vise, at $m_2 - m_1 > 0$ følger det heraf (se lige under t-Test) at vi skal se på Prob>t

Vi har følgelig, at P-værdi = 0.0609 .

Da P-værdi > 0.05 accepteres nulhypotesen, dvs. det er ikke muligt på dette grundlag at bevise, at tidsforbruget ved metode M₁ er 2 minutter mindre end ved metode M₂.

Dette fremgår også af konfidensintervallet, [-0.32 ; 2.50] som jo indeholder tallet 0.

Da P-værdien var så tæt ved 5%, vil vi nu forsøge med den stærkere t-test, hvor kravet er en fælles "pooled" spredning. Vi gør som før, men vælger nu Means/Anova/pooled t

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).

t Test			
m2-m1			
Assuming equal variances			
Difference	1,0900	t Ratio	1,628317
Std Err Dif	0,6694	DF	18
Upper CL Dif	2,4964	Prob > t	0,1208
Lower CL Dif	-0,3164	Prob > t	0,0604
Confidence	0,95	Prob < t	0,9396

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

5. 1 Faktor på 2 niveauer

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. ◆

Er de oprindelige tal ikke kendt, men kun gennemsnit og spredning må man anvende en t-test, hvor formlerne står i oversigt 5.5.1. Har man mange forsøg (over 30 for hver faktor) så er det beregningsmæssigt nemmere at anvende en z- test.

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 $s_2 = 14.2$

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

Løsning:

$H_0: \mu_1 + 10 = \mu_2$ H: $\mu_1 + 10 \leq \mu_2$ eller $H_0: \mu_1 = \mu_2 - 10$ H: $\mu_1 \leq \mu_2 - 10$

Benyttes oversigten 5.5 1 bliver man nødt til at sætte $d = -10$

Indlægges tallene i SAS.JMP fås

x1	s1	n1	x2	s2	n2	a	b	c	g	f	d	t	P-værdi
118,6	12,13	100	133,2	14,2	100	1,471369	2,01...	3,4877...	193,280...	194	-10	-2,463110...	0,0073226..

P-værdi = 0.00732 = 0.73%

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

Formler: $f = \text{Ceiling}\left(\frac{g}{c}\right)$ $t = \frac{(\bar{x}_1 - \bar{x}_2 - d)}{\sqrt{\frac{c}{n_1} + \frac{c}{n_2}}}$ P-værdi = t Distribution (f, f_1) ◆

Hvis nulhypotesen er H: $\mu_1 > \mu_2 + d$ må man beregne 1- den i programmet fundne P-værdi.

Test af Varians

Som det ses af eksempel 5.1 kan det sjældent blive nødvendig 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 5.3 har SAS.JMP et program til testningen, så det er ikke nødvendigt at anvende teststørrelsen direkte.

Eksempel 5.3 . 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:

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

Analyze ► Fit y by x ► Response: Tidsforbrug , Factor: Metode ► OK

Cursor på rød pil ► Un-Equal Variance

Blandt en række udskrifter forekommer nedenstående

Tests that the Variances are Equal

Level	Count	Std Dev	MeanAbsDif to Mean	MeanAbsDif to Median
m1	10	1,668965	1,412000	1,370000
m2	10	1,302135	0,960000	0,880000

Test	F Ratio	DFNum	DFDen	p-Value
O'Brien[.5]	0,7193	1	18	0,4075
Brown-Forsythe	1,2610	1	18	0,2762
Levene	1,6453	1	18	0,2159
Bartlett	0,5199	1	.	0,4709
F Test 2-sided	1,6428	9	9	0,4711

I udskrift for F - test ses, at alle -værdier er større end 0.05.

Følgelig accepteres H_0 , dvs. vi vil i en følgende test antage, at spredningerne er ens.

(selv om der kun var en enkelt der var over 0.05 ville man nok stadig antage, at spredningerne er ens) ◆

Dimensionering.**Eksempel 5.4 (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 $P(\text{fejl af type I}) = \alpha \leq 0.05$,

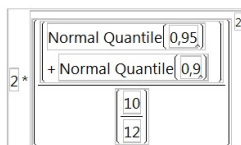
$P(\text{fejl af type II}) = \beta \leq 0.10$ og bagatelgrænsen $\Delta = 10$ minutter, idet man fra mange tilsvarende forsøg ved, at den fælles spredning er $\sigma = 12$ minutter.

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.

Løsning:

a) Af formlerne i oversigten i kapitel 5.5.1 fås
$$n \geq 2 \cdot \left(\frac{U_{1-\alpha} + U_{1-\beta}}{\frac{\Delta}{\sigma}} \right)^2 = 2 \cdot \left(\frac{U_{0,95} + U_{0,90}}{\frac{10}{12}} \right)^2 = 2 \cdot \left(\frac{1,645 + 1,282}{\frac{10}{12}} \right)^2$$

SAS.JMP: Kald søjle for n ► Placer cursor i n's hoved ► højre musetast ► Formula ► vælg formler som vist



= 24.66 dvs. der skal udføres i alt $n = 25$ delforsøg af hver behandling

Spørgsmål b)

Her kræves løst en kompliceret ligning (se igen oversigten), så man vil sædvanligvis sige man benytter svaret i spørgsmål a) og hvis de er under 30 forsøg, så siger man blot det skal benyttes 30 forsøg, da man så kan antage spredninger er rimelig eksakt. ◆

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

X_{gammel} = grad af slid ved anvendelse af sædvanlig lak B

X_{ny} = grad af slid ved anvendelse af ny lak A

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 $H : \mu > 0$.

Data indtastes

Ny lak	Gammel lak
20,3	19,5
25,1	28,4
21,8	21,6
19,6	22
18,9	20,9
23,5	25,8

Vælg Analyze ► Specialized Modeling ► Matched Pairs ► Y: Paired response: Ny lak og Gammel lak ► OK
► Rød pil: Fjern markering ved "Plot Diff by Mean"

Der fremkommer følgende tabel:

Matched Pairs			
Difference: Gammel lak-Ny lak			
Gammel lak	23,0333	t-Ratio	2,268219
Ny lak	21,5333	DF	5
Mean Difference	1,5	Prob > t	0,0726
Std Error	0,66131	Prob > t	0,0363*
Upper 95%	3,19996	Prob < t	0,9637
Lower 95%	-0,2		
N	6		
Correlation	0,89502		

Da gammel lak - ny lak = D og $H: \mu > 0$ så ses på Prob > t

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 formlerne i oversigten i afsnit 5.5.3 bygger på

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.

5. 1 Faktor på 2 niveauer

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) $H_0: p_1 - p_2 = 0$ mod $H: p_1 - p_2 < 0$ (da vi ønsker at vise, at $p_1 < p_2$).

Vi anvender nu formlerne i oversigt 5.5.3 side 56.

Vi ønsker at approksimere med normalfordeling.

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

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

er forudsætningerne for at approksimere med normalfordelingen opfyldt.

$$\text{Den har middelværdi 0 og spredning } s = \sqrt{p \cdot (1 - p) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{\frac{48}{279} \cdot \left(1 - \frac{48}{279}\right) \cdot \left(\frac{1}{139} + \frac{1}{140}\right)}$$

$$\text{P-værdi} = \text{Normal Distribution} \left(\frac{17}{139} - \frac{31}{140}, 0, \sqrt{\frac{48}{279} \cdot \left(1 - \frac{48}{279}\right) \cdot \left(\frac{1}{139} + \frac{1}{140}\right)} \right) = 0.01414 = 1.41\%$$

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

Konklusion:

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

2) 95% konfidensinterval;

Oversigterne er igen hentet fra oversigt 5.5.3

$$\text{Radius } r = \text{Normal Quantile}(0.975) \cdot \sqrt{\frac{\frac{17}{139} \cdot \left(1 - \frac{17}{139}\right)}{139} + \frac{\frac{31}{140} \cdot \left(1 - \frac{31}{140}\right)}{140}} = 0.08773$$

$$p_1 - p_2 = \frac{17}{139} - \frac{31}{140} = -0.0991$$

Konfidensinterval for $p_1 - p_2$: $[-0.0991 - 0.08773; -0.0991 + 0.08773] = \underline{[-0.187; -0.0111]}$
 eller for $p_2 - p_1$ $\underline{[0.011; 0.187]}$ ◆

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 :

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

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

er forudsætningerne for at approksimere med normalfordelingen opfyldt.

$$s = \sqrt{\frac{19}{39} \cdot \left(\frac{1}{20} + \frac{1}{19} \right)} \text{ og middelværdi } 0$$

Da $\bar{x}_1 > \bar{x}_2$ er P-værdi =  = 0.1502

Da P-værdi = 0.1502 > 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 , hvor f er det nærmeste hele tal, som er større end g .				
Forudsætninger	Alternativ hypotese H	P - værdi	Beregning	H_0 forkastes
σ_1, σ_2 ukendte $t = \frac{\bar{x}_1 - \bar{x}_2 - d}{\sqrt{c}}$	$\mu_1 > \mu_2 + d$	$P(T \geq t)$	1 - t Distribution(t, f) eller t-test: se eksempel 5.1	P - værdi $< \alpha$
	$\mu_1 < \mu_2 + d$	$P(T \leq t)$	t Distribution(t, f) eller t-test: 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 $< \frac{1}{2} \alpha$
100 · (1 - α)% konfidensinterval for differens $\mu_1 - \mu_2$: $\bar{x}_1 - \bar{x}_2 - t_{Quantile(1 - \frac{\alpha}{2}, f)} \cdot \sqrt{c} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{Quantile(1 - \frac{\alpha}{2}, f)} \cdot \sqrt{c}$				

Z-test (eller U-test): Forudsætning: σ_1 og σ_2 kendt eksakt , eller n_1 og n_2 er større end 30

Forkortelser: $\bar{x} = \bar{x}_1 - \bar{x}_2 - d$ $\sigma = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$				
Forudsætninger	Alternativ hypotese H	P - værdi	Beregning	H_0 forkastes
σ_1, σ_2 kendte	$\mu_1 > \mu_2 + d$	$P(Y \geq \bar{x})$	1-NormalDistribution($\bar{x}, 0, \sigma$)	P - værdi $< \alpha$
	$\mu_1 < \mu_2 + d$	$P(Y \leq \bar{x})$	NormalDistribution($\bar{x}, 0, \sigma$)	
	$\mu_1 \neq \mu_2 + d$	$P(Y \geq \bar{x})$ 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 $< \frac{1}{2} \alpha$
100 · (1 - α)% konfidensinterval for differens $\mu_1 - \mu_2$: $\bar{x}_1 - \bar{x}_2 - Normal\ Quantil\left(1 - \frac{\alpha}{2}\right) \cdot \sigma \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + Normal\ Quantil\left(1 - \frac{\alpha}{2}\right) \cdot \sigma$				

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.	$H: \mu > \mu_0$	$P(T \geq t)$	1 - t Distribution(t, f) eller Analyze, Matched pairs	P - værdi $< \alpha$
	$H: \mu < \mu_0$	$P(T \leq t)$	t Distribution(t, f) eller Analyze, Matched pairs	
	$t = \frac{(\bar{x} - \mu_0) \cdot \sqrt{n}}{s}$	$H: \mu \neq \mu_0$	$P(T \geq t)$ for $\bar{x} > \mu_0$ $P(T \leq t)$ for $\bar{x} \leq \mu_0$	som række 1 som række 2

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 størrelsen 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.
Se eksempel 5.3.

5. 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 se eksempel 5.3	P - værdi $< \alpha$
	$H: \sigma_1^2 < \sigma_2^2$	$P(Q \leq F)$	FCdf($-\infty, F, n_1 - 1, n_2 - 1$) eller se eksempel 5.3	
	$H: \sigma_1^2 \neq \sigma_2^2$	$P(Q \geq F)$ 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 $< \alpha$
100 · (1 - α)% Konfidensinterval for forhold $\frac{\sigma_1^2}{\sigma_2^2}$: $\frac{F}{F_{1-\frac{\alpha}{2}}(n_1 - 1, n_2 - 1)} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq F \cdot F_{\frac{\alpha}{2}}(n_2 - 1, n_1 - 1)$				

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 $F(f_T, f_N)$.

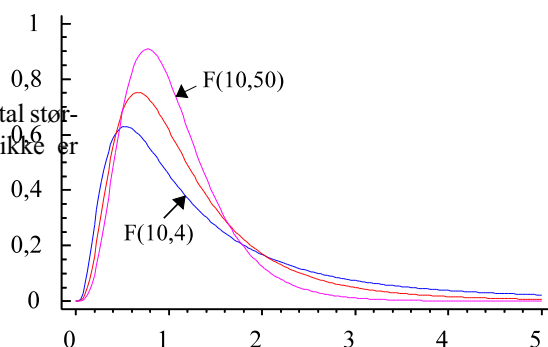
Det kan vises (se evt "afsnit til statistiske grundbegreber" 3C), at den variable $F = \frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}}$ er F - fordelt $F(f_1, f_2)$, hvor

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.



Numerator d.f., Denominator d.f.
 — 10,4
 — 10,10
 — 10,50

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
$n_1 \cdot \hat{p} \in [5; n_1 - 5]$ $n_2 \cdot \hat{p} \in [5; n_2 - 5]$	$H: p_1 > p_2$	$P(Y > \hat{p}_1 - \hat{p}_2)$	$1 - NormalDistribution(\hat{p}_1 - \hat{p}_2, 0, s)$	P - værdi $< \alpha$
	$H: p_1 < p_2$	$P(Y < \hat{p}_1 - \hat{p}_2)$	$NormalDistribution(\hat{p}_1 - \hat{p}_2, 0, s)$ se eksempel 5.6	
	$H: p_1 \neq p_2$	$P(Y > \hat{p}_1 - \hat{p}_2)$ 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 $< \frac{1}{2} \alpha$
100 · (1 - α)% konfidensinterval for differens $p_1 - p_2$ $\hat{p}_1 - \hat{p}_2 - NormalQuantile\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}} \leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + NormalQuantile\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$				

5.5.4 Test og konfidensinterval af $\mu_1 - \mu_2$ for 2 Poissonfordelte 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 $n(0, s)$ se evt. eksempel 5.7				
Forudsætninger	Alternativ hypotese H	P - værdi	Beregning	H_0 forkastes
$n_1 \cdot \bar{x} \geq 5$ $n_2 \cdot \bar{x} \geq 5$	$H: \mu_1 > \mu_2$	$P(Y > \bar{x}_1 - \bar{x}_2)$	$1 - NormalDistribution(\bar{x}_1 - \bar{x}_2, 0, s)$	P - værdi $< \alpha$
	$H: \mu_1 < \mu_2$	$P(Y < \bar{x}_1 - \bar{x}_2)$	$NormalDistribution(\bar{x}_1 - \bar{x}_2, 0, s)$	
	$H: \mu_1 \neq \mu_2$	$P(Y > \bar{x}_1 - \bar{x}_2)$ 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 $\mu_1 - \mu_2$. $\bar{x}_1 - \bar{x}_2 - NormalQuantile\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{\bar{x}_1}{n_1} + \frac{\bar{x}_2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + NormalQuantile\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{\bar{x}_1}{n_1} + \frac{\bar{x}_2}{n_2}}$				

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 diætenes 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 standarddosering 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

- og man ved, at spredningen $\sigma = 0.2$ ohm.
- og man har en forhåndsformodning om, at spredningen er ca. $\sigma = 0.2$ ohm.
- Hvilke konklusioner vedrørende behandlingernes virkning kan gøres, såfremt man ved testning af forsøgsresultaterne finder
 - signifikans
 - ingen signifikans
- 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 svovlmonooxid. 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%.

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 5 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 naturligvis 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.

SAS.JMP-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 (forklaring af metode og formler)

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: 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.

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. kapitel 10).

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

Pooling.

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

$$s_{\text{error}}^2 = \frac{s_1^2 + s_2^2 + s_3^2 + s_4^2}{4} = \frac{2.000^2 + 2.6458^2 + 2.5166^2 + 3.6056^2}{4} = 7583$$

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.

6. En faktor på mere end 2 niveauer

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 et regnemiddel findes $s_x^2 = 12.552$.

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

$$s_R^2 = 3 \cdot s_x^2 = 3 \cdot 12.552 = 37.6567. \text{ Frihedsgradstallet er } f_R = \text{antal niveauer} - 1 = 4 - 1 = 3.$$

Hvis nulhypotesen er sand burde $s_R^2 \approx s_{\text{error}}^2$ eller $\frac{s_R^2}{s_{\text{error}}^2} \approx 1$, mens hvis nulhypotesen er falsk (middelværdierne er forskellige) er $s_R^2 > s_{\text{error}}^2$ dvs. forholdet $F = \frac{s_R^2}{s_{\text{error}}^2}$ være signifikant større end 1.

Da $F = \frac{s_R^2}{s_{\text{error}}^2} = \frac{37.6567}{7.583} = 4.9656$ 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 $f_{\text{error}} = 8$ i nævneren kan vi afgøre dette ved at regne P - værdien ud.

P - værdi = $P(F > 4.9656) = 1 - P(F < 4.9656) =$

$$1 - \text{FDistribution}(4.9656, 3, 8) = 0.0311 = 3.11\%.$$

Da P - værdi = $0.031 < 0.05$ 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.

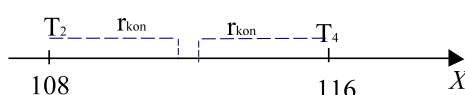
De "sædvanlige" konfidensintervaller for hvert niveau bestemmes ved (jævnfør eventuelt appendix 6.1):

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

$$\text{I vort tilfælde er } r_{kon} = t_{0.975}(8) \sqrt{\frac{7.583}{3}} = \text{tQuantile}(0.975, 8) \cdot \sqrt{\frac{7.583}{3}} = 3.667$$

og konfidensintervallet $\bar{x}_i \pm r_{kon} = \bar{x}_i \pm 3.667$

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 $2 \cdot r_{kon} = 2 \cdot 3.673 = 7.332$ 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 5.

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

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

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

Da $\sqrt{2} \cdot 3.666 = 5.19$ 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 $[108 - 3.666; 108 + 3.666] = [104.3 ; 111.7]$

6.2.3. Beregning af ensidet variansanalyse

Eksempel 6.2 (oprindelige data ikke kendt)

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. Gennemsnit og spredning er angivet i følgende skema

	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

Forudsætningerne for en variansanalyse antages 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 estimat for den mindste urenhed og 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".

Gennemsnit indtastes i tabel

Analyze ► Distribution ► marker gennemsnit og tryk på y, Columns ► ok

gennemsnit	
	110
	108
	113,33
	116

Man får en tabel, hvor man ser, at spredningen på gennemsnittene er 3.5429119

Summary Statistics	
Mean	111,8325
Std Dev	3,5429119
Std Err Mean	1,771456
Upper 95% Mean	117,47006
Lower 95% Mean	106,19494
N	4

Da gennemsnit har 3 gange så lille varians som de oprindelige tal multipliceres varians med 3

$$\text{variens-gennemsnit} = 3 \cdot 3,5429119^2 = 37,6567$$

$$\text{Varians} = \frac{(2^2 + 2,6458^2 + 2,5166^2 + 3,6056^2)}{4} = 7,8347 \quad \text{F-værdi} = \frac{\text{variens-gen}}{\text{variens}} = 4,9656$$

$$P\text{-værdi} = P(F > 4,9656) = 1 - P(F < 4,9656) = 1 - F\text{Distribution}(F, 3, 8) = 0,0311$$

Resultaterne samlet i skemaet

variens-gen	variens	F	P-værdi
37,656674193	7,58347114	4,9656250414	0,0311056874

Da $P\text{-værdi} = 0,0311 < 0,05$ forkastes nulhypotesen $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

Konklusion: De fire producenteres batterier har ikke samme virkning.

Mindst 2 af producenteres batterier er forskellige.

- I skemaet er nu beregnet radius i konfidensinterval og LDD-interval

variens-gen	variens	F	P-værdi	r	r-lsd
37,656674193	7,58347114	4,9656250414	0,0311056874	3,6663462179	2,5924982728

$$r = r_{kon} = t_{1-\frac{\alpha}{2}}(f_{error}) \cdot \sqrt{\frac{s_{error}^2}{n}} = \text{tQuantile}(0,975, 8) \cdot \sqrt{\frac{\text{varians}}{3}}$$

Da $2 \cdot r = 7.332$ ses, at T_2, T_1 og T_3 ikke kan adskilles (de giver begge mindst urenhed)

Benyttes LSD, ses, at nu er det T_2 og T_1 der giver mindst urenhed.

Et 95% konfidensinterval for T_2 er $[108 - 3.667; 108 + 3.667] = [104.3; 111.7]$

Eksempel 6.3 (eksempel 6.1 fortsat)

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

- Kontroller om betingelserne om normalitet og varianshomogenitet er rimelig 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 estimat for den mindste urenhed og et 95% konfidensinterval for middelværdien.

Løsning:

Data indtastes, "tilsætningsstof" skal være "rød" (se under Columns) og urenhed skal være "blå".

tilsætningsstof	urenhed
t1	108
t1	110
t1	112
t2	105
t2	110
t2	109
t3	116
osv	

a) Kontrol af forudsætninger:

Selv om det er muligt i SAS.JMP at kontrollere forudsætningen om normalfordeling, behøver man ikke nødvendigvis at gøre det, da testen er robust overfor afvigelse.

Normalfordelingsplot:

Vælg Analyze ► Fit Y by X ► markér "Urenhed" og tryk på Y Response ► markér "Tilsætningsstof" og tryk på X Factor ► OK

Der fremkommer et "scatterplot", hvor man kan se afbildet de tre værdier af hver af de 4 tilsætningsstoffer

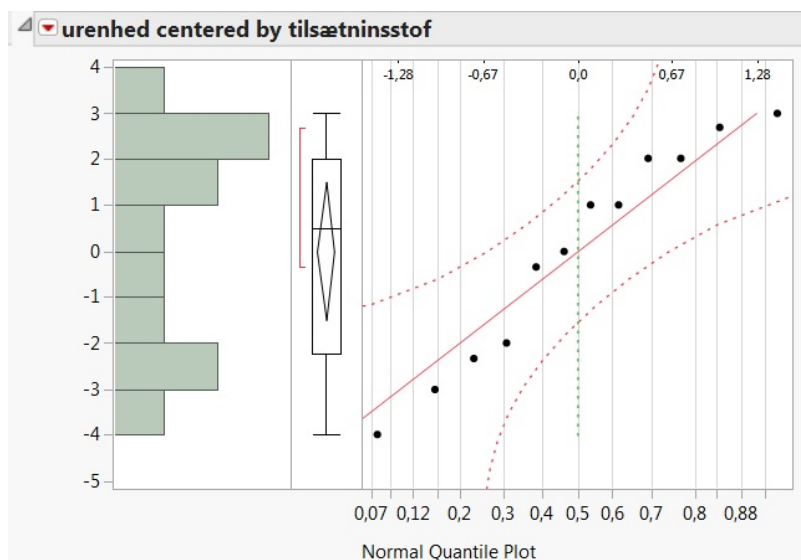
rød pil ved figur ► rullemenu Save ► Save Residual

Under data kommer nu en ekstra søjle med residualerne (overskrift "urenhed centered by tilsætningsstof").

Vælg Analyze ► Distribution ► Indsæt Residual søjlen i Y ► OK ► Cursor på rød pil ► "normal Quantile plot"

Der fremkommer et histogram med indtegnet normalfordelingskurve

6. En faktor på mere end 2 niveauer



Da vi jo kun har 12 residualer kan vi ikke forvente et særligt klart billede, men i princippet burde de på normalfordelingsplottet ligge nogenlunde på en ret linie. De ligger i hvert fald alle indenfor konfidens-grænserne, så vi kan tillade os at antage fordelingen er approksimativt normalfordelt.

Kontrol af varianshomogenitet:

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

X_i antages approksimativt normalfordelt med middelværdien μ_i og spredning σ_i .

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

Sæt cursor på rød pil ved tegningen for Scatterplot, og vælg fra rullemenuen "Unequal Variances".

Tests that the Variances are Equal

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	0,3812	3	8	0,7694
Brown-Forsythe	0,4211	3	8	0,7430
Levene	0,5917	3	8	0,6376
Bartlett	0,1991	3	.	0,8971

Warning: Small sample sizes. Use Caution.

Da P - værdierne alle er over 0.05 accepteres nulhypotesen H_0 : De 4 varianser er ens.

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

Sæt cursor på rød pil ved scatterplot, og vælg fra rullemenuen "Mens/Anova".

Der fremkommer så en tegning og følgende udskrift:

Oneway Analysis of urenhed By tilsætningsstof

Oneway Anova
Summary of Fit

Rsquare	0,650672
Adj Rsquare	0,519674
Root Mean Square Error	2,753785
Mean of Response	111,8333
Observations (or Sum Wgts)	12

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
tilsætningsstof	3	113,00000	37,6667	4,9670	0,0311*
Error	8	60,66667	7,5833		
C. Total	11	173,66667			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
t1	3	110,000	1,5899	106,33	113,67
t2	3	108,000	1,5899	104,33	111,67
t3	3	113,333	1,5899	109,67	117,00
t4	3	116,000	1,5899	112,33	119,67

Af variansanalysetabellen fremgår, at P -værdi ($\text{Prob}>F$) = 0.0311 < 0.05, dvs. nulhypotesen $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ forkastes (svagt)

Konklusion: De fire tilsætningsstoffer har ikke samme virkning.
Mindst 2 af middelurenhederne er forskellige.

Forklaring på de enkelte størrelser i variansanalysetabellen kan man finde i appendix 4.1.

- c) Konfidensintervallerne ses i ovenstående udskrift, og umiddelbart ses, at der er et overlap mellem den mindste værdi t2 og t1 og t3.

For at få et lidt bedre skøn, ses på LSD-intervaller.

Tegningen der fremkom skal forstås på følgende måde:

“Diamanterne” angiver 95% konfidensintervaller.

Den midterste vandrette streg angiver gennemsnittet og de to andre vandrette streger angiver LSD -intervaller.

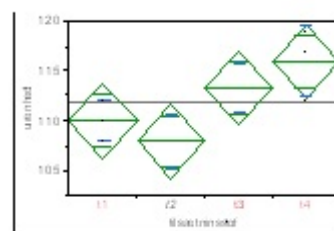
Af figuren ses derfor straks, ved at se på LSD intervallerne, at t2 er signifikant mindre end t4, mens det er vanskeligere at se om t2 og t3 kan adskilles.

LSD-intervaller fås ved fra rullemenuen at vælge

“Compare Means” ► Each Pair, Students t-tast

Blandt en række udskrifter findes denne

Level	Mean
t4 A	116,00000
t3 A B	113,33333
t1 B C	110,00000
t2 C	108,00000



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

Af den forrige udskrift ses, at et estimat for mindsteværdi er 108.0
og et 95% konfidensinterval er [104.3 ; 111.7] ◆

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 inddele 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

6. En faktor på mere end 2 niveauer

((stor “støj”). Dette kan bevirke, 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 bevirke, at forsøgsfejls spredning bliver så stor, at intet kan påvises (forsøget drukner i støj). Da grise fra samme kuld må forventes at være mere ensartede, vælger man at lave et randomiseret blokforsøg med kuld som blokfaktor.

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.

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”.

Indtastning af data

	Fodertype	Kuld	Vægt
1	A	k1	7
2	A	k2	16
3	A	k3	10.5
4	A	k4	13.5
5	B	k1	14
6	B	k2	15.5
7	B	k3	15
8	B	k4	21
9	C	k1	8.5
10	C	k2	16.5
11	C	k3	9.5
12	C	k4	13.5

Vælg Analyze ► Fit Y by X ► markér “Vægt” og tryk på Y Response ► markér “Fodertype” og tryk på X Factor ► Marker “Kuld” og tryk på blok ► OK

Der fremkommer et “scatterplot”, hvor man kan se afbildet de fire værdier for hver af de 3 fodertyper.

Rød pil ► “Means/Anova”.

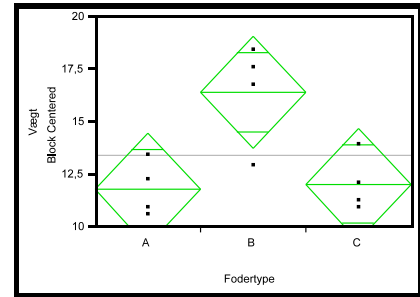
Der fremkommer så følgende tegning og udskrift:

Oneway Analysis of Vægt By Fodertype

Oneway Anova

Summary of Fit

Rsquare	0,83413
Adj Rsquare	0,695904
Root Mean Square Error	2,168269
Mean of Response	13,375
Observations (or Sum Wgts)	12



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Fodertype	2	54,12500	27,0625	5,7563	0,0402
Kuld	3	87,72917	29,2431	6,2201	0,0285
Error	6	28,20833	4,7014		
C. Total	11	170,06250			

Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
A	4	11,7500	1,0841	9,097	14,403
B	4	16,3750	1,0841	13,722	19,028
C	4	12,0000	1,0841	9,347	14,653

Std Error uses a pooled estimate of error variance

Block Means

Kuld	Mean	Number
k1	9,8333	3
k2	16,0000	3
k3	11,6667	3
k4	16,0000	3

Da P - værdi for fodertype er $0.0402 < 0.05$ forkastes nulhypotesen H_0 : Ingen forskel på fodertyper, dvs. , at der på et signifikansniveau på 5 % er signifikant forskel på fodertyperne (mindst én afviger fra de øvrige).

Vi ser endvidere, at det var fornuftigt at dele op i kuld, da der også er signifikans for kuld.

Vi er imidlertid ikke interesseret i at finde ud af hvilket kuld der er det bedste, da vi jo blot har taget nogle tilfældige kuld ud. Konfidensintervallerne viser et svagt overlap.

Vælger nu LSD-intervaller

“Compare Means” ► each pair students t-test

Comparisons for each pair using Student's t

Confidence Quantile

t	Alpha
2,44691	0,05

Connecting Letters Report

Level	Mean
B A	16,375000
C B	12,000000
A B	11,750000

Levels not connected by same letter are significantly different.

Af LSD-intervallerne ses, at fodertype B giver den største vægtforøgelse.



6. En faktor på mere end 2 niveauer

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-værdi = 0.0402 < 0.05

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

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

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 $f = n - 1 = 7$

$$P\text{-værdi} = P(\chi^2 > 14.67) = 1 - \text{ChiSquare Distribution}(14.67, 7) = 0.0405$$

Da P-værdi < 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 kanyle til udtagning af blodprøven ankommer de hvide blodlegemer "tilfældigt" i tiden. Det mulige antal blodlegemer er næsten ubegrænset.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$$

Antal elementer i hver stikprøve er 1, dvs. i 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 = $P(\chi^2 > 7.1) = 1 - \text{ChiSquare Distribution}(7, 1, 5) = 0.2133 > 0.05$ 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. OVERSIGT 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".

Data indtastes i 2 lister, med overskrift A og observationer.

Analyze ► **Fit Y by** ► **X** ► marker observationer og tryk på Y Response ► marker A og tryk på X Factor ► OK

Der fremkommer bl.a. et scatterplot.

Sæt Cursor på rød pil ved scatterplot og vælg fra rullemenu Means/Anova

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

Konfidensintervaller

I Means/Anova ses konfidensintervaller

LSD - intervaller fås ved at vælge Compare Means

Er $|\bar{x}_1 - \bar{x}_2| > d_{LSD}$ hvor d_{LSD} er det dobbelte af radius i LSD-intervallerne antages niveauerne

A₁ og A₂ 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}

Indtaster og ordrer som før, men marker overskrift "blokke" og tryk på blok

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

Konfidensintervaller

Samme ordrer som før

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:	Nullhypotese	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 $n_k \cdot \hat{p} \in [5; n_k - 5]$ 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	Nullhypotese	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 - værdi $< \alpha$ hvor P - 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
B2	3200	3300	2975	3250
B3	2800	2900	2985	3050
B4	2600	2700	2600	2765

- Kontroller om betingelserne om normalitet og varianshomogenitet er rimelig opfyldt.
- Undersøg på et signifikansniveau på 5% om forskellen i blandingsteknik har betydning for trykstyrken.
- Find i bekræftende fald den (de) blandingsteknik(er) der har størst trykstyrke, og angivet 95% konfidensinterval for middelværdien.

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	4	5	6
0.76	0.84	0.83	1.00	0.88	0.86
0.82	0.79	0.78	0.90	0.90	0.89
0.79	0.82	0.97	0.92	1.03	0.87
0.86	0.79	0.88	0.88	0.87	0.84

Analysens forudsætninger om varians og normalitet antages opfyldt.

Vurdér på grundlag af dette materiale en antagelse om, at alkoholkoncentrationen i blodet ikke afhænger af andre faktorer end den indtagne alkoholemængde.

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

Undersøg om det kan antages, at de 5 spolers modstande er ens.

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 studentereksamensgennemsnit.

Navn	JK	AL	TS	BS	DT	HN	MO	FD	PJ	KM	SR	RA
Snit	6.3	6.8	7.3	7.3	7.9	8.2	8.4	8.5	9.0	10.2	11.1	11.2

- 1) Hvordan ville du opdele disse studenter på tre hold med 4 på hver hold?
- 2) Hvordan ville du gøre det, hvis karaktererne gik fra 7.8 til 8.2 ?

Opgave 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 benzinforsøget.



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 En 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 benzinforbruget. Derefter varieres den anden faktor (her karburator), mens olieblandingen fastholdes på O_2 , dvs. de sidste 4 biler udstyres med karburator K_2 og olieblanding O_2 . Igen gennemkøres den udvalgte strækning, og benzinforbruget måles. Indtegnes for hver karburator det gennemsnitlige benzinforbrug mod olie-blandingen kunne vi eksempelvis få tegningen på fig. 7.1.

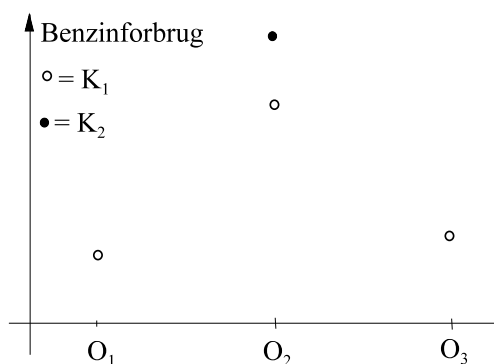


Fig 7.1 Skitse af benzinforbrug

Umiddelbart ses, at K_1 giver lavest benzinforbrug, og O_1 (eller O_3) skal foretrækkes. Hvad nu med benzinforbruget i karburator K_2 , hvis vi anvender olieblanding O_1 eller O_3 ? Kan man slutte, at benzinforbruget ved olieblanding 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
Olieblanding	O_1	× ×	× ×
	O_2	× ×	× ×
	O_3	× ×	× ×

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

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.

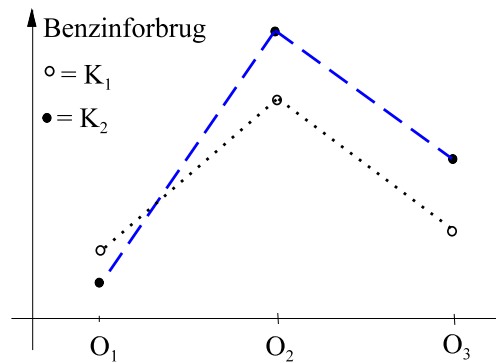


Fig 7.2. Vekselvirkning

Vi ser, i modstrid med hvad vi antog ud fra "en faktor ad gangen forsøget", at kombinationen af katalysator K_2 og olieblending O_1 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 O_1 , O_2 , og O_3 , og 2 karburatorer K_1 og K_2 påvirker benzinforgbruget.

Forsøgsresultaterne er følgende:

		Karburator			
		K_1		K_2	
Olieblending	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 følgelig $N = r \cdot q \cdot n$.

I eksempel 7.4 er $R = \text{"olieblending"}$, $r = 3$, $C = \text{"karburator"}$, $q = 2$, $n = 2$ 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 (pooled) 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.

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 faktoreres 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 SS=SAK'erne er temmelig omfattende, og næppe giver en dybere forståelse, vil vi dog sædvanligvis benytte programmer som findes i SAS.JMP hertil. Programmerne dækker dog ikke alle muligheder, så i enkelte tilfælde må man derfor benytte formlerne for eksempelvis konfidensintervaller.

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

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

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

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

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

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

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

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

I praksis vil dette naturligvis aldrig være tilfældet på grund af den tilfældige variation (støj), men udregnes kvadratet på afvigelse (SAK), og disse afvigelser ikke er større end hvad er rimeligt i forhold til støjen (s_0), vil vi kunne konkludere at der ikke kan konstateres nogen vekselvirkning.

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 SAS.JMP.

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 O_1 , O_2 , og O_3 , og 2 karburatorer K_1 og K_2 påvirker benzinforsøget.

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

- Kontroller om betingelserne for normalitet og varianshomogenitet er rimelig opfyldt
- Angiv hvilke kombinationer af karburator og olieblending der giver det laveste forbrug, og giv et estimat og et konfidensinterval for dette forbrug.

Løsning:

Data indtastes på sædvanlig måde .

Karburator	olieblanding	benzinforbrug
k1	o1	830
k1	o1	860
k1	o2	940
k1	o2	990
k1	o3	855
k1	o3	815
k2	o1	810 osv.

a) Kontrol af normalfordeling

For at kunne beregne residualer fra hver celle dannes en ekstra søjle, som angiver cellenummer.

Karburator	Olieblanding	Benzinforbrug	celler
k1	o1	830	c1
k1	o1	860	c1
k1	o2	940	c2
k1	o2	990	c2
k1	o3	855	c3
k1	o3	815	c3
k2	o1	810	c4
k2	o1	840	c4
k2	o2	1050	c5
k2	o2	1020	c5
k2	o3	930	c6
k2	o3	910	c6

Vælg Analyze ► Fit Y by X ► markér “Benzinforbrug” og tryk på Y Response ► markér “celler” og tryk på X Factor ► OK

Der fremkommer et “scatterplot”

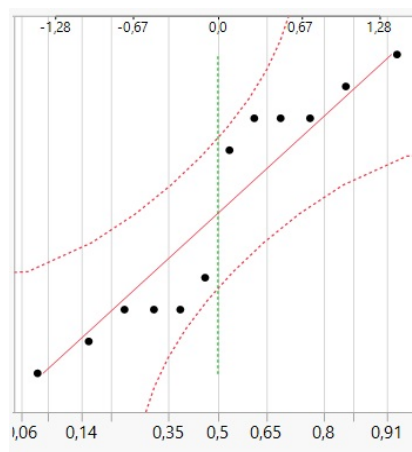
rød pil ved figur ► rullemenu Save ► Save Residual

Under data kommer nu en ekstra søjle med residualerne(overskrift “urenhed centered by tilsætningsstof”).

Vælg Analyze ► Distribution ► Indsæt Residual søjlen i Y ► OK ► Cursor på rød pil ► “normal Quantile plot”

Der fremkommer et histogram med indtegnet normalfordelingskurve

Da vi jo kun har 12 residualer kan vi ikke forvente et særligt klart billede, men i princippet burde de på normalfordelings-plottet ligge nogenlunde på en ret linie. De ligger i hvert fald alle indenfor konfidensgrænserne, så vi kan tillade os at antage fordelingen er approksimativt normalfordelt.

**Varianshomogenitet.**

Lad varianserne i de 6 celler være $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5$ og σ_6 .

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2 = \sigma_6^2$ H: Mindst en varians er forskellig fra en af de øvrige

Da varianserne i hver af 6 celler skal være ens, karakteriserer man disse ved at vælge “celler”

Vælg Analyze ► Fit Y by X ► markér “benzinforbrug” og tryk på Y Response ► markér “celler” og tryk på X Factor ► OK cursor på rød pil på tegning og vælg “Unequal Variances”

Oneway Analysis of Benzinförbrug By Celler

Tests that the Variances are Equal

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	0,0000	-1	0	0,0000*
Brown-Forsythe	.	5	6	.
Levene	.	5	6	.
Bartlett	0,1354	5	.	0,9842

Warning: Small sample sizes. Use Caution.

Da Bartletts test giver en P -værdi = 0.9642 > 0.05 fås en accept af nulhypotesen.

Kravet er rimeligt opfyldt.

b) Først testes $H_0: R * C = 0$ mod $H: R * C \neq 0$

Vælg Analyze ► Fit Model ► Indsæt for Y: benzinförbrug ,Indsæt for ADD: karburator og olieblanding ► Indsæt for CROSS: karburator, olieblanding (marker begge) ► Emphasis's rullemenu: vælg Minimal Report ► Run

(Minimal report er kun valgt for at undgå nogle i denne forbindelse overflødige figurer)

Man får (blandt andet)

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
karburator	1	1	6075,000	10,1250	0,0190
Olieblanding	2	2	58716,667	48,9306	0,0002
karburator*Olieblanding	2	2	6450,000	5,3750	0,0460

Ud for "Karburator*olieblanding" findes P - værdi = 0.0460.

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

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

For at finde hvilke kombinationer der giver lavest benzinförbrug vælges

"Effect Details" ► rød pil ved "karburator x temperatur" ► Vælg "LS means Plot"

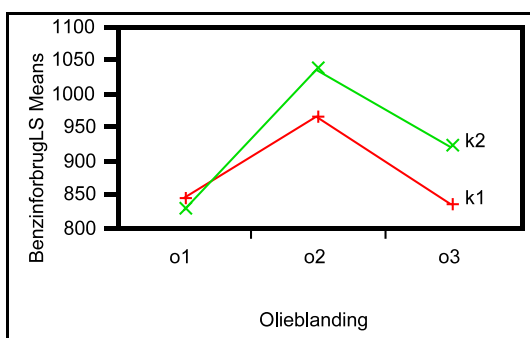
Vi får følgende Udskrift+tegning :

Karburator*Olieblanding

Least Squares Means Table

Level	Least Sq Mean	Std Error
k1,o1	845,0000	17,320508
k1,o2	965,0000	17,320508
k1,o3	835,0000	17,320508
k2,o1	825,0000	17,320508
k2,o2	1035,0000	17,320508
k2,o3	920,0000	17,320508

LS Means Plot



Umiddelbart ses af figuren, at man ikke bør vælge olieblending O_2 .

Derimod er det uklart hvilken af kombinationer (se tabellen) med de mindste “means”, der giver det laveste olieforbrug.

Dette kan afklares ved på ovennævnte rullemenu at vælge “LSMeans students t”

Det giver en stor tabel (som kan fjernes ved med cursor på overskrift, højre musetast at fjerne markeringen ved “Crosstab Report”).

Under den findes følgende lille tabel

Level		Least Sq Mean
k2,o2	A	1035,0000
k1,o2	B	965,0000
k2,o3	B	920,0000
k1,o1	C	845,0000
k1,o3	C	835,0000
k2,o1	C	825,0000

Levels not connected by same letter are significantly different

Heraf ses, at kombinationen $K_2 O_1$ giver det laveste benzinformbrug (825), men, at der ingen signifikant forskel er mellem $K_2 O_1$, $K_1 O_3$ og $K_1 O_1$.

Konfidensintervaller

Ønskes fundet 95% konfidensintervaller

musetast i skemaet ovenfor ► tryk højre musetast ► Vælg “Columns” ► vælg “lower” og derefter “upper”.

		Least		
Level		Sq Mean	Lower 95%	Upper 95%
k2,o2	A	1035,0000	992,62	1077,4
k1,o2	B	965,0000	922,62	1007,4
k2,o3	B	920,0000	877,62	962,4
k1,o1	C	845,0000	802,62	887,4
k1,o3	C	835,0000	792,62	877,4
k2,o1	C	825,0000	782,62	867,4

Levels not connected by same letter are significantly different.

Heraf ses, at for kombinationen $K_2 O_1$ er konfidensintervallet = [782.6 ; 867.4]



Eksempel 7.3 Additiv model: To signifikante hovedvirkninger

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											
		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.

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:

Lad starten af indtastningen i regnearket være

cementblandere	cementknusere	brudstyrke
c1	k1	147
c1	k1	175
c1	k1	130
c1	k2	99
c1	k2	85

osv.

1) Variansanalysetabel opstilles.

Vælg Analyze ► Fit Model ► Indsæt for Y: Brudstyrke ,Indsæt for AD:Cementblandere og Cementknusere ► Indsæt for CROSS: Cementblandere,Cementknusere (marker begge) ► Run

Der fremkommer en række figurer og tabeller.

Blandt disse er følgende tabel

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Cementblander	2	2	8706,056	4,9116	0,0163*
Cementknuser	3	3	51995,222	19,5557	<,0001*
Cementblander*Cementknuser	6	6	7122,611	1,3394	0,2787

a) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning) $H: R * C \neq 0$

For "Cementblandere*cementknusere" findes P -værdi = 0.2787.

Da P -værdi = 0.2787 > 0.05 accepteres H_0 .

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

b) Vekselvirkningen "pooles" ned i "error".

Marker under den øverste Source: Cementblander*Cementknusere ► Remove

Blandt mange tabeller findes

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Cementblander	2	2	8706,056	4,5994	0,0181*
Cementknuser	3	3	51995,222	18,3125	<,0001*

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

$H_0: R = 0$ (Cementblander har ingen virkning) forkastes, da P -værdi = 0.0181 < 0.05

Konklusion: Cementknuserne har en stærk signifikant virkning.
Cementblanderne har en signifikant virkning.

2) Under "cementblandere" kan på samme måde som i forrige eksempel bl.a. findes følgende tabeller

Level		Least Sq Mean
c2	A	140,08333
c1	B	110,66667
c3	B	104,41667

Levels not connected by same letter are significantly different.

Heraf ses, at cementblander 2 må foretrækkes

Under cementknusere fås

Level		Least Sq Mean
k1	A	151,33333
k4	A	150,22222
k2	B	114,11111
k3	C	57,88889

Levels not connected by same letter are significantly different

Cementknuser 1 og 4 må foretrækkes.

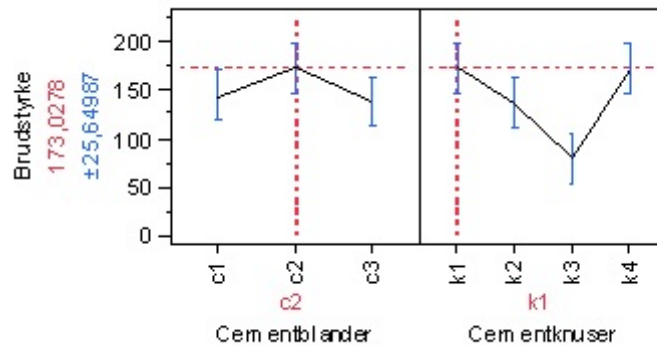
Konklusion: Størst middelbrudstyrke fås i kombinationen

cementknuser 1 og cementblander 2 eller cementknuser 4 og cementblander 2

Et estimat $\tilde{\mu}_{12}$ for største middelbrudstyrke på basis af cementknuser 1 og cementblander 2:

Rød pil ved Response Brudstyrke ► Factor profiling ► Profiler

Ved de fremkomne figurer flyttes linier



$$\tilde{\mu}_{12} = \underline{\underline{173.0}}$$

$$95\% \text{ Konfidensinterval : } [173.03 - 25.65; 173.03 + 25.65] = \underline{\underline{[147.4; 198.7]}}$$



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:

a) $H_0: R * C = 0$ (Ingen signifikant vekselvirkning) $H: R * C \neq 0$

Fremgangsmåden 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.

- b) Marker under Source: Cementblander*Cementkuserne ► Remove
Får nu

Effect Tests					
Source	Nparm	DF	Sum of		
			Squares	F Ratio	Prob > F
Cementblander	2	2	4299,389	2,2713	0,1206
Cementkuser	3	3	51995,222	18,3125	<,0001*

$H_0: C = 0$ (Cementkuser har ingen virkning) mod $H: C \neq 0$

Da P-værdi for C er under 0.0001 må $H_0: C = 0$ forkastes, dvs. Cementkuserne har en (stærk) virkning

$H_0: R = 0$ (Cementblander har ingen virkning) mod $H: R \neq 0$

$H_0: R = 0$ (Cementblander har ingen virkning) accepteres, da P-værdi = 0.1206 > 0.05

Konklusion: Cementkuserne har en (stærk) virkning
Cementblandere har ingen virkning

Marker under Source: Cementblander ► Remove

Under "effect details" ► rød pil ved "Cementkuser" ► vælges lsMean Students t-test

Det resulterer i følgende tabel:

Level		Least Sq Mean
k1	A	148,00000
k4	A	146,88889
k2	B	110,77778
k3	C	54,55556

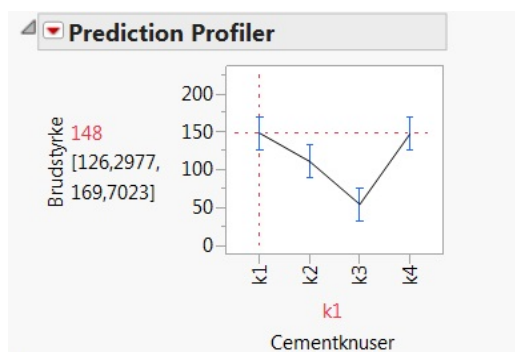
Levels not connected by same letter are significantly different.

Konklusion: Cementkuser 1 og 4 må foretrækkes

Konfidensinterval:

Rød pil ved Response Brudstyrke ► Factor profiling ► Profiler

Ved den fremkomne figur flyttes linie.



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 inddelle 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 karburatorer 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

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.

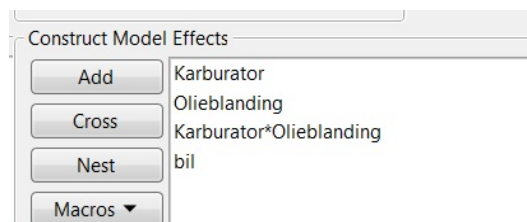
Lad den første seddel der trækkes være O_1K_2 . Det betyder nu, at bil 1 skal forsynes med karburator 2 og olieblending 1 og køre dag 1. Lad den næste seddel der trækkes være O_3K_2 . Det betyder tilsvarende at bil 1 skal forsynes med karburator 2 og olieblending 3 og køre dag 2. Således fortsættes indtil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

Derefter fortsættes med at trække sedler fra dåsen med bil 2.

- 2) Data indtastes

	Karburator	Olieblending	Benzinforsøg	bil
1	k1	o1	830	bil1
2	k1	o1	860	bil2
3	k1	o2	940	bil1
4	k1	o2	990	bil2
5	k1	o3	855	bil2
6	k1	o3	815	bil1
7	k2	o1	810	bil1
8	k2	o1	840	bil2
9	k2	o2	1050	bil2
10	k2	o2	1020	bil1
11	k2	o3	930	bil2
12	k2	o3	910	bil1

Vælg Analyze ► Fit Model ► Indsæt for Y: Benzinförbrug ,Indsæt for AD:Cementblandere og Cementknusere ► Indsæt for CROSS: Cementblandere,Cementknusere (marker begge) ► Indsæt for AD: bil ► Run



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 benzinförbrug end bil 2).

Der fremkommer bl.a. følgende udskrift

Source	Nparm	DF	Sum of		
			Squares	F Ratio	Prob > F
Karburator	1	1	6075,000	113,9062	0,0001*
Olieblanding	2	2	58716,667	550,4687	<,0001*
Karburator*Olieblanding	2	2	6450,000	60,4687	0,0003*
bil	1	1	3333,333	62,5000	0,0005*

Idet vi som sædvanlig antager at signifikansniveauet er 5 % fås, at da P -værdi = 0.03% < 5% forkastes H_0

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

P-værdien er væsentlig mindre end i eksempel 7.2, dvs. blokken (bilerne) har haft betydning

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 Andre varianter af variansanalyse

7.6.1 Manglende observationer (ubalancerede forsøg).

Vi har hidtil forudsat, at forsøgsplanen var et fuldstændigt faktorforsøg med lige mange gentageser af hver behandling. Dette vil i praksis ikke altid være tilfældet. Mislykkes således et delforsøg, opdager man det måske først så sent, at det ikke er muligt at foretage et nyt forsøg under de samme betingelser. Man har så et ikke-balanceret forsøg. Analysen foretages på samme måde, som hvis dataene var balancerede, men forsøget er nu mere følsomt overfor, om forudsætningen om varianshomogenitet er opfyldt.

7.6.2 Trinvis variansanalyse

Trinvis niveauer. At en faktor B er undertrin (engelsk: nested factor) til faktor A's niveauer, betyder at ethvert niveau af B kun kombineres med ét niveau af A.

Et eksempel på en trinvis variansanalyse er følgende

Eksempel 7.6 (trinvis variansanalyse)

I hvert af 5 tilfældigt valgte autoriserede reparationsværksteder udvalgte tilfældigt 2 mekanikere, og for hver af disse målte 3 gange den tid, som en bestemt reparation tog. Følgende resultater fandtes (kodede tal):

Værksteder							
1		2		3		4	
Mekaniker		Mekaniker		Mekaniker		Mekaniker	
1	2	3	4	5	6	7	8
6	1	10	4	10	7	5	0
2	10	7	1	11	12	10	8
0	0	12	9	6	15	8	6

værksted	mekaniker	tid
v1	m1	6
v1	m1	2
v1	m1	0
v1	m2	1
v1	m2	10
v1	m2	0
v2	m3	10
v2	m3	7
v2	m3	12
v2	m4	4
v2	m4	1
v2	m4	9
v3	m5	10
v3	m5	11
v3	m5	6
v3	m6	7
v3	m6	12
v3	m6	15
v4	m7	5
v4	m7	10
v4	m7	8
v4	m8	0
v4	m8	8
v4	m8	6

Udfør en trinvis "random factor" variansanalyse og drag konklusioner.

Løsning

Data indtastes

Vælg Analyze ► Fit Model ► Indsæt for Y: tid ,Indsæt for ADD: værksted og mekaniker ► tryk på mekaniker under effects og på værksted under "select columns ► Tryk på Nest ► Emphasis's rullemenu: vælg Minimal Report ► Run

Resultat:

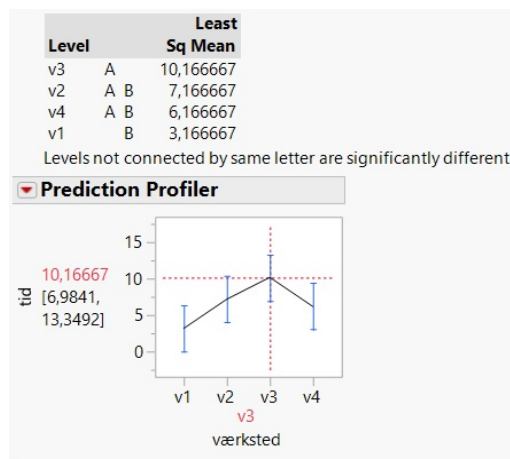
Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
værksted	3	3	150,00000	3,6585	0,0351*
mekaniker[værksted]	4	4	60,66667	1,1098	0,3860

Det ses, at der er forskel på værkstederne, da P-værdi = 0.035 < 0.05.

Derimod er der ikke forskel på mekanikerne.

Pooler man mekanikerne væk, fås ved students t-test osv

Heraf ses, at v3,v2 og v4 ikke kan adskilles. v1 er det dårligste værksted.



I visse situationer, kan man som det følgende eksempel viser frygte, at der er mere end 1 betydende støjkilde. Et såkaldt romersk kvadratforsøg kan så være en mulighed for at dæmpe støjen.

Eksempel 7.7 (romersk kvadratforsøg) 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. Forsøget planlægges som et fuldstændigt faktorforsøg idet hvert niveau skal gentages mindst 4 gange. Dette betyder at der skal udføres 12 delforsøg.

Hvert delforsøg med én bil tager 1 dag. Man frygter, at forskellige dage med deres forskellige vejrforhold kan give stor variation i resultaterne. Endvidere frygter man også, at forskellige biler (med tilhørende chauffør) frygtes give systematisk forskellige resultater.

Da der er i alt 6 forskellige behandlinger ønskes foretaget et **romersk kvadratforsøg** med anvendelse af 6 biler og 6 dage som blokke.

Fordele og ulemper ved planen

Fordelen ved en sådan plan er, at støj der kan skyldes eventuelle forskelle mellem biler og mellem dage elimineres. Ulempe er, at man let kommer op på mange delforsøg (her 36), og en ret kompliceret forsøgsplan. Det er således vigtigt at man foretager en randomisering, og det kan også være ret besværligt.

Eksempel 7.8 (romersk kvadratforsøg)

Der ønskes undersøgt hvorledes udbyttet af et biprodukt ved en fabrikationsproces afhænger af fabrikationsmetode og katalysator type. Der skal vælges mellem 2 forskellige fabrikationsmetoder, og 3 katalysatorer. Det kan ikke udelukkes, at de to faktorer vekselvirker, så man vælger som forsøgsplan en fuldstændig faktorstruktur.

Lad A_i betegne fabrikationsmetode i , og lad B_j betegne katalysator j . Der er følgende 6 behandlinger $A_1 B_1$, $A_1 B_2$, $A_1 B_3$, $A_2 B_1$, $A_2 B_2$, $A_2 B_3$.

Hvert delforsøg varer en uge, og for at forsøgene ikke skal tage for lang tid, er man nødt til at anvende flere apparater samtidig. Da det kan tænkes, at såvel fabrikationstidspunkt som apparatvalg kan påvirke udbyttet, og dermed give en for stor spredning, vælges som forsøgsplan et romersk kvadratforsøg. Da der er 6 behandlinger bliver det et 6×6 kvadrat.

Forsøgsplan og forsøgsresultater var følgende:

		Apparater					
		1	2	3	4	5	6
Uger	1	$A_2 B_1$: 41.3	$A_2 B_2$: 52.6	$A_1 B_2$: 39.7	$A_2 B_3$: 40.8	$A_1 B_3$: 27.6	$A_1 B_1$: 29.5
	2	$A_1 B_1$: 9.3	$A_1 B_3$: 18.7	$A_2 B_3$: 19.5	$A_1 B_2$: 21.1	$A_2 B_2$: 31.9	$A_2 B_1$: 33.2
	3	$A_2 B_3$: 30.5	$A_2 B_1$: 41.2	$A_1 B_3$: 28.7	$A_2 B_2$: 32.4	$A_1 B_1$: 31.9	$A_1 B_2$: 32.2
	4	$A_1 B_3$: 50.7	$A_1 B_2$: 48.3	$A_2 B_1$: 51.6	$A_1 B_1$: 40.9	$A_2 B_3$: 60.4	$A_2 B_2$: 49.8
	5	$A_2 B_2$: 40.8	$A_2 B_3$: 39.6	$A_1 B_1$: 30.7	$A_2 B_1$: 31.8	$A_1 B_2$: 29.6	$A_1 B_3$: 39.8
	6	$A_1 B_2$: 42.7	$A_1 B_1$: 52.4	$A_2 B_2$: 62.1	$A_1 B_3$: 53.6	$A_2 B_1$: 63.6	$A_2 B_3$: 58.7

Foretag en statistisk analyse af forsøget, idet man ønsker så lavt et udbytte som muligt.

Løsning:

Analyseres svarende til en firesidet variansanalyse uden vekselvirkning mellem blokke og faktorer.

Data indtastes

	metode	katalysator	uge	apparat	udbytte
1	a2	b1	u1	a1	41,3
2	a2	b2	u1	a2	52,6
3	a1	b2	u1	a3	39,7
4	a2	b3	u1	a4	40,8
5	a1	b3	u1	a5	27,6
6	a1	b1	u1	a6	29,5
7	a1	b1	u2	a1	9,3

Vælg Analyze ► Fit Model ► Indsæt for Y: udbytte ,Indsæt for ADD: metode, katalysator,apparat, uge ► Indsæt for CROSS: metode, katalysator (marker begge) ► Emphasis's rullemenu: vælg Minimal Report ► Run

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
metode	1	1	662,2044	21,9115	0,0001*
katalysator	2	2	27,8956	0,4615	0,6369
uge	5	5	4385,5389	29,0224	<,0001*
apparat	5	5	181,1189	1,1986	0,3454
metode*katalysator	2	2	61,4756	1,0171	0,3796

7 To faktorer på 2 eller flere niveauer

Da P-værdi for metode*katalysator = 0.3796 > 0.05 ses heraf, at der ikke er vekselvirkning mellem metode*katalysator

Vekselvirkningen "pooles" ned i "error".

Marker under Source: metode*katalysator ► Remove

Effect Tests					
Source	Nparm	DF	Sum of		
			Squares	F Ratio	Prob > F
metode	1	1	662,2044	21,8775	0,0001*
katalysator	2	2	27,8956	0,4608	0,6367
uge	5	5	4385,5389	28,9774	<,0001*
apparat	5	5	181,1189	1,1967	0,3431

Det ses, at faktor katalysator ingen rolle spiller, mens metoden har stor betydning

Det ses endvidere, at det var fornuftigt at foretage en blokopdeling efter uge, mens det var unødigt at dele op efter apparater.

For at finde hvilken metode der giver det mindste udbytte

"Effect Details" ► rød pil ved metode ► Vælg "LS means student, s -t"

Level	Least	
		Sq Mean
a2	A	43,433333
a1	B	34,855556

Levels not connected by same letter are significantly different.

Det ses, at metode 1 giver det laveste udbytte

Da størrelsen af udbyttet afhænger af blokken har det ingen mening at sige, at udbyttet er 34.85

Man kan naturligvis godt undersøge grafisk om data er normalfordelt, men da der ikke er gentagelser kan forudsætningen om varianshomogenitet ikke undersøges.

7.7. 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

7.8. OVERSICHT 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: $R * C \neq 0$

Data indtastes i 3 søjler R, Q og data.

Vælg Analyze ► Fit Model ► Indsæt for Y: data , Indsæt for ADD: R og Q ► Indsæt for CROSS: R og Q (marker begge)

► Emphasis's rullemenu: vælg Minimal Report ► Run Model

(Minimal report er kun valgt for at undgå nogle i denne forbindelse overflødige figurer)

Der fremkommer en lang række tal.

Under "Effects Tests" findes P-værdi for vekselvirkning R*Q.

1) P - værdi $< \alpha$: H_0 forkastes, dvs. Faktorerne R og C vekselvirker

Konfidens og d intervaller opstilles til afgørelse af hvilken kombination, idet man af faktorer, der eksempelvis giver mindst resultat.

"Effect Details" ► rød pil ved "R x Q ► Vælg "LS means Plot"

Der fremkommer en udskrift og en tegning.

Muligvis kan man ud fra tegning se hvilken kombination der eksempelvis giver lavest værdi. Ellers kan det afklares ved på ovennævnte rullemenu at vælge "LSMeans students t"

Det giver en stor tabel (som kan fjernes ved med cursor på overskrift, højre musetast at fjerne markeringen ved "Crosstab Report).

Under den findes en lille tabel hvor man kan se resultatet.

Ønskes fundet 95% konfidensintervaller

rød pil ved "karburator x olieblending" ► "LS Means Table" ► Højremustast i tabel: Vælg "Columns"

► vælg "lower" og derefter "upper".

2) P - værdi $> \alpha$: H_0 accepteres, dvs, der kan ikke påvises en signifikant vekselvirkning

Vekselvirkningen "pooles" ned i "error".

Blandt mange tabeller findes under "Effect Tests" P-værdier for faktorerne R og C

a) Er de begge $< \alpha$ har begge en virkning

Under faktoren R kan man som i forrige tilfælde vælge "LSMeans students t" og finde hvilket niveau der giver det bedste resultat.

Samme gøres under den anden faktor

Et estimat for bedste værdi fås ved

Rød pil ved Response Profiler ► Factor profiling ► Profiler

Ved de fremkomne figurer flyttes linier, så man får bedst resultat.

Rød pil ved Predicter Profiler ► Confidence Intervals

Man kan nu beregne konfidensintervallerne

b) Kun en P-værdi $< \alpha$.

Beregningerne ganske analoge med forrige tilfælde.

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 middeludbyttets (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

- 1) Kontroller om betingelserne om normalitet og varianshomogenitet er rimelig opfyldt. Idet de sædvanlige variansanalyseforudsætninger antages opfyldt, ønskes følgende spørgsmål belyst:
- 2) Angiv hvilke faktorer der har en virkning.
- 3) 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 laboratoriemæssig undersøgelse af mulighederne for at forøge procesudbyttet foretoges bl.a. et forsøg, hvor man dels sammenlignede virkningen af tilsætning af 3 forskellige katalysatorer, dels undersøgte, om udbyttet afhang af, om den nødvendige PH - værdi opnåedes ved tilsætning af HCl i stedet for som hidtil H_2SO_4 .

Forsøgsresultaterne var (udbytteprocenter):

		Tilsat syre			
		HCl		H_2SO_4	
Katalysatorer	1	27.0	27.7	30.1	29.1
	2	25.5	27.0	30.6	28.9
	3	26.5	25.0	25.8	28.0

- 1) Foretag en statistisk analyse af forsøgsresultaterne og drag konklusioner.
- 2) Estimer under hensyn til resultatet af den under punkt 1) foretagne analyse procesudbyttet ved benyttelse af katalysator 2 under tilsætning af HCl og opstil et 95% - konfidensinterval for dette udbytte.

Opgave 7.5

I et fremstillings proces, er en plastikstang fremstillet ved opvarmning af granuleret plastik til en smeltet tilstand og derefter samlet under tryk (pound per square inches) gennem en dyse. For at undersøge hvorledes disse 2 faktorer påvirker den samlingshastighed (målt i inches per. second) hvormed plastikstangen når tilstrækkelig styrke, laves et fuldstændigt faktorforsøg.

Tre slags granuleret plastik anvendes, og man vælger at benytte dem som en blokke i forsøget

Resultaterne blev

		blok					
		1		2		3	
		tryk		tryk		tryk	
		40	60	40	60	40	60
temperatur	200 ^o	1.35	1.74	1.31	1.67	1.40	1.86
	300 ^o	2.48	3.63	2.29	3.30	2.14	3.27

Foretag en statistisk analyse af forsøgsresultaterne og drag konklusioner.

Opgave 7.6

Hver af tre laboranter har bestemt hydroquinons smeltepunkt (^o 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 7.7.

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 målte.

Nedenfor er anført en skematisk oversigt over forsøgsresultaterne:

		Vitaminbehandlinger			
		1	2	3	4
Giftstoffer	1	3.1	8.2	4.3	4.5
		4.5	11.0	4.5	7.1
	2	2.2	3.0	2.3	3.0
		2.1	3.7	2.5	3.6

Teoretiske overvejelser i forbindelse med tidligere lignende forsøg har vist, at variabeltransformationen $\sqrt{\frac{1}{x}}$ sikrer den for analysen nødvendige varianshomogenitet.

- 1) Analyser forsøgsresultaterne og drag konklusioner med hensyn til faktorernes 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.

Opgave 7.8.

Ved en enzymproduktion havde man valget mellem anvendelse af 4 forskellige råmaterialer R_1 , R_2 , R_3 og R_4 . Man ønskede ved et forsøg at foretage en sammenligning mellem disse med hensyn til det opnåede procesudbytte.

Hvert delforsøg krævede en laborants fuldtidsmedvirken og varede 5 dage. Af ressourcernæssige grunde måtte forsøgsarbejdet fordeles på flere uger, idet man kunne få 4 laboranter stillet til disposition. Man kunne på forhånd ikke udelukke, at forsøgsresultaterne kunne afhænge af forsøgstidspunktet og af, hvilken laborant der medvirkede.

Forsøgsplan og forsøgsresultater var (kodede tal):

		Laboranter							
		1		2		3		4	
Uger	1	R_1	58	R_4	108	R_2	75	R_3	76
	2	R_4	60	R_3	26	R_1	-15	R_2	20
	3	R_3	-16	R_2	23	R_4	24	R_1	-24
	4	R_2	17	R_1	31	R_3	17	R_4	41

- 1) Foretag en statistisk analyse til undersøgelse af, om de 4 råmaterialer i middel giver samme udbytte
- 2) Find i benægtende fald den type råmateriale der giver størst udbytte.
- 3) Undersøg tillige, om procesudbyttet i middel må antages at være påvirket af forsøgstidspunkt og/eller laborant.

Opgave 7.9.

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 middeltallet af garnbrud afhænger af vævemetoder og/eller materialetyper.

8 FLERE END TO FAKTORER, SCREENINGS-FORSØG

8.1 INDLEDNING

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.

8.2 Flere end 2 faktorer med begrænset antal forsøg.

Er antallet af faktorer, niveauer og gentagelser begrænset, så antallet af forsøg er overkommeligt, kan man godt naturligvis bruge samme metode som ved den tosidede variansanalyse.

Følgende eksempel viser hvordan.

Eksempel 8.1 (4-sidet variansanalyse med $n = 1$)

Virksomheden af fire faktorer på hærningstiden for en bestemt cementtype ønskes undersøgt.

De 4 faktorer var

A:Omrøringshastighed:2 niveauer A_1, A_2 C:Formalingsgrad af bestemt reaktant 2 niveauer C_1, C_2

B: Fabrikationsmetoder: B_1, B_2, B_3 D: Råvaretype 4 niveauer D_1, D_2, D_3, D_4 .

Forsøgsplanen var et fuldstændigt randomiseret forsøg med en fuldstændig faktorstruktur med 1 gentagelse af hver behandling. Følgende observationer af Y (kodede tal) fandtes:

		D_1		D_2		D_3		D_4	
		C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2
B_1	A_1	23.8	24.4	25.3	21.5	23.1	15.3	4.12	11.6
	A_2	45.8	47.7	35.5	26.1	10.0	20.0	5.30	13.2
B_2	A_1	16.0	18.9	11.6	11.8	22.7	20.9	28.9	21.1
	A_2	27.0	31.0	31.6	31.6	19.8	13.8	15.5	26.8
B_3	A_1	10.1	9.3	13.2	19.0	30.3	27.9	34.1	35.5
	A_2	17.0	27.8	24.6	25.2	17.0	27.2	21.1	34.7

Det antages, at forudsætningerne for at udføre en variansanalyse er opfyldt.

- 1) Undersøg hvilke af de fire faktorer, der har virkninger eventuelt i form af vekselvirkninger.
- 2) Angiv hvilke niveauer faktorerne skal stilles på for at give den mindst mulige størkningstid..
- 3) Angiv et estimat for den mindste størkningstid, og et 95% konfidensinterval herfor.

Løsning

- 1) Data indtastes på sædvanlig måde

	A	B	C	D	Hærningstid
1	A1	B1	C1	D1	23,8
2	A1	B1	C2	D1	24,4
3	A1	B1	C1	D2	25,3
4	A1	B1	C2	D2	21,5
5	A1	B1	C1	D3	23,1
6	A1	B1	C2	D3	15,3
7	A1	B1	C1	D4	4,12
8	A1	B1	C2	D4	11,6
9	A2	B1	C1	D1	45,8

Først testes om der er vekselvirkninger

Vælg Analyze ► Fit Model ► Indsæt for Y: Hærningstid , Indsæt for ADD: A, B, C, D ► Indsæt for CROSS: de 6 vekselvirkninger (Man indsætter eksempelvis AC ved under Construct Model Effect at markere A og på Columns markere A og tryk på Cross, så fremkommer AC ► Emphasis's rullemenu: vælg Minimal Report ► Run

Resultat:

Det ses, at kun vekselvirkningerne AD og BD har betydning.

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
A	1	1	274,9461	14,3733	0,0009*
B	2	2	22,7433	0,5945	0,5601
C	1	1	49,7761	2,6021	0,1204
D	3	3	140,3956	2,4465	0,0896
A*B	2	2	52,6883	1,3772	0,2723
B*C	2	2	49,7123	1,2994	0,2920
C*D	3	3	81,2751	1,4163	0,2635
A*C	1	1	77,3176	4,0419	0,0563
A*D	3	3	1009,3437	17,5884	<,0001*
B*D	6	6	2065,2491	17,9941	<,0001*

Vi “excluderer” nu alle tofaktorvekselvirkningerne, der ikke har betydning (gøres som under eksempel 7.3)

Det ses, at faktor C ingen rolle spiller, mens A, B og D alle har betydning i form af vekselvirkning.

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
A	1	1	274,9461	12,1595	0,0015*
B	2	2	22,7433	0,5029	0,6096
C	1	1	49,7761	2,2014	0,1480
D	3	3	140,3956	2,0697	0,1245
A*D	3	3	1009,3437	14,8795	<,0001*
B*D	6	6	2065,2491	15,2227	<,0001*

Af principielle hensyn slettes faktor C i modellen (det har i praksis kun meget ringe betydning, da vi jo allerede har 31 frihedsgrader i error).

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
A	1	1	274,9461	11,7196	0,0017*
B	2	2	22,7433	0,4847	0,6203
D	3	3	140,3956	1,9948	0,1345
A*D	3	3	1009,3437	14,3411	<,0001*
B*D	6	6	2065,2491	14,6718	<,0001*

2) “Effect Details” ► rød pil ved “B x D ► Vælg “LS Means Students t ”

Det giver en stor tabel. Under den findes følgende tabel

Level	Least Sq Mean
B1,D1 A	35,425000
B3,D4 A B	31,350000
B1,D2 B C	27,100000
B3,D3 B C D	25,600000
B2,D1 C D E	23,225000
B2,D4 C D E	23,075000
B2,D2 C D E F	21,650000
B3,D2 C D E F	20,500000
B2,D3 D E F	19,300000
B1,D3 E F	17,100000
B3,D1 F	16,050000
B1,D4 G	8,555000

Levels not connected by same letter are significantly different.

Heraf ses, at man skal vælge B på niveau B1 og D på niveau D4, så bliver størkningstiden mindst.

Level	Least Sq Mean
A2,D1 A	32,716667
A2,D2 A	29,100000
A1,D3 B	23,366667
A1,D4 B C	22,553333
A2,D4 B C	19,433333
A2,D3 B C	17,966667
A1,D1 C	17,083333
A1,D2 C	17,066667

Levels not connected by same letter are significantly different.

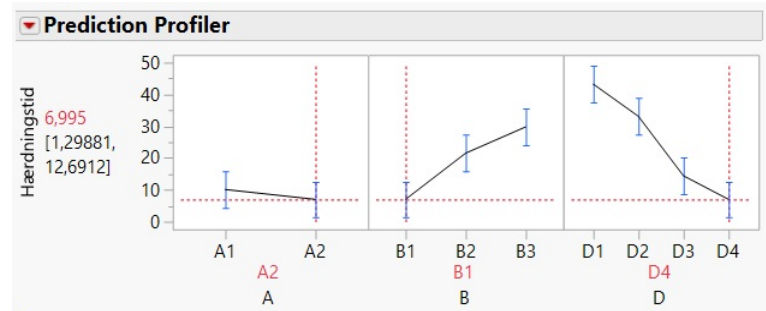
Af den tilsvarende tabel for AD ses, at et valg af D på niveau 4 gør, at man frit kan vælge A på såvel højt som lavt niveau

Konklusion: A (omrøringshastighed) vælges frit, B (fabrikationsmetode) på niveau 1, C (formalingsgrad) har ingen betydning og D (råvaretype) på niveau 4.

3) Rød pil ved Response Brudstyrke ► Factor profiling ► Profiler

8. Flere end to forsøg, Screeningsforsøg

Ved de fremkomne figurer flyttes linier



Mindste tørkningstid 7.0 95% konfidensinterval er [1.3 ; 12.7]

8.3 Screeningsforsøg

Har man mange faktorer, som de følgende eksempler viser, så vil den hidtidige variansanalyse forudsættet mange hundrede forsøg, hvilket naturligvis ikke er praktisk muligt.

Eksempel 8.2. (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: Syrekonzentrationen

C: Katalysatoren

D: Reaktionstid

E: Reaktionstemperatur.

Eksempel 8.3. (3 faktorer)

Antallet af brud i en stålfjeder er rigelig stor. Følgende 3 faktorer har muligvis betydning:

A: Stålets procentiske indhold af kul.

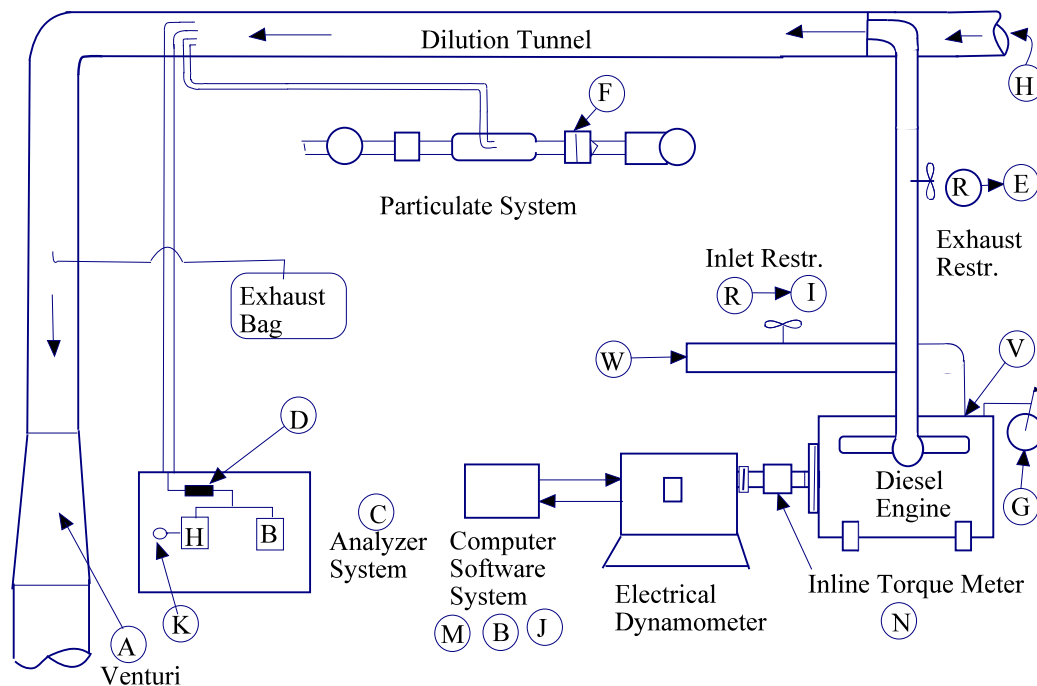
B: Temperaturen af det oliebad, som fjederen dyppes ned i under hærningen

C: Fjederens temperatur lige før den nedsænkes i oliebadet.

Eksempel 8.4 (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,



Formål med screeningsforsøg

Hvordan kan man med det *mindst* mulige antal forsøg dels *finder hvilke* faktorer der har en virkning, dels *finder hvorledes* disse faktorer påvirker resultatet (ændres middelværdien?, ændres spredningen?, øges robustheden overfor påvirkninger fra omgivelserne eller fra "dårlige" komponenter?).

Her er det, at de i de følgende afsnit omtalte screeningsforsøg er centrale.

Man kan ved denne metode med forholdsvis få delforsøg undersøge mange faktorer 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. Der gælder altid den regel, at man bør højst bruge 25% af det planlagte budget på det første forsøg.

Den statistiske metode blev i 30-erne udviklet af en berømt statistiker R. Fisher, (F- fordelingen er opkaldt efter ham) .

Metoden fik imidlertid først sin helt store udbredelse, da en japaner "Taguchi" formåede at gøre metoden let forståelig og derved sikrede dens udbredelse. Endvidere udvidede han metoden til også at finde hvilke faktorer der bidrog med den største spredning. Dette var bl.a. medvirkende til at sikre japanske bilers overlegne kvalitet i forhold til eksempelvis de amerikanske biler.

Hvilke faktorer kan have betydning

Man starter med at søge at finde alle de faktorer, som på nogen mulig måde kan tænkes at have en indflydelse på resultatet.

Her kan "brainstorm" være værdifuld, idet man "uden hæmminger" skal foreslå mulige faktorer. Ved at lade en gruppe af "sagkyndige" på denne måde inspirerer hinanden, vil mange

utraditionelle forslag komme frem.

Man må passe på ikke at affærdige foreslåede faktorer som værende helt hen i vejret. Det er trods alt bedre at tage mange faktorer med og så opdage, at de fleste ingen rolle spiller, end at risikere at udskyde en væsentlig faktor.

Udførelse af forsøg

Det har stor betydning for forsøgets succes, at man er meget omhyggelig med at sammensætte den gruppe af mennesker, som skal planlægge og udføre forsøget. Hvis det drejer sig om udvikling af et nyt produkt vil det ofte være udviklingsafdelingen der både planlægger og udfører forsøget. Hvis det drejer sig om et problem i en eksisterende produktion, som måske er koncentreret om en bestemt maskine (eller del af produktionen), er det vigtigt at gruppen har medlemmer (driftsingeniører, værkførere, opstillere), som arbejder tæt på maskinen (den del af produktionen) til daglig. Det sikrer også, at de forstår ideen i forsøget, og derved sikrer at forsøget bliver udført som planlagt.

Man har eksempler på, at hvis man ikke har forklaret metoden, så har man "på gulvet" strøget nogle forsøg, fordi man fandt dem ganske overflødige, hvilket dermed ødelagde hele screeningen.

Screeningsforsøg

Har man mange faktorer der skal undersøges, må man for at begrænse antallet af delforsøg nøjes med at undersøge **hver faktor på 2 niveauer** "lavt" niveau og "højt" niveau. Disse niveauer forsøges valgt som "yderpunkter" med hensyn til faktorerens formodede virkning. Har man ingen formodning herom, vælges sædvanligvis to faktorniveauer, som "afviger" mest muligt.

For at reducere antallet af forsøg mest muligt antages endvidere, **at alle behandlinger kun udføres 1 gang.**

Har man k faktorer, vil antallet af forsøg være 2^k , deraf navnet " 2^k faktorforsøg".

Har man 4 faktorer er antallet af delforsøg $2^4=16$, hvilket jo ikke er mange, men hvis eksempelvis der er 7 faktorer, er antallet af delforsøg $2^7=128$, hvilket begynder at være uoverskueligt mange.

For at reducere antallet af delforsøg forudsættes derfor yderligere, **at der ikke forekommer vekselvirkninger mellem mere end 2 faktorer.**

Det er en erfarings sag, at der meget sjældent forekomme signifikante vekselvirkninger mellem 3 faktorer, og vekselvirkninger mellem mere end tre faktorer er aldrig forekommet.

Man kan derfor poole disse "ledige" vekselvirkninger sammen til et estimat for spredningen.

Har vi eksempelvis 4 faktorer A, B, C og D, og hver faktor er på 2 niveauer, har de alle 1 frihedsgrad.

Vi skal finde følgende 10 faktorer A, B, AB, C, AC, BC, D, AD, BD, CD

Det giver 10 frihedsgrader. Da de 16 forsøg i alt har 15 frihedsgrader, så er der følgelig 5 frihedsgrader til rådighed til bestemmelse af den fælles varians.

Et forsøg som ovenstående kaldes et **fuldstændigt 2^4 - faktorforsøg**, da man er interesseret i at teste alle 2-faktor-vekselvirkninger.

Hvis man har mange faktorer giver det dog stadig alt for mange delforsøg

Man kan imidlertid reducere antal forsøg kraftigt ved såkaldte "partielle 2^k faktorforsøg.

Sædvanligvis ved man (måske fra tidligere forsøg), at visse 2-faktorvekselvirkninger ikke er signifikante. Dette kan man som der vises i det følgende udnytte til at reducere antallet af delforsøg kraftigt.

8.4 NOMENKLATUR

Faktorerne omdøbes til A, B, C osv. og de antages alle på 2 niveauer, et "højt niveau og et lavt niveau

En særdeles bekvem notation er følgende:

(1): Alle faktorer er på lavt niveau

a: A på højt niveau, alle andre på lavt niveau.

b: B på højt niveau, alle andre på lavt niveau.

ab: A og B på højt niveau, alle andre på lavt niveau.

c: C på højt niveau, alle andre på lavt niveau.

ac: A og C på højt niveau, alle andre på lavt niveau osv.

Nomenklaturen anvendes i en dobbelt betydning, idet eksempelvis a også betyder forsøgsresultaterne med denne behandling.

Faktorerne A, B, C ... på lavt niveau benævnes A_1, B_1, C_1, \dots og på højt niveau A_2, B_2, C_2, \dots

Nomenklatur anskueliggøres i følgende eksempel.

Eksempel 8.5 2^3 faktorforsøg.

Ved et fuldstændigt randomiseret laboratorieforsøg undersøges, hvorledes udbyttet af en nitreringsproces, hvis resultat (i udbytteprocenter) indgår i en fabrikation af farvestoffer, afhænger af 3 faktorer

	Lav	Høj
A: Den tilføjede salpetersyre	kold	varm
B: Omrøringstiden	kort	lang
C: Resteffekt	Renset beholder før delforsøget	urensset beholder

Symbolikken fremgår af følgende skema

	C_1 (renset beholder)		C_2 (urensset beholder)	
	B_1 (kort)	B_2 (lang)	B_1 (kort)	B_2 (lang)
A_1 (kold)	(1)	b	c	bc
A_2 (varm)	a	ab	ac	abc

Beregninger

Forsøgsresultaterne indføres nu i skemaet, hvor også nomenklatur og gennemsnit er anført.

	C_1 (renset beholder)		C_2 (urensset beholder)		gennemsnit
	B_1 (kort)	B_2 (lang)	B_1 (kort)	B_2 (lang)	
A_1 (kold)	(1) 87.2	b 82.0	c 86.7	bc 83.4	84.825
A_2 (varm)	a 88.4	ab 83.0	ac 89.2	abc 83.7	86.075

- Angiv faktor A's hovedvirkning og tegn den
- Angiv vekselvirkningen BC og tegn den

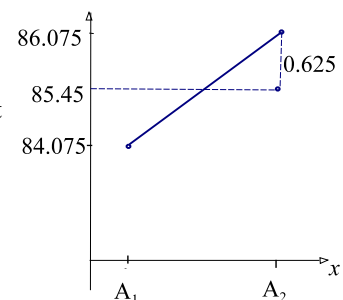
Løsning (håndregning)

Hovedvirkning

Ved A's hovedvirkning forstås effekten af en ændring ud fra gennemsnitsniveauet til niveauet, når A er på højt niveau.

Et estimat (skøn) for A's hovedvirkning er i eksempel 8.5 $\tilde{A} = 86.075 - 85.45 = 0.625$. (se figuren).

eller i symboler



8. Flere end to forsøg, Screeningsforsøg

$$\tilde{A} = \frac{\frac{a+ab+ac+abc}{4} - \frac{(1)+b+c+bc}{4}}{2} = \frac{-(1)+a-b+ab-c+ac-bc+abc}{8}$$

Bemærk: Et estimat for A skrives kort \tilde{A} osv. (der sættes en "tilde" over bogstavet).

B og C's hovedvirkning defineres analogt

Vekselvirkning:

Lad os betragte vekselvirkningen AB.

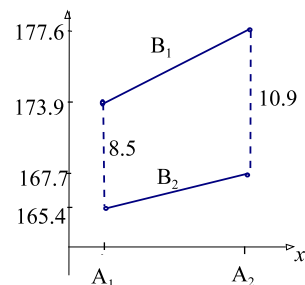
$$A_2B_2 = \frac{ab+abc}{2} = \frac{83+83.7}{2} = 83.35,$$

$$A_2B_1 = \frac{a+ac}{2} = \frac{88.4+89.2}{2} = 88.8$$

Vi har

$$A_1B_2 = \frac{b+bc}{2} = \frac{82+83.4}{2} = 82.7$$

$$A_1B_1 = \frac{(1)+c}{2} = \frac{87.2+86.7}{2} = 86.95$$



Tegnes forbindelseslinierne (se figuren) ses, at linierne ikke er parallelle.

$$\text{"Hældning" på } B_2 \text{ linie} = \frac{\frac{ab+abc}{2} - \frac{b+bc}{2}}{2} = \frac{83.35 - 82.7}{2} = 0.325$$

$$\text{"Hældning" på } B_1 \text{ linie} = \frac{\frac{a+ac}{2} - \frac{(1)+c}{2}}{2} = \frac{88.8 - 86.95}{2} = 0.925$$

Vekselvirkningen AB er defineret som den gennemsnitlige forskel på hældningen

$$AB = \frac{\frac{\frac{ab+abc}{2} - \frac{b+bc}{2}}{2} - \frac{\frac{a+ac}{2} - \frac{(1)+c}{2}}{2}}{2} = \frac{+(1)-a-b+ab+c-ac-bc+abc}{8} = \frac{0.325 - 0.925}{2} = -0.3$$

Mange statistikprogrammer tegner disse skrå figurer, som gør det lettere at tolke eksempelvis vekselvirkninger

Generelle formler for hoved- og vekselvirkninger

Fortegnsmatrix: Et vigtigt hjælpemiddel er de såkaldte fortegnsmatricer:

Disse er opbygget således, at

- 1) A, B og C søjlerne har +, hvis det til faktoren svarende lille bogstav indgår, ellers -.
- 2) AB, AC, BC og ABC søjlerne er produktet af tilsvarende fortegn i hovedsøjlerne.

Eksempel: Skemaet

	C ₁ (renset beholder)		C ₂ (urenset beholder)	
	B ₁ (kort)	B ₂ (lang)	B ₁ (kort)	B ₂ (lang)
A ₁ (kold)	(1) 87.2	b 82.0	c 86.7	bc 83.4
A ₂ (varm)	a 88.4	ab 83.0	ac 89.2	abc 83.7

kan beskrives

	I	A	B	AB	C	AC	BC	ABC	Resultat
(1)	+	-	-	+	-	+	+	-	87.2
a	+	+	-	-	-	-	+	+	88.4
b	+	-	+	-	-	+	-	+	82.0
ab	+	+	+	+	-	-	-	-	83.0
c	+	-	-	+	+	-	-	+	86.7
ac	+	+	-	-	+	+	-	-	89.2
bc	+	-	+	-	+	-	+	-	83.4
abc	+	+	+	+	+	+	+	+	83.7

For at beregne A's virkning betragtes søjlen under A. Tælleren er så netop tallene i højre kolonne med det dertil svarende fortegn.

Eksempel : $\tilde{A} = \frac{-(1)+a-b+ab-c+ac-bc+abc}{2^3}$ (nævner er det totale antal delforsøg).

$$\tilde{A} = \frac{-87.2+88.4-82.0+83-86.7+89.2-83.4+83.7}{2^3} = \frac{5}{8} = 0.625$$

$$AB = \frac{87.2-88.4-82+83+86.7-89.2-83.4+83.7}{2^3} = -\frac{2.4}{8} = -0.3$$

Det ses, at vi får værdierne fra før.

Resultaterne kan generaliseres til et vilkårligt antal faktorer.

8.5 Definitionsrelationer og aliasrelationer.

Lad os indledningsvis som eksempel betragte et fuldstændigt 2^2 -faktorforsøg.

Fortegnsmatricen for et sådant forsøg er

	I	A	B	AB
(1)	+	-	-	+
a	+	+	-	-
b	+	-	+	-
ab	+	+	+	+

Ved man, at vekselvirkningen AB er nul, kan man benytte fortegnssøjlen for AB til beregning af et estimat for en ny faktor C.

Faktor C kan indføres i den fuldstændige 2^2 -faktorstruktur på en måde, som vi betegner $C = AB$. Ved denne lader vi faktor C indgå på sit høje niveau i en behandling, når der i fortegnsmatricen for den fuldstændige 2^2 -faktorstruktur er et + i AB-søjlen, og vi lader faktor C indgå i behandlingen på sit lave niveau, når der er et minus (se følgende skema).

	I	A	B	AB	C=AB
(1)	+	-	-	+	+
a	+	+	-	-	-
b	+	-	+	-	-
ab	+	+	+	+	+

Med 3 faktorer skulle vi normalt have udført $2^3 = 8$ forsøg,

I ovenstående skema kan vi se, at netop halvdelen af alle behandlingerne i en fuldstændig 2^3 -

faktorstruktur udføres.

Den fremkomne “**partielle 2^3 - faktorstruktur**” kaldes derfor for en $\frac{1}{2} \cdot 2^3$ -faktorstruktur.

Vi opskriver nu i en fortegnsmatrix samtlige virkninger

Fortegnssøjlerne bestemmes analogt med tidligere: I A-, B- og C - søjlerne sættes + eller - eftersom det tilsvarende lille bogstav indgår i behandlingsnotationen eller ikke gør det. De øvrige søjler udfyldes efter produktreglen, hvor eksempelvis fortegnene i ABC - søjlen fås ved produkt af fortegnene i A-, B- og C-søjlerne.

Underliggende struktur	Behandlinger	I	A	B	AB	C=AB	AC	BC	ABC
-1	c	+	-	-	+	+	-	-	+
a	a	+	+	-	-	-	-	+	+
b	b	+	-	+	-	-	+	-	+
ab	abc	+	+	+	+	+	+	+	+

Af fortegnsmatricen ses, at der følger en række relationer mellem fortegnssøjlerne, som symbolsk kan skrives:

$$A = BC \quad B = AC \quad C = AB \quad I = ABC$$

Det kan vises, at dette betyder, at vi ingen mulighed har for at estimere virkning ABC, og at de linearkombinationer, der benyttes til estimation af virkningerne A, B og C, i virkeligheden estimerer summer af virkninger: A+BC, B+AC og C+AB.

Vi benytter den terminologi, at A og BC er hinandens aliaser, B og AC er hinandens aliaser, og at C og AB er hinandens aliaser. Vi siger også, at virkningerne er sammenblandede, og at **aliasrelationerne** er $A = BC$, $B = AC$ og $C = AB$.

Vi kalder endvidere $I=ABC$ for den partielle faktorstrukturs **definitionsrelation**.

Vi kunne i stedet have sat $C = -AB$, dvs. at vi lader C-søjlen være AB - søjlen med modsat fortegn. Det giver følgende fortegnsskema

Underliggende struktur	Behandlinger	I	A	B	AB	C = - AB	AC	BC	ABC
(1)	(1)	+	-	-	+	-	+	+	-
a	ac	+	+	-	-	+	+	+	-
b	bc	+	-	+	-	+	-	+	-
ab	ab	+	+	+	+	-	-	-	-

Analogt med før ses, at der symbolsk gælder en række relationer:

$$A = -BC \quad B = -AC \quad C = -AB \quad I = -ABC$$

I dette tilfælde er $I=-ABC$ den partielle faktorstrukturs definitionsrelation.

Vi kan af relationerne se, at hvis faktoren C vekselvirker med de oprindelige faktorer A og B, så kan vi ikke estimere A og B rent, da de er sammenblandede med disse vekselvirkninger. Da vi naturligvis ønsker at kunne estimere faktorerne, så er den ovenstående forsøgsplan kun acceptabel, hvis alle 2-faktorvekselvirkninger på forhånd kan antages at være 0.

Produktregler

Ved betragtning af fortegnssøjlerne i den tilsvarende fortegnsmatrix indses let, at der for "multiplikation" af fortegnssøjler gælder en række simple "produktregler", hvoraf vi eksempelvis med let forståelig symbolik anfører:

$$A \cdot A = B \cdot B = C \cdot C = I,$$

$$I \cdot A = A \cdot I = A, \text{ (dvs. } I \text{ svarer til tallet 1 ved normal multiplikation)}$$

$$AB = A \cdot B = B \cdot A$$

$$ABC = A \cdot B \cdot C = B \cdot A \cdot C = C \cdot A \cdot B \quad \text{osv.}$$

Ved benyttelse af disse produktregler fås aliasrelationerne umiddelbart ud fra definitionsrelationen $I=ABC$:

$$\text{Ved multiplikation med } A: \quad A = BC.$$

$$\text{Ved multiplikation med } B: \quad B = AC$$

$$\text{Ved multiplikation med } C: \quad C = AB.$$

Omvendt kan definitionsrelationen fås ud fra en vilkårlig aliasrelation ved en multiplikation, f.eks. fås ud fra $C=AB$ ved multiplikation med C : $I=ABC$.

Foranstående mønster kan generaliseres, og der kan udarbejdes en fortegnsmatrix for enhver 2^k -faktorstruktur.

8.6 Planlægning af et partielt forsøg

Et partielt 2^k faktorforsøg kan opdeles i tre faser

- 1) **Planlægning:** Angivelse af hvilke forsøg der skal laves.
- 2) **Udførelse:** Forsøgene udføres randomiseret
- 3) **Beregning:** Beregne P-værdier, angive faktorer med signifikante virkninger, evt. hvilke niveauer faktorerne skal indstilles på for at give maksimal virkning.

Vi vil i dette afsnit give et par eksempler på hvorledes man planlægger et partielt forsøg ud fra givne oplysninger, mens vi i afsnit 8.8 vil vise hvorledes man ved hjælp af SAS.JMP udfører beregningerne.

Eksempel 8.6. Planlægning af et partielt 2^k faktorforsøg

Virkningerne af 4 faktorer A, B, C og D ønskes undersøgt ved anvendelse af færrest mulige delforsøg.

Man kan ikke på forhånd udelukke, at faktoren A vekselvirker med faktorerne B og C, mens det anses for udelukket, at der kan forekomme andre vekselvirkninger.

Opstil en forsøgsplan, og angiv herunder

- a) Den underliggende fuldstændige faktorstruktur, og relationerne hvormed nye faktorer er indført i den underliggende struktur
- b) Antal frihedsgrader til "error" (forsøgsfejls varians)
- c) Forsøgets behandlinger.

Løsning:

- a) Der skal undersøges virkningerne af A, B, C, D, AB og AC, dvs. i alt 6 virkninger

Man kan derfor muligvis nøjes med at lave 8 delforsøg, dvs. vi vil udføre et partielt $\frac{1}{2} 2^4$ faktorforsøg, da et sådant forsøg jo har 7 frihedsgrader. Der er dermed 1 frihedsgrad til "error".

8. Flere end to forsøg, Screeningsforsøg

Vi opskriver samtlige hoved- og vekselvirkninger i den underliggende fuldstændige 2^3 struktur, og sætter x ved de virkninger man foreløbig skal undersøge

	Beslaglagt virkning
A	x
B	x
AB	x
C	x
AC	x
BC	
ABC	

Derefter indsættes faktoren D i modellen.

Sættes $D = ABC$ ses, at $AB = DC$ og $AC = BD$.

Da DC og BD antages at være forsvindende, kan AB og AC estimeres rent.

(En anden mulighed: Sæt $D = BC$, hvilket gav $B = CD$ og $C = BD$, hvilket også er i orden.)

Tabellen nedenfor angiver relationen hvormed D er indført.

	Beslaglagt virkning
A	x
B	x
AB	x
C	x
AC	x
BC	
ABC = D	x

b) Antal frihedsgrader til error : 1

c) De behandlinger (forsøg der skal laves) bestemmes af følgende fortegnsskema:

	A	B	C	D=ABC	
(1)	-	-	-	-	(1)
a	+	-	-	+	ad
b	-	+	-	+	bd
ab	+	+	-	-	ab
c	-	-	+	+	cd
ac	+	-	+	-	ac
bc	-	+	+	-	bc
abc	+	+	+	+	abcd

En kontrol er, at alle faktorer skal være 4 gange på højt og 4 gange på lavt niveau.

Forsøgene udgøres nu (randomiseret) efter denne plan.

Har man ved lodtrækning fundet, at man først skal lave forsøget "ad", så laves altså et forsøg med faktorerne A og D på højt niveau og B og C på lavt niveau. ◆

Har man adgang til et statistikprogram som eksempelvis SAS.JMP er det ikke nødvendigt selv at lave et sådant skema, da det automatisk dannes, når man har indlagt sin plan. En fordel er desuden, at der sker randomiseret, så forsøgene automatisk bliver lavet i tilfældig orden.

Eksempel 8.7 Konstruktion af forsøgsplan

Virkningerne af 7 faktorer $T_1, T_2, T_3, T_4, T_5, T_6$ og T_7 ønskes undersøgt ved et screeningsforsøg. Man kan ikke på forhånd udelukke, at faktoren T_6 har 2-faktorvekselvirkninger med faktorerne T_2, T_3 og T_4 . Endvidere kan faktoren T_2 også tænkes at have 2-faktorvekselvirkninger med faktorerne T_3 og T_4 . Endelig er det ikke utænkelig, at T_5 vekselvirker med T_7 .

Alle andre vekselvirkninger kan derimod anses for udelukket.

Opstil en forsøgsplan med færrest mulige delforsøg, og angiv herunder

- Den underliggende fuldstændige faktorstruktur, samt relationerne hvormed nye faktorer er indført i den underliggende struktur
- Antal frihedsgrader til "error" (forsøgsfejls varians)
- Forsøgets behandlinger.

Løsning

- Vekselvirkninger opskrives: $T_6 T_2, T_6 T_3, T_6 T_4, T_2 T_3, T_2 T_4, T_5 T_7$.

Den underliggende struktur skal altid indeholde de faktorer, som har flest mulige vekselvirkninger med andre.

T_1 indgår i 0 vekselvirkning, T_2 indgår i 3 vekselvirkninger, T_3 indgår i 2 vekselvirkninger, T_4 indgår i 2 vekselvirkninger, T_5 indgår i 1 vekselvirkning, T_6 indgår i 3 vekselvirkninger, T_7 indgår i 1 vekselvirkning

Den mest "aktive" faktor benævnes med A, den næstmest aktive med B osv.

Vi foretager følgelig følgende omdøbning: $A = T_6, B = T_2, C = T_3, D = T_4, E = T_5, F = T_7$ og $G = T_1$.

Der skal i alt estimeres 7 hovedvirkninger og 6 vekselvirkninger AB, AC, AD, BC, BD, EF.

Disse 13 virkninger kan muligvis estimeres ud fra en underliggende 2^4 faktorstruktur, da $13 < 15$.

Vi prøver derfor om det er muligt at konstruere en $\frac{1}{2^3} 2^7$ -faktorstruktur.

Der opskrives samtlige hoved- og vekselvirkninger i den underliggende fuldstændige 2^4 -faktorstruktur.

	Beslaglagt virkning		Beslaglagt virkning
		D	x
A	x	AD	x
B	x	BD	x
AB	x	ABD	
C	x	CD	
AC	x	ACD	
BC	x	BCD	
ABC		ABCD	

I skemaet er med "x" markeret de "beslaglagte" virkninger, dvs. virkninger som kan være forskellig fra 0.

Vi skal nu indføre E og F så EF ikke er beslaglagt.

Efter at have prøvet forskellige kombinationer fandtes, at sættes $E = ABC$ og $F = CD$ er $EF = ABD$, som ikke er beslaglagt (ses ved, at $EF = ABCCD = ABD$). Da G ikke vekselvirker med noget, kan den sættes på en vilkårlig ledig plads.

Resultatet kan samles i tabellen

8. Flere end to forsøg, Screeningsforsøg

	Beslaglagt virkning		Beslaglagt virkning
		D	x
A	x	AD	x
B	x	BD	x
AB	x	ABD=EF	x
C	x	CD = F	x
AC	x	ACD	
BC	x	BCD	
ABC=E	x	ABCD = G	x

Vi har dermed vist, at en $\frac{1}{2^3} 2^7$ - faktorstruktur med ovenstående struktur kan benyttes til testning af de mulige hoved- og 2-faktorvekselvirkninger.

b) Da der i alt skal estimeres 7 hovedvirkninger og 6 vekselvirkninger er der $15 - 13 = 2$ frihedsgrader til error

c) Behandlinger

Underliggende struktur	E=ABC	F=CD	G=ABCD	Behandlinger
(1)	-	+	+	fg
a	+	+	-	aef
b	+	+	-	bef
ab	-	+	+	abfg
c	+	-	-	ce
ac	-	-	+	acg
bc	-	-	+	bcg
abc	+	-	-	abce
d	-	-	-	d
ad	+	-	+	adeg
bd	+	-	+	bdef
abd	-	-	-	abd
cd	+	+	+	cdefg
acd	-	+	-	acdf
bcd	-	+	-	bcdf
abcd	+	+	+	abcdefg



8.7 KONFUNDERING AF BLOKKE

Blokke dannes som bekendt for at forminske “støjen”. Eksempelvis kan det være nødvendigt af tidsmæssige grunde at benytte to ovne ved forsøget, og den ene ovn er af ældre dato, tager længere tid at varme op, og man kan derfor risikere, at forsøgsresultaterne generelt er “lavere” end et tilsvarende forsøg i den anden ovn.

Man kan så lave fuldstændige blokke, dvs. alle behandlinger skal være repræsenteret i begge blokke. Uden en sådan balance er det klart ikke muligt at give en “fair” vurdering af behandlingerne. Imidlertid koster det jo mange forsøg, idet eksempelvis 2 blokke jo så kræver, at alle behandlinger skal udføres 2 gange.

Skal antal delforsøg reduceres, er det derfor væsentligt, at have en metode til at lave lave balancerede ufuldstændige blokke. Efter nogenlunde samme teknik som ved indførelsen af faktorer

i partielle forsøg, kan dette lade sig gøre. Man siger så, at man har konfunderet blokkene i strukturen.

Metoden vises i det følgende eksempel.

Eksempel 8.8. Konstruktion af et konfunderet partielt forsøg.

Virkningerne af 5 faktorer T_1, T_2, T_3, T_4 og T_5 ønskes undersøgt ved et screeningsforsøg. Man ved at ingen af faktorerne T_2 og T_4 vekselvirker med andre faktorer. Man er nødt til at lave et blokforsøg, med 4 forsøgseenheder i hver blok.

Opstil en forsøgsplan med færrest mulige delforsøg, og angiv herunder

- Den underliggende fuldstændige faktorstruktur, samt relationerne hvormed nye faktorer og blokke er indført i den underliggende struktur.
- Forsøgets behandlinger, og de 4 blokkes placering.
- Antal frihedsgrader til "error" (forsøgsfejls varians)

Løsning:

- Den underliggende struktur skal altid indeholde de faktorer, som har flest mulige vekselvirkninger med andre. Dette er i dette tilfælde T_1, T_3 og T_5 .

Den mest "aktive" faktor med A, og de øvrige med B, C osv.

Vi foretager følgende omdøbning: $A = T_1, B = T_3, C = T_5, D = T_2$ og $E = T_4$.

Der skal i alt estimeres 5 hovedvirkninger og 3 vekselvirkninger AB, AC, BC.

Da blokkene er på 4 delforsøg, må vi så også indlægge 4 blokke. Dette koster 3 frihedsgrader, dvs. i alt $5 + 3 + 3 = 11$ frihedsgrader.

Vi prøver derfor om det er muligt at konstruere en konfunderet $\frac{1}{2}2^5$ -faktorstruktur.

Vi indfører nu eksempelvis den "nye" faktor E ved aliasrelationen $E = ABCD$ (som jo ikke er beslåglagt). Endvidere konfunderer vi ABD og CD (og dermed også ABC).

Vi opskriver den beskrevne struktur i følgende skema:

	x = Beslåglagt virkning k = konfunderet virkning		x = Beslåglagt virkning k = konfunderet virkning
		D	x
A	x	AD	
B	x	BD	
AB	x	ABD	k
C	x	CD	k
AC	x	ACD	
BC	x	BCD	
ABC	k	ABCD=E	x

Vi har dermed vist, at ovenstående model kan anvendes.

- Man vælger 3 (tilfældigt) fortegnene ++ til blok 1 + - til blok 2 osv.

Forsøgets behandlinger og fordeling på blokke er

	E=ABCD	behandlinger	ABD	CD	Blokke
(1)	+	e	-	+	3
a	-	a	+	+	1
b	-	b	+	+	1
ab	+	abe	-	+	3
c	-	c	-	-	4
ac	+	ace	+	-	2
bc	+	bce	+	-	2
abc	-	abc	-	-	4
d	-	d	+	-	2
ad	+	ade	-	-	4
bd	+	bde	-	-	4
abd	-	abd	+	-	2
cd	+	cde	+	+	1
acd	-	acd	-	+	3
bcd	-	bcd	-	+	3
abcd	+	abcde	+	+	1

- c Da der i alt skal estimeres 7 hovedvirkninger, 1 vekselvirkning og 4 blokke bliver antal frihedsgrader $7 + 1 + 3 = 11$ dvs der er $15 - 11 = 4$ frihedsgrader til error ◆

8.8 Beregning af et partielt forsøg

Lad i det følgende signifikansniveauet være 5%.

Eksempel 8.9. Beregning af partiel faktorforsøg

Virkningerne af 7 faktorer ønskedes undersøgt ved et partielt faktorforsøg. Om 3 af faktorerne kunne forudsættes, at kun hovedvirkninger kunne være forskellige fra nul, medens for de 4 øvrige også 2-faktorvekselvirkninger eventuelt kunne være forskellige fra nul. Disse faktorer benævnes A,B,C og D, og man skal undersøge om en eller flere af vekselvirkningerne AB, AC, AD, BC, BD, CD kunne forekomme. De tre øvrige faktorer benævnes med bogstaverne E,F og G.

Der udførtes et fuldstændigt randomiseret forsøg med en faktorstruktur, fremkommet ved, at faktorerne E,F og G indførtes i en fuldstændig 2^4 -faktorstruktur med faktorerne A,B,C og D ved relationerne: $E = ABC$ $F = BCD$ $G = ABCD$

Behandlingerne anført i standardrækkefølge efter A,B, C og D, og forsøgsresultaterne var følgende:

(1)	A	B	AB	C	AC	BC	ABC	D	AD	BD	ABD	CD	ACD	BCD	ABCD
g	ae	bef	abfg	cef	acfg	bcg	abce	df	adefg	bdeg	abd	cdeg	acd	bcd	abcdfg
15.3	18.4	26.1	26.3	13.5	15.7	18.8	17.3	21.0	22.3	18.9	15.5	9.6	10.5	23.1	25.0

- Find, hvilke faktorer, der har virkning
- Find de niveauer de pågældende faktorer skal indstilles på, for at give det største resultat.
- Angiv et estimat for dette største middelresultat, og angiv et 95% Konfidensinterval herfor.

Løsning

Vi skal undersøge hovedvirkningerne + vekselvirkningerne AB, AC, AD, BC, BD, CD

- a) Doe ► Classical ► Screening Design ► under Y skal stå "Maximize" ► Under Factors vælg "Add N Factors", Skriv 7 ► Vælg Categorical: 2 level ► ADD ► ændre navnene fra x1, x2 ... til A, B, ... ► Continue ► Continue ► Vælg 16 Run, Fractional Factorial (uden blokke) ► Continue
Vi ser nu følgende forslag til struktur (efter at have valgt nedenfor nævnte boks)

Factors

Continuous Discrete Numeric Categorical Remove Add N Factors 7

Name	Role	Values
A	Categorical	L1 L2
B	Categorical	L1 L2
C	Categorical	L1 L2
D	Categorical	L1 L2
E	Categorical	L1 L2

Display and Modify Design

Change Generating Rules

Factors	E	F	G
A	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
B	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
C	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
+/-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Change Generating Rules

Factors	E	F	G
A	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
B	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
C	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
D	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
+/-	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Apply

Aliasing of Effects

Effects	Aliases
A	= F*G
D	= E*G
E	= D*G
F	= A*G
G	= A*F = D*E
A*B	= C*E
A*C	= B*E
A*D	= E*F
A*E	= B*C = D*F
B*D	= C*F
B*F	= C*D

Skemaet skal læses $E = BCD$, $F = ACD$, $G = ABD$

I "Changing Generating Rules" ændres nu planen ved at sætte passende krydser. ► Apply

Vi får planen

Det ses, at planen opfylder kravene, idet vi dog skal flytte BC og CD frem

Vi går til "Output Options" ► Randomize (eller vælg Keep The Same) ► Make Tabel

Der fremkommer en tabel som skal udfyldes (se nedenfor)

Vælges "Randomize" står behandlingerne i tilfældig rækkefølge, så man kunne benytte den til at sende den til de der skal lave forsøgene. De skal så laves i den angivne rækkefølge .

Vælges "Keep the Same" er tabellen hurtigere at udfylde.

8. Flere end to forsøg, Screeningsforsøg

Marker Model ► Højre musetast ► Edit

Vi flytter nu BC og CD frem ved at skrive B*C fremfor A*E og C*D fremfor B*F. samt sletter de to sidste vekselvirkninger BG og CG, Run

Vi er nu tilbage ved Model Specifikation, hvor vi ses, at den endelige model stemmer SAS.JMP laver en tabel som skal udfyldes.

Vi udfylder tabellen

Pattern	A	B	C	D	E	F	G	Y
-----+	L1	L1	L1	L1	L1	L1	L2	15,3
+----+-	L2	L1	L1	L1	L2	L1	L1	18,4
--+++-+	L1	L1	L2	L2	L2	L1	L2	9,6
-+-++-+	L1	L2	L1	L2	L2	L1	L2	18,9
+++++++	L2	L2	L2	L2	L2	L2	L2	25
+--++++	L2	L1	L1	L2	L2	L2	L2	22,3
++---++	L2	L2	L1	L1	L1	L2	L2	26,3
-++---+	L1	L2	L2	L1	L1	L1	L2	18,8
--+-++-	L1	L1	L2	L1	L2	L2	L1	13,5
++-+---	L2	L2	L1	L2	L1	L1	L1	15,5
+++ +--	L2	L2	L2	L1	L2	L1	L1	17,3
---+--+	L1	L1	L1	L2	L1	L2	L1	21
+--++-	L2	L1	L2	L1	L1	L2	L2	15,7
-+++--+	L1	L2	L2	L2	L1	L2	L1	23,1
+--++--	L2	L1	L2	L2	L1	L1	L1	10,5
-+---++-	L1	L2	L1	L1	L2	L2	L1	26,1

Forsøget skal nu analyseres. Vi får

Rød pil i tabel ved Design ► Model ► Run Script ► *Emphasis vælg minimal report Run*

Blandt mange andre udskrifter fås følgende tabeller:

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Ratio	
Model	13	395,54313	30,4264	15,0579	
Error	2	4,04125	2,0206		Prob > F
C. Total	15	399,58438		0,0639	

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
A	1	1	1,38062	0,6833	0,4954
B	1	1	124,88063	61,8030	0,0158*
C	1	1	57,38063	28,3975	0,0335*
D	1	1	1,89063	0,9357	0,4354
E	1	1	1,50062	0,7427	0,4796
F	1	1	148,23063	73,3588	0,0134*
G	1	1	2,64062	1,3068	0,3714
A*B	1	1	6,63062	3,2815	0,2118
A*C	1	1	0,33062	0,1636	0,7250
A*D	1	1	0,68062	0,3368	0,6203
B*C	1	1	39,37563	19,4869	0,0477*
B*D	1	1	2,64063	1,3068	0,3714
C*D	1	1	7,98063	3,9496	0,1852

Heraf ses, at BC er den eneste signifikante vekselvirkning, da for BC er P-værdi=0.0477<0.05. De 5 øvrige vekselvirkninger pooler ned i error. Dette er vigtigt, da error kun har 2 frihedsgrader, så en forøgelse til 7 giver sikrere resultater.

De øvrige vekselvirkninger pooler ned i “Error”, ved at slette dem i Design ► Model ► Edit ► Run

Vi får nu følgende udskrift

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
A	1	1	1,38062	0,4333	0,5314
B	1	1	124,88063	39,1925	0,0004*
C	1	1	57,38063	18,0083	0,0038*
D	1	1	1,89063	0,5934	0,4663
E	1	1	1,50062	0,4710	0,5146
F	1	1	148,23063	46,5207	0,0002*
G	1	1	2,64063	0,8287	0,3929
B*C	1	1	39,37563	12,3576	0,0098*

Vi interesserer os ikke for B og C, da de jo indgår i vekselvirkningen, men ser, at blandt de øvrige hovedvirkninger har kun F en signifikant virkning.

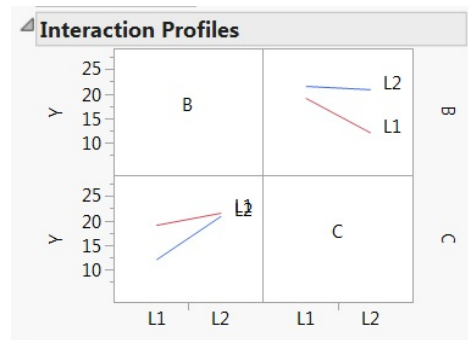
Konklusion: B, C og F har en virkning B og C i form af en vekselvirkning.

b) De øvrige hovedvirkninger pooler ned i “Error”. Igen vigtigt for at øge frihedsgradstallet.

Vælg derefter minimal Report, Run

Rød pil ved “Response Y” ► “Factor Profiling” ► vælg “Interaction Plots”

Man kan nu finde hvilke niveauer faktorerne skal sættes på for at få størst resultat, ved at se på følgende figur.



Lad os betragte tegningen øverst til højre.

Tegningen skal forstås på den måde, at nederst ses C på lavt niveau L1 og højt niveau L2 Til venstre hvor der står B skal det forstås, at L1 linien (rød) er B på lavt niveau og L2 linien (blå) er B på højt niveau

Det ses, at den største værdi ligger øverst til venstre, på blå linie. Det punkt er B_2C_1 , dvs. B på højt niveau og C på lavt niveau.

Man kan dog se, at punkterne B_2C_2 og B_1C_1 ligger næsten lige så højt, så sandsynligvis er den eneste man ikke skal vælge er B_1C_2

Betragtes tegningen nederst til venstre, så aflæses B fornedet, og rød linie er C på lavt niveau. Heraf aflæses igen at B skal være på højt niveau, og C på rød linie er L1, dvs. på lavt niveau.

Konklusion: C på lavt niveau og B på højt niveau (en mere detaljeret undersøgelse med flere forsøg er nok nødvendigt)

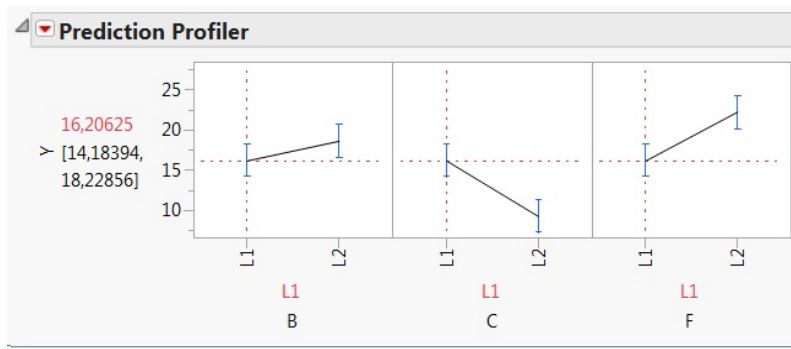
Nu skal man se hovedvirkninger

Rød pil ved “Response Y” ► “Factor Profiling” ► vælg “Profiler”

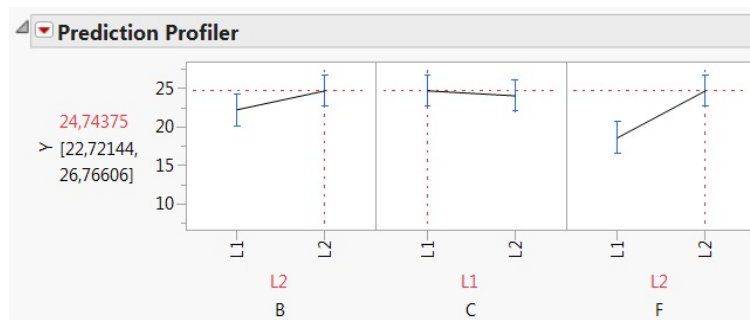
Under rød pil “Prediction Profiler” slet kryds ved “Desirability Functions”

Af Prediction profiles ses, at man skal vælge F på højt niveau. (L2)

8. Flere end to forsøg, Screeningsforsøg



c) Man kan nu rykke på de røde linier, så de står på B højt, C lavt og F højt.



Det ses, at den største middelværdi er 24.74,
og et 95% konfidensinterval er [22.72 ; 26.76]



Eksempel 8.10. Beregning af et konfunderet partielt 2^{5-1} faktorforsøg

(fortsættelse af eksempel 8.8)

Virkningen af 5 faktorer A, B, C, D og E ønskes undersøgt. Man ved, at ingen af faktorerne D og E vekselvirker med andre faktorer. Da man kun kan udføre 4 forsøg pr. apparat, indføres blokke på 4 forsøgsenheder.

I en fuldstændig 2^4 struktur med faktorerne A, B, C og D indføres $E = ABCD$. Endvidere indføres blokkene ved at konfundere ABD, CD og ABC.

Her vælges (tilfældigt) fortegnene ++ til blok 1 + - til blok 2 osv.

Forsøgsplanen (opskrevet på sædvanlig måde i standardorden) og forsøgsresultaterne er:

behandling	e	a	b	abe	c	ace	bce	abc	d	ade	bde	abd	cde	acd	bcd	abcde
blok	3	1	1	3	4	2	2	4	2	4	4	2	1	3	3	1
resultat	9	16	11	13	10	14	6	17	11	14	7	14	9	16	8	5

- Find hvilke faktorer der har virkning
- Angiv de niveauer de pågældende faktorer skal indstilles på, for at give det største middelresultat.

Løsning:

- Doe ► Screening Design ► under Y skal stå "Maximize" ► Under Factors vælg "Add N Factors", Skriv 5 ► Vælg Catagorical: 2 level ► ADD ► ændre navnene fra x1, x2 ... til A, B, ... ► Continue ► Vælg 16 Run, block 4, Fractional Factorial ► Continue
Vi får en udskrift
Vælg "Change Generating Rules"

Vi ændrer nu planen ved at sætte og fjerne passende krydser i overensstemmelse med ovenstående plan.

► Apply

Change Generating Rules

Factors	E	Block	Block
A	X	X	
B	X	X	
C	X		X
D	X	X	X

Aliasing of Effects

Effects	Aliases	Block Aliases
C*D		= Block
C*E		= Block
D*E		= Block

Vi ser planen stemmer

“Output Options” ► Randomize within blocks. ► Make Tabel

Der udskrives nu en tabel, efter hvilken rækkefølgen af forsøgene er bestemt (randomiseret)

Vi indsætter nu forsøgsresultaterne:

Bloknumrene svarer ikke til dem vi indførte, men da man jo ikke er interesseret i hvilke blokke der er “bedst”, er det ligegyldigt. Eksempelvis svarer SAS blok 1 til vor nr 4

Pattern	Block	A	B	C	D	E	Y
-- + --	1	L1	L1	L2	L1	L1	10
- + - ++	1	L1	L2	L1	L2	L2	7
+++ --	1	L2	L2	L2	L1	L1	17
+ - - ++	1	L2	L1	L1	L2	L2	14
--- + -	2	L1	L1	L1	L2	L1	11
- ++ - +	2	L1	L2	L2	L1	L2	6
+ - + - +	2	L2	L1	L2	L1	L2	14
++ - +-	2	L2	L2	L1	L2	L1	14
---- +	3	L1	L1	L1	L1	L2	9
+ - ++ -	3	L2	L1	L2	L2	L1	16
++ - - +	3	L2	L2	L1	L1	L2	13
- +++ -	3	L1	L2	L2	L2	L1	8
-- +++	4	L1	L1	L2	L2	L2	9
+ ----	4	L2	L1	L1	L1	L1	16
- + ---	4	L1	L2	L1	L1	L1	11
+++++	4	L2	L2	L2	L2	L2	5

Analyse af data (variationsanalyse)

Rød pil i tabel ved Design, Model ► Edit

I den fremkomne tabel fjernes nu alle de vekselvirkninger, man mener der er 0.

Rød pil i tabel ved Design, Model ► Run Script

Man ser nu den endelige model.

Vi får bl.a. følgende udskrift.

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	11	179,50000	16,3182	2,0722
Error	4	31,50000	7,8750	Prob > F
C. Total	15	211,00000		0,2516

8. Flere end to forsøg, Screeningsforsøg

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	11,25	0,701561	16,04	<,0001*
A	2,375	0,701561	3,39	0,0276*
B	-1,125	0,701561	-1,60	0,1841
C	-0,625	0,701561	-0,89	0,4233
D	-0,75	0,701561	-1,07	0,3453
E	-1,625	0,701561	-2,32	0,0815
Block[1]	0,75	1,215139	0,62	0,5705
Block[2]	0	1,215139	0,00	1,0000
Block[3]	0,25	1,215139	0,21	0,8470
A*B	-0,25	0,701561	-0,36	0,7396
A*C	0	0,701561	0,00	1,0000
B*C	-0,5	0,701561	-0,71	0,5154

Det ses, at alle vekselvirkninger er 0.

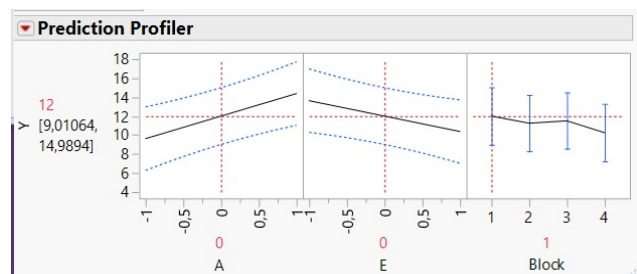
De pooles væk.

Marker dem under "Model "Edit". ok og vælg Model Run script, run

Konklusion: A og E har en virkning

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	11,25	0,57087	19,71	<,0001*
A	2,375	0,57087	4,16	0,0042*
B	-1,125	0,57087	-1,97	0,0894
C	-0,625	0,57087	-1,09	0,3098
D	-0,75	0,57087	-1,31	0,2303
E	-1,625	0,57087	-2,85	0,0248*
Block[1]	0,75	0,988776	0,76	0,4729
Block[2]	0	0,988776	0,00	1,0000
Block[3]	0,25	0,988776	0,25	0,8077

- b) B, C og D pooles ned i "error" hvorved vi får den endelige model. Derefter
 Rød pil ved "Response Y" ► Factor Profiling ► Profiler ► Rød pil ved "Prediction Profiler"
 Slet desirability functions
 Der fremkommer følgende figur



A skal på højt niveau, og E skal på lavt niveau for at give det største middelværdi

8.9. SEKVENTIEL FORSØGSSTRATEGI

I en række situationer vil man selv med partielle 2^k - faktorforsøg og med rimelige forhåndsforudsætninger om hvilke vekselvirkninger der er nul få mange behandlinger og dermed mange delforsøg. Er delforsøgene dyre og budgettet begrænset, kan man prøve at formindske antallet ved at udføre færre forsøg end nødvendigt til at estimere alle virkninger, og dog håbe at nå et fornuftigt resultat.

Man sammenblander eksempelvis to vekselvirkninger $AB=CD$ uden at vide om en af dem er 0. Hvis man så opdager at den sammenblandede faktor $AB+CD$ ikke er signifikant, så vil den enkleste model være, at ingen af dem har en virkning. Det behøver naturligvis ikke at være tilfældet, men det er en erfaring, at man ofte kan få en korrekt fortolkning af forsøgsresultaterne ved at vælge den, som giver den enkleste model. (den simpleste models princip).

8.10 OVERSIGT OVER FREMGANGSMÅDE VED PARTIELT 2^k FAKTORFORSØG.

- 1) Faktorerne omdøbes, så den faktor der indgår i flest 2-faktorvekselvirkninger kaldes A, næstflest kaldes B osv.
- 2) Antallet af faktorer og 2-faktorvekselvirkninger, der ønskes testet, optælles.
Lad antallet være t .
- 3) $t < 7$: Start med en 2^{t-1} faktorstruktur (8 delforsøg)
 $7 \leq t < 15$: Start med en 2^{5-1} faktorstruktur (16 delforsøg) osv.
- 4) Den grundlæggende struktur opskrives, og man markerer hvilke af disse der skal undersøges
- 5) De resterende faktorer indføres, idet man opskriver aliasstrukturene, og dermed sikrer, at der ikke sker sammenblanding af virkninger.
Er det ikke muligt at undgå, må man fordoble antallet af forsøg (eksempelvis gå fra 8 til 16 forsøg).
- 6) Behandlingerne bestemmes ved at lave et fortegnsskema
(kontrol: Alle faktorer skal være lige mange gange på højt og lavt niveau)
- 7) Forsøgene udføres (randomiseret)
I opgaver bliver de opgivet, ofte i standardorden
- 8) Beregningerne foretages via SAS.JMP (se eksempel 8.10 og 8.11)

OPGAVER

Opgave 8.1

Ved en produktion ønskedes virkningen af 3 faktorer undersøgt:

A: Maskiner (3 forskellige anvendes) B: Benyttet halvfabrikata (2 muligheder)

C: Produktionstidspunkt (2 muligheder)

Der var på forhånd en vis formodning om, at tidsfaktoren ingen virkning har.

Følgende observationer af produktionsudbyttet foretoges (kodede tal).

		Faktor C			
		C ₁		C ₂	
		B ₁	B ₂	B ₁	B ₂
Faktor A	A ₁	35	24	32	27
		37	28	38	30
		38	26	36	24
		37	24	35	26
	A ₂	42	36	38	33
		40	35	43	35
		39	36	42	39
		44	37	43	34
	A ₃	42	41	50	40
		47	47	43	44
		48	40	48	39
		49	48	47	50

Det antages, at forudsætningerne for at udføre en variansanalyse er opfyldt.

- 1) Undersøg hvilke af de 3 faktorer, der har virkninger eventuelt i form af vekselvirkninger.
- 2) Angiv hvilke niveauer faktorerne skal stilles på for at give det størst mulige udbytte.
- 3) Angiv et estimat for det største udbytte, og et 95% konfidensinterval herfor.

Opgave 8.2

Virkningerne af 6 faktorer T_1, T_2, T_3, T_4, T_5 og T_6 ønskes undersøgt ved et screeningsforsøg. Det vides, at kun hovedvirkningerne og 2-faktorvekselvirkningerne

$T_1T_2, T_1T_3, T_1T_6, T_2T_3, T_2T_6, T_4T_5$ kan være forskellig fra nul.

Foretag en hensigtsmæssig omdøbning til A, B osv. og opstil derefter en forsøgsplan med færrest mulige delforsøg. Angiv herunder

- a) Den underliggende fuldstændige faktorstruktur, og de relationer hvormed nye faktorer er indført i den underliggende struktur
- b) Forsøgets behandlinger.
- c) Angiv antal frihedsgrader for forsøgsfejlen (error)

Opgave 8.3.

Virkningen af 8 faktorer $T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8$ ønskes undersøgt ved et screeningsforsøg.

Det vides på forhånd, at faktoren T_2 muligvis vekselvirker med faktorerne T_1, T_3 og T_7 samt at T_4 muligvis vekselvirker med T_7 . Alle andre vekselvirkninger er 0.

Foretag en hensigtsmæssig omdøbning til A, B osv. og opstil en forsøgsplan med færrest mulige delforsøg.

8. Flere end to forsøg, Screeningsforsøg

Angiv herunder

- Den underliggende fuldstændige faktorstruktur, samt relationerne hvormed nye faktorer er indført i den underliggende struktur.
- Forsøgets behandlinger
- Angiv frihedsgraden for forsøgsfejlen (error)

Opgave 8.4

Virkningerne af 6 faktorer T_1, T_2, T_3, T_4, T_5 og T_6 ønskes undersøgt ved et screeningsforsøg. Man har kun råd til at lave 16 delforsøg, og det er af tidsmæssige grunde nødvendigt at arbejde med 4 blokke på 4 forsøgsenheder. Det vides at kun hovedvirkninger og 2 - faktorvekselvirkningerne T_4T_1, T_4T_2, T_4T_3 og T_4T_5 kan være forskellig fra nul.

Foretag en hensigtsmæssig omdøbning til A, B osv. og opstil derefter en forsøgsplan med færrest mulige delforsøg. Angiv herunder

- Den underliggende fuldstændige faktorstruktur, samt relationerne hvormed nye faktorer er indført i den underliggende struktur
- Forsøgets behandlinger.
- Angiv antal frihedsgrader for forsøgsfejlen (error).

Opgave 8.5

Som led i en serie forsøg med henblik på optimering af styrken af et plastprodukt ønskedes ved et fuldstændigt 2^4 faktorforsøg virkningen af følgende faktorer undersøgt:

	Lavt niveau	Højt niveau
A: Materialetype	Type 1	Type 2
B: Fabrikationsmetode	Metode 1	Metode 2
C: Kemisk farvestof	Ikke tilsat	Tilsat
D: Opbevaring	Kort tid	Lang tid

Ved forsøget målttes styrken af produktet

Forsøgsplan og forsøgsresultater (kodede tal) var følgende (i randomiseret rækkefølge):

a	b	ab	bcd	ac	bc	(1)	ad	abc	acd	bd	c	cd	abd	abcd	d
30.9	39.8	50.2	26.9	26.0	30.4	29.2	23.5	28.4	21.0	40.0	23.9	20.4	50.5	24.7	24.3

Det forudsættes, at kun hovedvirkninger og 2-faktorvekselvirkninger kan være forskellig fra nul.

- Hvilke faktorer har en signifikant virkning (indgår eventuelt i en vekselvirkning?)
- Bestem, idet man ønsker den størst mulige middelværdi, de(n) optimale kombination(er) af de 4 faktorerers niveauer.

Opgave 8.6

Virkningen af 5 faktorer A, B, C, D og E, som ikke vekselvirker, skal undersøges ved et fuldstændigt randomiseret forsøg.

Faktorerne D og E indføres ved aliasrelationerne $D = ABC$ og $E = BC$.

Forsøgsplanen (opskrevet på sædvanlig måde i standardorden) og forsøgsresultaterne er:

	I	A	B	C	D = ABC	E = BC	behandlinger	resultat
(1)	+	-	-	-	-	+	e	7
a	+	+	-	-	+	+	ade	19
b	+	-	+	-	+	-	bd	25
ab	+	+	+	-	-	-	ab	15
c	+	-	-	+	+	-	cd	54
ac	+	+	-	+	-	-	ac	46
bc	+	-	+	+	-	+	bce	12
abc	+	+	+	+	+	+	abcde	31

- Hvilke faktorer har en signifikant virkning
- Angiv de niveauer de pågældende faktorer skal indstilles på, for at give det største middeludbytte, og angiv et 95% konfidensinterval for dette middeludbytte.

Opgave 8.7.

Virkningen af 4 faktorer A, B, C og D påtænkes undersøgt ved et fuldstændigt randomiseret forsøg med en partikulær 2^k faktorstruktur. Det antages, at bortset fra hovedvirkninger, kan der kun tænkes at forekomme 2-faktorvekselvirkning AC.

Der udføres et 2^{5-2} faktorforsøg med aliasrelationerne $D = ABC$.

Opstillet i standardrækkefølge (efter den underliggende fuldstændige faktorstruktur i faktorerne A, B og C) er forsøgsresultaterne på responsvariablen Y:

(1)	ad	bd	ab	cd	ac	bc	abcd
77.2	74.4	80.9	67.3	70.3	75.0	62.1	86.7

- Hvilke faktorer har en signifikant virkning (indgår eventuelt i en vekselvirkning?), idet signifikansniveauet sættes til 5%.
- Bestem, idet man ønsker den størst mulige middelværdi, de(n) optimale kombination(er) af de 4 faktorerers niveauer, samt et 95% konfidensinterval.

Opgave 8.8

Virkningen af 6 faktorer A, B, C, D, E og F ønskes undersøgt. Det antages, at bortset fra hovedvirkninger, kan der kun tænkes at forekomme 2-faktorvekselvirkning AC.

Da man kun kan udføre 4 forsøg pr. apparat, indføres blokke på 4 forsøgsheder.

I en fuldstændig 2^4 struktur med faktorerne A, B, C og D indføres $E = ABCD$ og $F = ACD$.

Endvidere indføres blokkene ved at konfundere ABC og CD.

Opstillet i standardrækkefølge (efter den underliggende fuldstændige faktorstruktur i faktorerne A, B, C og D) er forsøgsresultaterne på responsvariablen Y:

e	af	b	abef	cf	ace	bcef	abc	df	ade	bdef	abd	cde	acdf	bcd	abcdef
27.9	43.7	24.3	46.1	37.2	13.3	20.7	29.6	27.5	25.0	30.3	20.7	17.4	52.8	35.2	36.6

- Hvilke faktorer har en signifikant virkning.
- Bestem, idet man ønsker den størst mulige middelværdi, de(n) optimale kombination(er) af de

8. Flere end to forsøg, Screeningsforsøg

4 faktorerers niveauer.

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 støjkluder. 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 $y = a + bx^e$.

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..

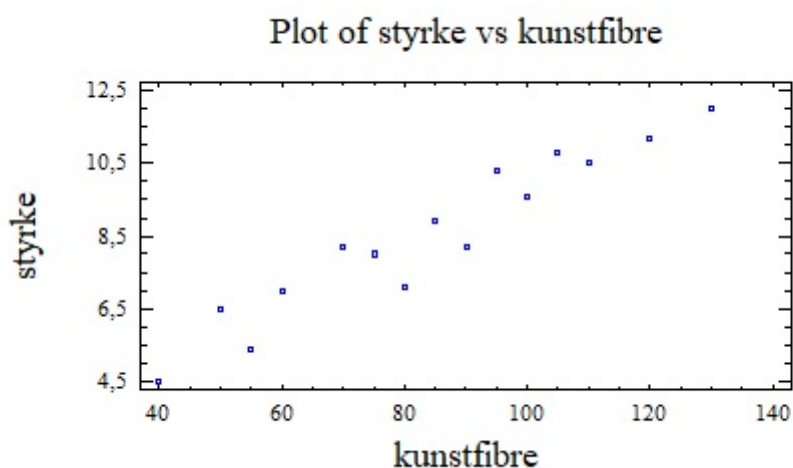
Som eksempler herpå kan nævnes $y = a + bx$ og $\ln y = a + b \cdot \ln x$.

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**.

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 taleksempel. Man benytter hertil et matrixprogram. De angivne metoder kan umiddelbart generaliseres til mere komplicerede eksempler.

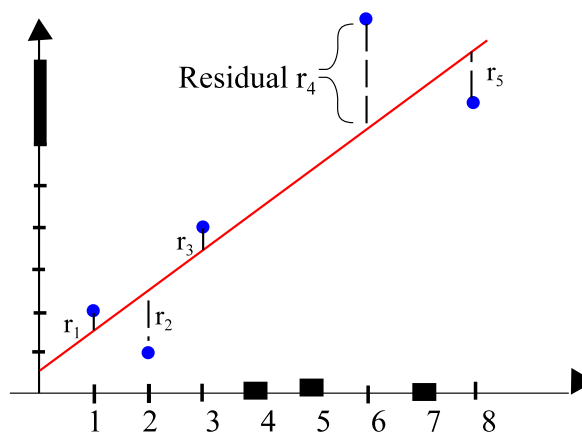
Eksempel 9.2. Bestemmelse af regressionskoefficienter ved mindste kvadraters metode.

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).



På figur 9.1 er afsat de 5 punkter, og indtegnet en ret linie.

Figur 9.1 Residualer

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

$$B = \left(\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 6 & 8 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 6 \\ 1 & 8 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 6 & 8 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 4 \\ 9 \\ 7 \end{pmatrix} = \begin{pmatrix} 0.6 \\ 1.0 \end{pmatrix}$$

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 som SAS.JMP, 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**.

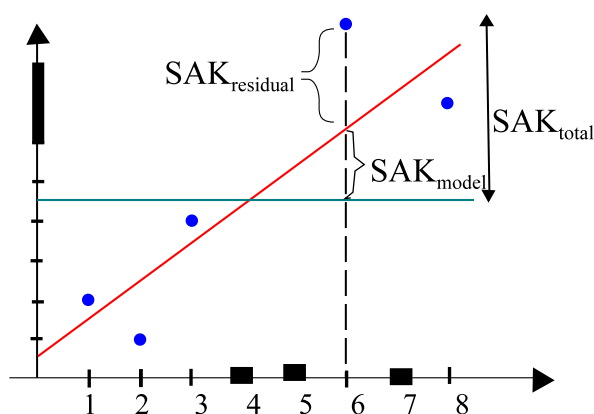
Outliers er en betegnelse for enkelte målinger som afviger så kraftigt fra den almindelige tendens, at det kan skyldes fejlmålinger. Sådanne punkter kan i uheldige tilfælde på grund af et stort bidrag til residualsommen få regressionslinien til at dreje. Det er dog klart, at man ikke blot kan stryge 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 $y = \bar{y}$

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"

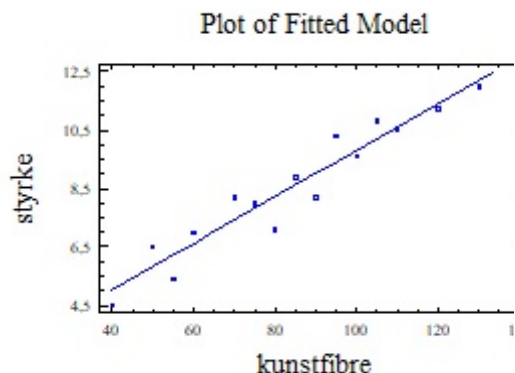
I eksempel 9.2 forklarer den fundne model således "75.2% af den totale variation.

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.

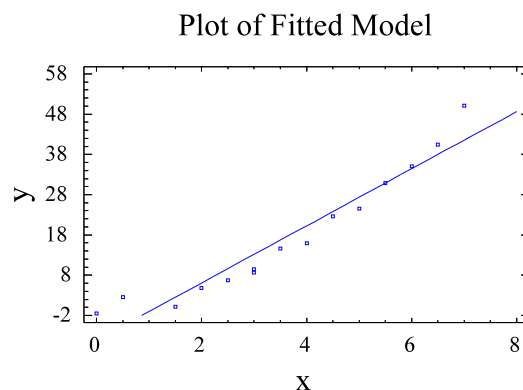
At man ikke alene kan stole på forklaringsgraden illustreres ved følgende eksempel.

Eksempel 9.3 .Grafisk vurdering af model.

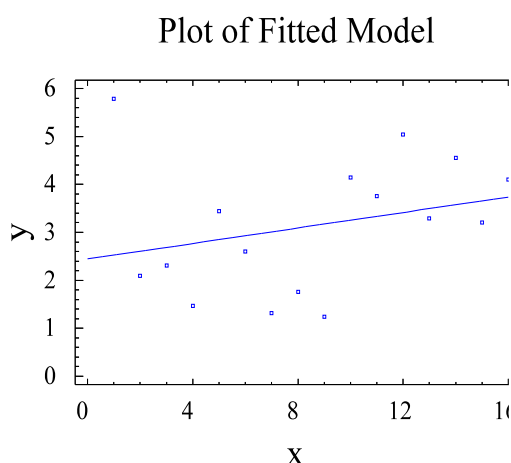
De følgende 4 figurer afspejler forskellige muligheder.



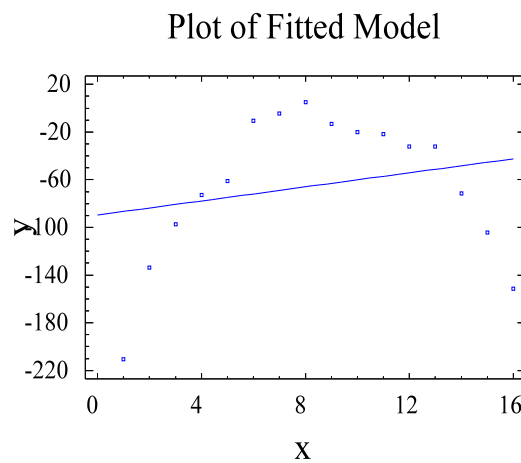
Figur 9.2a: $r = 0.959$ $r^2 = 91.9\%$



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).

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 tegning med indlagt ret linie (eller residualernes fortegn) til vurdering af modellen.
- 2) Opskriv regressionsligningen.

Løsning

1) Data indtastes

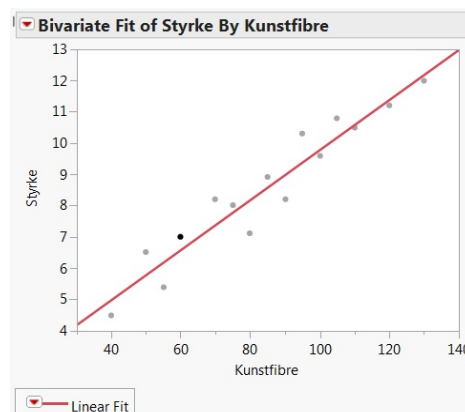
```
kunstfibre  styrke
40          4,5
50          6,5
osv.
```

Man kan ved analysen vælge 2 modeller, enten "fit Y by X " som giver en forholdsvis simpel og overskuelig analyse, eller "Fit model" som er nødvendig ved mere specielle analyser.

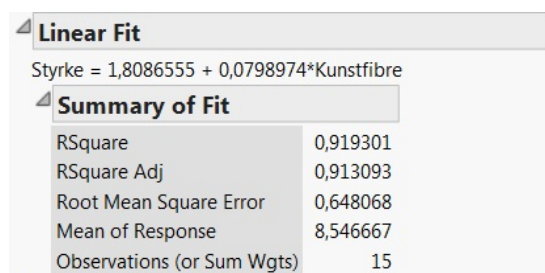
Vælg Analyze ► Fit Y by X ► markér "Styrke" og tryk på Y Response ► markér "kunstfibre" og tryk på X Factor ► OK
Der fremkommer et "scatterplot", hvor man i et koordinatsystem kan se punkterne afbildet.
Rød pil ► vælg fra rullemenuen "Fit Line".

Der fremkommer så følgende tegning og udskrift:

Af figuren ses, at punkterne fordeler sig tilfældigt omkring linien.



Af udskriften ses, at forklaringsgraden “RSquare” er 91,93 % , hvilket er tilfredsstillende, da modellen altså “forklarer” 91,93% af variationen.



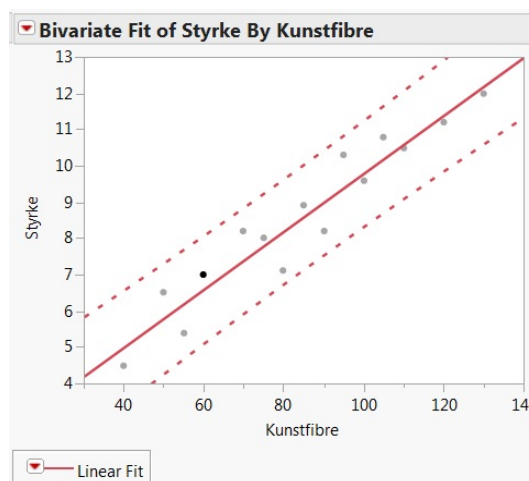
Outliers. Af ovenstående figur ses, at der næppe er nogen “outliers” (punkter der afviger så kraftigt fra det generelle billede, at man kunne frygte de var fejlmålinger).

En undersøgelse af om der er outliers er vigtigt.

En (lidt usikker) metode er, at få tegnet såkaldte 95% predikationskurver, og se om praktisk taget alle punkter ligger indenfor disse.

Rød pil under tegningen ved “linear fit” ► vælg “Confid. Curves indv”

Da alle punkter ligger indenfor grænserne, tyder det ikke på, at der er outliers.



En lidt sikrere metode er, at lade SAS-JMP beregne såkaldte “Studentized Residuals”, som tager i betragtning, at spredningen er mindre ved “midtpunktet” end langt fra det.

Vælg Analyze ► Fit model ► markér “Styrke” og tryk på Y ► markér “kunstfibre” og tryk ADD ► I “Emphasis” vælg “Minimal Report” (for at begrænse udskrifterne) ► Run

Vi får tegninger og udskrifter nogenlunde magen til før.

Rød pil ► vælg Save Columns ► Studentized residuals

Der bliver nu tilføjet en ekstra kolonne til data

Heraf fremgår, at da ingen “Studentized Residuals”, numerisk er større end 3 (det er tilladt, at nogle få er større end 2) er der ingen outliers.

Af fortegnene på Studentized Residuals kan man se at de har ret skiftende fortegn hvilket jo (måske lettere) viser at punkterne fordeler sig tilfældigt omkring linien.

Konklusion: Modellen synes tilstrækkelig godt at beskrive data indenfor måleområdet.

Kunstfibre	Styrke	Studentized Resid Styrke
40	4,5	-0,910214041
50	6,5	1,1923782679
55	5,4	-1,347927857
60	7	0,6566305986
70	8,2	1,2900712404
75	8	0,3194449843
80	7,1	-1,759454913
85	8,9	0,479282925
90	8,2	-1,279096524
95	10,3	1,4483232937
100	9,6	-0,321237896
105	10,8	0,9851242793
110	10,5	-0,161471825
120	11,2	-0,33813566
130	12	-0,355350614

- 2) Under “Linear Fit” ses regressionsligningen Styrke = 1.8087 + 0.0799*kunstfibre.



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 101)

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 prædiktionsintervaller.

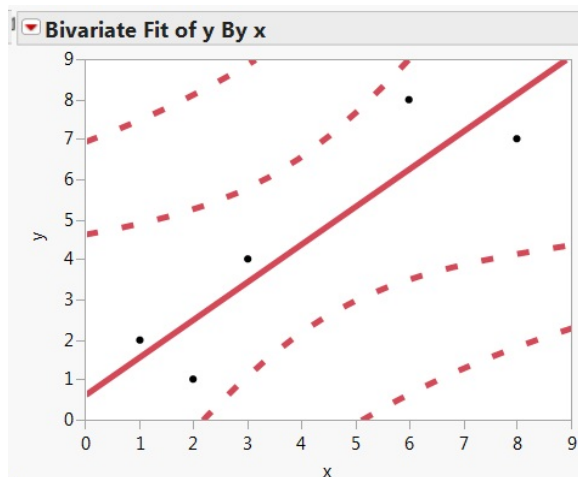
Et led i analysen kan være, at udregne et 95% konfidensinterval for β_1 .

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% prædiktionsinterval for én ny observation.

På figur 9.4 er her tegnet 95% konfidensintervaller for middelværdierne (de inderste buede kurver), og 95% prædiktionsintervaller (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.



Figur 9.4 Konfidens- og prædiktionsintervaller

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:

Data er indtastet som i eksempel 9.4

- 1) $H_0: Y$ er uafhængig af $x \Leftrightarrow H_0: \text{Regressionslinien er vandret} \Leftrightarrow H_0: \beta_1 = 0$.

Vælg Analyze ► Fit model ► markér "Styrke" og tryk på Y ► markér "kunstfibre" og tryk på Add ► Run
Der fremkommer så blandt andet denne udskrift:

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	62,197436	62,1974	148,0919
Error	13	5,459897	0,4200	Prob > F
C. Total	14	67,657333		<,0001*

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	1,8086555	0,578421	3,13	0,0080*
Kunstfibre	0,0798974	0,006565	12,17	<,0001*

Det ses, ud for "Model", at

F - Ratio = 148.09 og at P -value = 0.0001

(kan også ses ud for "kunstfibre")

Heraf fås, at

H_0 forkastes

Y er ikke uafhængig af x .

- 2) Konfidensinterval for hældningskoefficienten β_1 :

Cursor placeres i tabel for "Parameter Estimates", højre musetast ► Columns ► "Lower 95%" ► "Upper 95%"

Man får bl.a. følgende tabel:

Parameter Estimates						
Term	Estimate	Std Error	t Ratio	Prob> t	Lower 95%	Upper 95%
Intercept	1,8086555	0,578421	3,13	0,0080*	0,5590522	3,0582587
Kunstfibre	0,0798974	0,006565	12,17	<,0001*	0,0657135	0,0940812

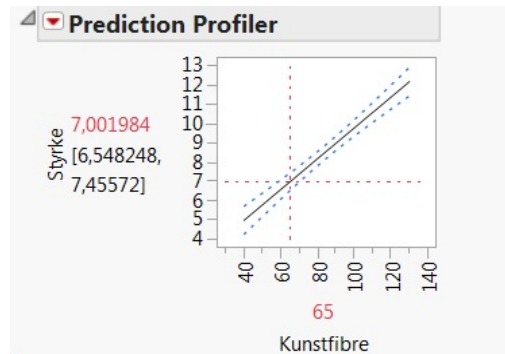
Heraf aflæses [0.0657 ; 0.0940]

9. Regressionsanalyse

- 3) 95% konfidensinterval for middeltrækstyrken svarende til x - værdien 65.
rød pil ved "Response, y " ► vælg "Factor Profiling" ► Profiler ► sæt cursor på tallet i bunden og skriv 65 og aflæs

$$x = 65 \quad Y = 7.002$$

$$95\% \text{ konfidensinterval} = [6.55; 7.46]$$



- 4) Indsæt 65 nederst i x-kolonne i tabel
rød pil ved response ► save Columns ► Individ Confidence Interval
I tabel fremkommer nu prædiktionsintervaller [5.53 ; 8.47]



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 .

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:

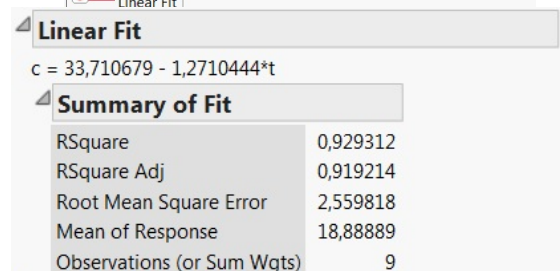
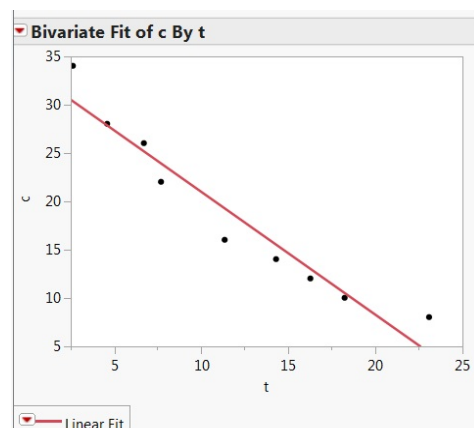
Data indtastes

t	c
2,67	34
4,59	28
6,75	26
7,67	22
11,34	16
14,34	14
16,25	12
18,25	10
23,09	8

- 1) Analyze ► Fit Y by X ► markér "c" og tryk på Y Response ► markér "t" og tryk på X Factor ► OK
Der fremkommer et "scatterplot", hvor man i et koordinatsystem kan se punkterne afbildet.
Rød pil ► Fit line

Der fremkommer følgende figur og udskrift:

Det ses, at punkterne ikke fordeler sig jævnt om linien

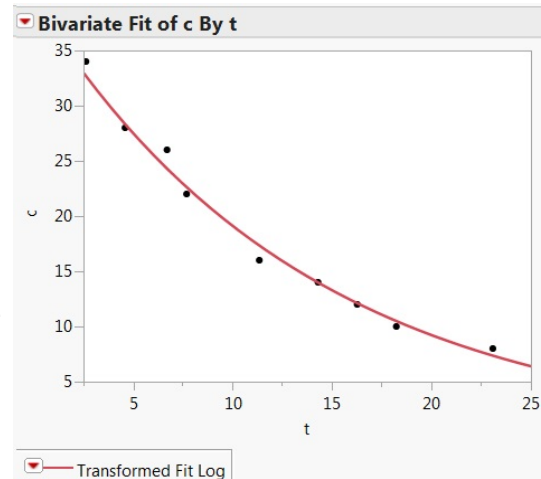


Forklaringsgraden 91.9% er høj

9. Regressionsanalyse

Vi gentager nu, idet vi nu vælger
Fit special ► Marker "Natural logarithm" for y ► OK

Vi ser, at punkterne fordeler sig tilfældigt omkring kurven,



og at RSquare = 0.988 er høj, så
model2 (eksponentiel model) må være den bedste model.

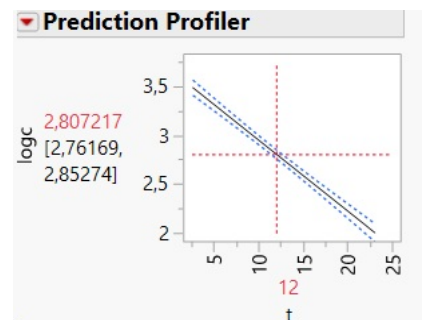
Transformed Fit Log	
Log(c) = 3,6780207 - 0,072567*t	
Summary of Fit	
RSquare	0,988291
RSquare Adj	0,986618
Root Mean Square Error	0,057679
Mean of Response	2,831809
Observations (or Sum Wgts)	9

2) Af $\text{Log}(c) = 3,6780207 - 0,072567*t$ fås (med regnemaskine)

$$c = e^{3,6780 - 0,07257 \cdot t} = \underline{\underline{39,57 \cdot e^{-0,07257 \cdot t}}}$$

3) Danner en ny kolonne med navnet logc ► cursor på navn, højre musetast, Formula ► transcental ► vælg c ► ln ► ok

Analyze ► Fit model ► Marker logc og vælg y ► marker t og vælg add ► Emphasis, Minimal report ► Run
Rød pil ved "Response Inc" ► Factor profiling ► Profiler
Der fremkommer en tegning, t-værdi sættes til 12



Heraf ses, at for $t = 12$ er $\text{log}(c) = 2,8072 \Leftrightarrow c = e^{2,8072} = \underline{\underline{16,56}}$

95% konfidensinterval: $[e^{2,76169}; e^{2,8527}] = \underline{\underline{[15,83; 17,33]}}$



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	7.4	8.8	13.6	13.1	14.9	20.0	23.1	27.5	32.9
	4.9	6.9	8.2	12.0	12.4	16.8	21.2	25.2	25.1	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 y	4.2	7.4	8.8	13.6	13.1	14.9	20.0	23.1	27.5	32.9
	4.9	6.9	8.2	12.0	12.4	16.8	21.2	25.2	25.1	32.4

- 1) Foretag en testning af forudsætningen om varianshomogenitet.
- 2) 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.
- 3) Bestem i bekræftende fald ligningen for den fundne regressionslinie.
- 4) Det påstås i litteraturen, at hældningskoefficienten β_1 er 0.15
Test om dette på et signifikansniveau på 5% kan være sandt.
- 5) Angiv et 95% konfidensinterval for middelværdien af tykkelsen y , når $t = 110$ minutter.

Løsning

Data indtastes på sædvanlig måde:

t	y
20	4,2
20	4,9
30	7,4
30	6,9
40	8,8
40	8,2
60	13,6

osv.

- 1) Undersøgelse af varianshomogenitet.

Test af nulhypotesen $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{10}^2$

Man gør som beskrevet under ensidet variansanalyse, dvs. vi vælger at gøre t til "nominal" (rød)

Vælg Analyze ► Fit Y by X ► markér "y" og tryk på Y Response ► markér "t" og tryk på X Factor ► OK
Sæt cursor på overskrift, højre musetast ► vælg fra rullemenuen "Unequal Variances".

Test	F Ratio	DFNum	DFDen	Prob > F
O'Brien[.5]	0,0000	-1	0	0,0000
Brown-Forsythe	-1,9e+16	9	10	0,0000
Levene	-1,9e+16	9	10	0,0000
Bartlett	0,4990	9	.	0,8763

Warning: Small sample sizes. Use Caution.

Da vi kun har 2 gentagelser for hver t-værdi kan kun Bartletts test anvendes.

Da P -værdien $= 0.8763 > 0.05$ accepteres H_0 , dvs. vi vil i det følgende antage, at kravet om varianshomogenitet er opfyldt.

2) H_0 : Lineær model gælder $\Leftrightarrow H_0: (x_i, \mu_i)$ ligger på en ret linie

Man sørger nu for, at t er "continuous", og vælger nu forfra

Analyze ► Fit Y by X ► markér "y" og tryk på Y Response ► markér "t" og tryk på X Factor ► OK

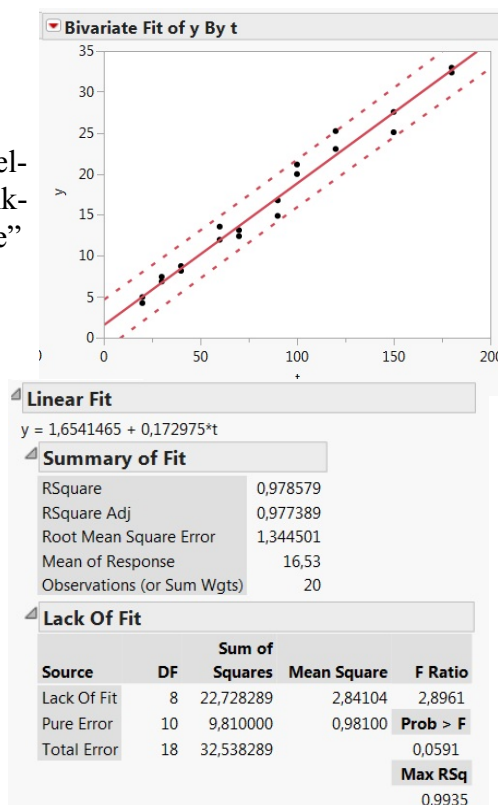
Der fremkommer et "scatterplot", hvor man i et koordinatsystem kan se punkterne afbildet.

Sæt cursor på overskrift, højre musetast ► vælg fra rullemenuen "Fit Line".

Tryk på den røde pil under tegningen ved "linear fit" ► vælg "Confid. Curves indv"

Der fremkommer følgende udskrift:

Af figuren ses, at "gennemsnitspunkterne" ligger tilfældigt omkring linien, og der næppe er outliers, da punkterne næsten alle falder indenfor "prediction linierne"



Det ses, at forklaringsgraden RSquare er høj

Af udskriften for "Lack of fit" ses, at P -value er 0.0591.

På et signifikansniveau på 5%, ses, at H_0 må accepteres, dvs. vi kan antage, at indenfor måleområdet giver førstegradsmodellen en rimelig god beskrivelse af resultaterne.

3) Af udskriften ses, at regressionsligningen bliver $\hat{y} = 16542 + 0.1730 \cdot x$

4) $H_0: \beta_1 = 0.15$

Sæt cursor på et vilkårligt tal under "Parameter estimates", højre musetast ► Columns ► vælg "lower" ► gentag men vælg "Upper"

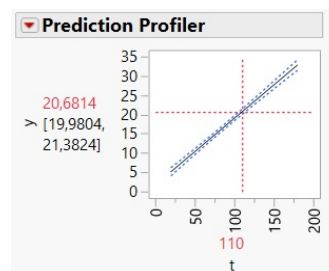
9. Regressionsanalyse

Parameter Estimates						
Term	Estimate	Std Error	t Ratio	Prob> t	Lower 95%	Upper 95%
Intercept	1,6541465	0,599582	2,76	0,0129*	0,3944709	2,9138221
t	0,172975	0,006032	28,68	<,0001*	0,160302	0,1856481

Heraf ses, at et 95% konfidensinterval for β ikke indeholder 0.15.

dvs. data giver ikke den i litteraturen angivne hældningskoefficient..

- 5) Find det til $t = 110$ svarende 95% konfidensinterval for tykkelsen y .
Analyze ► Fit model ► Marker y og vælg y ► marker t og vælg add ► Minimal report ► Run
Rød pil ved "Response y ► Factor Profiling ► Profiler ► på figur skriv 110



Vi får y 's værdi for $x = 110 = \underline{20.68}$.
95% konfidensinterval [19.98 ; 21.38]



9.7 MULTIPLE REGRESSIONS ANALYSE.

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 SAS.JMP

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 værdien og et 95% konfidensinterval for Y i punktet $(x_1, x_2, x_3, x_4) = (0, 20, 90, 100)$

Løsning :

Data indtastes

	x1	x2	x3	x4	y
1	-4	22	91	100	836
2	-1	20	90	95	789
3	7	21	88	110	883
4	16	19	87	88	790
5	18	20	91	94	816
6	23	19	94	99	859
7	27	23	87	97	831
8	29	21	86	96	832
9	24	22	88	110	897
10	16	23	91	105	872
11	10	20	90	100	842
12	3	20	89	98	821

- 1) Vælg Analyze ► Fit model ► markér "y" og tryk på Y Response ► markér "x1, x2, x3, x4" og tryk ADD ► Emphasis: Minimal report ► Run
Der fremkommer bl.a. følgende udskrift

Response y	
Summary of Fit	
RSquare	0,965448
RSquare Adj	0,945705
Root Mean Square Error	7,90936
Mean of Response	839
Observations (or Sum Wgts)	12

9. Regressionsanalyse

Vælg rød pil ► Save Columns ► Studentized residuals
I datatabel kan man nu yderligere finde følgende

x1	x2	x3	x4	y	Studentized Resid y
-4	22	91	100	836	1,7900396908
-1	20	90	95	789	-2,254767828
7	21	88	110	883	0,1362598723
16	19	87	88	790	1,4978025262
18	20	91	94	816	-0,633566393
23	19	94	99	859	0,5274844434
27	23	87	97	831	-0,653345676
29	21	86	96	832	0,1146976544
24	22	88	110	897	-0,330042097
16	23	91	105	872	0,2078113731
10	20	90	100	842	0,2523615824
3	20	89	98	821	-0,012553137

Da kun en enkelt værdi numerisk er større end 2 og ingen er over 3, antages, at der ikke er outliers
Da yderligere forklaringsgraden = 0.965 er tæt ved 1 vurderes modellen at være rimelig god.

2) Mulig reduktion af modellen

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$, H: Mindst en af regressionskoefficienterne er forskellig fra 0.

I samme udskrift som under "Summary of Fit" fandtes

Response y				
Summary of Fit				
Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	12236,094	3059,02	48,8990
Error	7	437,906	62,56	Prob > F
C. Total	11	12674,000		<,0001*
Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	175,49949	113,8626	1,54	0,1671
x1	1,0266425	0,226819	4,53	0,0027*
x2	-0,793015	2,120284	-0,37	0,7195
x3	1,9113126	1,145911	1,67	0,1393
x4	4,9822626	0,44901	11,10	<,0001*

Af ovenstående udskrift ses for model, at P -Value $< 0.0001 < 0.05$.

Heraf følger, at H_0 forkastes (stærkt), dvs. mindst en af regressionskoefficienterne er forskellig fra 0.

Vi ser nu regressionskoefficienterne

Den β størrelse, der har størst P -værdi er β_2 .

$H_0: \beta_2 = 0$ accepteres, da P -værdien = $0.7195 > 0.05$.

x2-leddet bortkastes.

Bemærk, at man kun eliminerer én variabel ad gangen.

Vi eliminerer nu x_2 :
(slettes under "ADD")

Da P -værdien for x_3 er $0.0849 > 0.05$ eliminerer vi nu x_3 .

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
x1	1	1	1274,177	22,8216	0,0014*
x3	1	1	215,715	3,8636	0,0849
x4	1	1	10796,609	193,3764	<,0001*

Det er nu ikke muligt at reducere modellen mere.

Effect Tests					
Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
x1	1	1	1083,972	14,7285	0,0040*
x4	1	1	11089,679	150,6814	<,0001*

3) For at kunne angive regressionsligningen betragtes følgende udskrift.

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	335,65234	40,26796	8,34	<,0001*
x1	0,9016966	0,234953	3,84	0,0040*
x4	4,9401735	0,40245	12,28	<,0001*

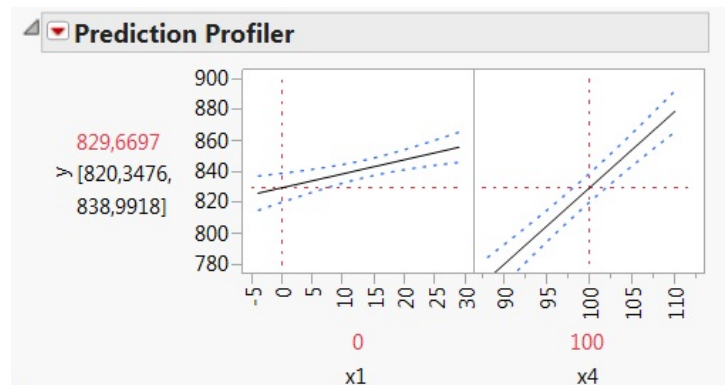
Ligningen bliver $\hat{y} = 335.65 + 0.9017 \cdot x_1 + 4.9402 \cdot x_4$

4) Cursor i tabellen ovenfor, højre musetast ► Columns ► Upper 95% ► lower 95%

Parameter Estimates						
Term	Estimate	Std Error	t Ratio	Prob> t	Lower 95%	Upper 95%
Intercept	335,65234	40,26796	8,34	<,0001*	244,55989	426,74479
x1	0,9016966	0,234953	3,84	0,0040*	0,370196	1,4331972
x4	4,9401735	0,40245	12,28	<,0001*	4,0297676	5,8505795

Konfidensintervallerne bliver $\beta_1 : [0.3702 ; 1.4332]$, $\beta_4 : [4.030 ; 5.851]$

5) Cursor på rød pil "Response Y"
► vælg "Factor Profiling" ► Profiler ►
cursor på det røde tal foruden ved x_1 , skriv 0, Cursor på rødt tal ved x_4 , skriv 100



Vi har derfor, at $\hat{y} = 829.67$ og et 95% konfidensinterval er $[820.35 ; 838.99]$ ◆

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 specialtilfælde. 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 andengradspolynomium $Y = \beta_0 + \beta_1 x + \beta_2 x^2$ 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 $T = a_0 + b_1 x + b_2 x^2 + b_3 x^3$

til $T = a_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$ 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 $Y = \beta_0 + \beta_1 x + \beta_2 x^2$, beregner forklaringsgraden r^2 , derefter betragter et polynomium af tredje grad, og beregner igen forklaringsgraden, så er den steget, selv om trediegradsmodellen 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.

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).

Fremgangsmåde

Man starter med at tegne punkterne, eventuelt sammen med en lineær model af første grad $y = \alpha + \beta_1 x$ (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 $y = \alpha + \beta_1 x + \beta_2 x^2$ og sammenligner den med en af tredje grad $y = \alpha + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$

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 trediegradspolynomiet 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 trediegradspolynomiet med et polynomium af fjerdegrad.

Backward

Her starter man med en model af fjerde grad $y = \alpha + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4$

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 r^2 " for at se om der er sket en væsentlig ændring. Således fortsættes indtil man får en forkastelse af nulhypotesen, samt ser, at "adjusted r^2 " 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 $T = \alpha_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \dots + \beta_p x^p$ 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.

9. Regressionsanalyse

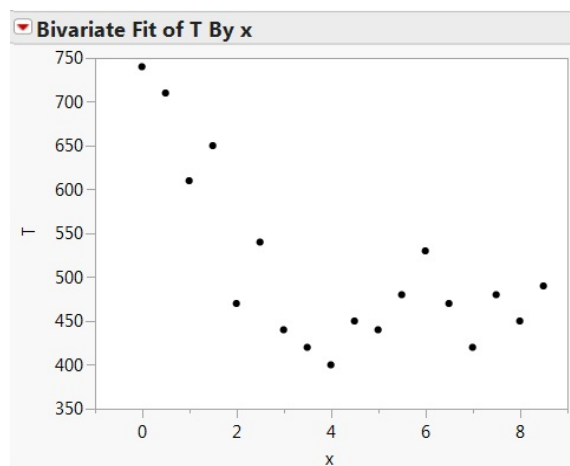
3) Beregn værdien af T for $x = 6.2$, og angiv et 95% konfidensinterval for T for $x = 6.2$.

Løsning:

1) Data indtastes

a) Analyze ► Fit Y by X ► markér "T" og tryk på Y Response ► markér "x" og tryk på X Factor ► OK

Der fremkommer et "scatterplot", hvor man i et koordinatsystem kan se punkterne afbildet.



Cursor på rød pil ved overskrift på tegning, højre musetast ► Fit Polynomial

Man kan nu vælge, hvilken grad polynomiet skal have.

Ud fra scatterplottet synes en andengradsmodel ikke at være en god model

Vi vælger dog først en andengradsmodel 2.quadratic
Der fremkommer blandt andet følgende udskrift:

Summary of Fit	
RSquare	0,787571
RSquare Adj	0,759247
Root Mean Square Error	49,711
Mean of Response	510,5556
Observations (or Sum Wgts)	18

Vi vælger nu en trediegradsmodel 3,cubic

Vi ser, at R-squared (adjusted) er steget fra 0.759 til 0.807.

Summary of Fit	
RSquare	0,841256
RSquare Adj	0,807239
Root Mean Square Error	44,48109
Mean of Response	510,5556
Observations (or Sum Wgts)	18

Vi vælger nu en fjerdegradsmodel 4,quartic

Vi ser, at RSquar Adj (adjusted) er faldet fra 0.807 til 0.799

Summary of Fit	
RSquare	0,846337
RSquare Adj	0,799056
Root Mean Square Error	45,41548
Mean of Response	510,5556
Observations (or Sum Wgts)	18

Heraf må slutes, at fjerdegradsmodellen ikke har givet et væsentligt forbedret bidrag til forklaring af data.

For at lave tests, konfidensintervaller m.m. må man indføre 2 nye kolonner $x_2 = x^2$ og $x_3 = x^3$ og gå over i multipel analyse.

Hertil benyttes formula (Cursor på kolonneoverskrift ► højre musetast ► formula) ► vælg potensformel i oversigt

Vælg Analyze ► Fit model ► markér "T" og tryk på Y Response ► markér "x, x2, x3" (vælg minimal report) og tryk ADD ► Run

Man får bl.a.

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	770,70175	34,52201	22,32	<,0001*
x	-179,2699	36,21048	-4,95	0,0002*
x2	31,336429	10,07039	3,11	0,0077*
x3	-1,692466	0,777816	-2,18	0,0472*

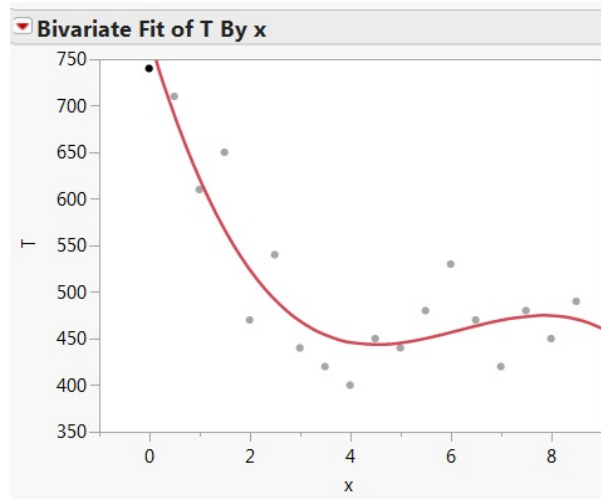
Da vi ser, at P-værdien for $\beta_3 = 0.047 < 0.05$ forkastes $H_0: \beta_3 = 0$, dvs.

Vi kan ikke bortkaste trediegradsleddet.

Heraf slutes, at en trediegradsmodel må være det foreløbig bedste bud

Grafen for trediegradsmodellen bliver

Da punkterne ligger tilfældigt omkring kurven finder vi, at trediegradsmodellen er en acceptabel model



For at vurdere om der er outliers, vælges studentized residuals

Vælg rød pil (ved response) ► Save Columns ► Studentized residuals

Resultat blev følgende tabel

9. Regressionsanalyse

Den viser, at ingen numerisk er over 3, og kun en enkelt på 2.05 er numerisk over 2

Konklusion: Trediegradsmodellen er den bedste model

x	T	x2	x3	Studentized Resid T
0	740	0	0	-1,094539224
0,5	710	0,25	0,125	0,5615462107
1	610	1	1	-0,273772112
1,5	650	2,25	3,375	2,0531565861
2	470	4	8	-1,337807591
2,5	540	6,25	15,625	1,19147889
3	440	9	27	-0,717878586
3,5	420	12,25	42,875	-0,838797886
4	400	16	64	-1,123753994
4,5	450	20,25	91,125	0,1366190753
5	440	25	125	-0,150579153
5,5	480	30,25	166,375	0,7109070448
6	530	36	216	1,7941894139
6,5	470	42,25	274,625	0,133409821
7	420	49	343	-1,250031893
7,5	480	56,25	421,875	0,1274797235
8	450	64	512	-0,672767712
8,5	490	72,25	614,125	0,6567235345

2) Regressionsligningen ses under parameter Estimates

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	770,70175	34,52201	22,32	<,0001*
x	-179,2699	36,21048	-4,95	0,0002*
x2	31,336429	10,07039	3,11	0,0077*
x3	-1,692466	0,777816	-2,18	0,0472*

$$\hat{T} = 770.70 - 179.27x + 31.336x^2 - 1.629x^3$$

3) Beregn værdien af T for x = 6.2, og angiv et 95% konfidensinterval for T for x = 6.2.

Skriv 6.2 nederst i datatabel ► cursor på overskrift på x2, højre musetast ►

Nu kommer automatisk i kolonnerne 6.2^2 osv.

Vælg Analyze ► Fit model ► OSV

Rød pil ved Response T ► Save Columns ► Predicted value

derefter det samme men nu vælges ► Mean Confidence Interval

Nu dannes i tabellen tre nye kolonner op, hvoraf man aflæser det ønskede.

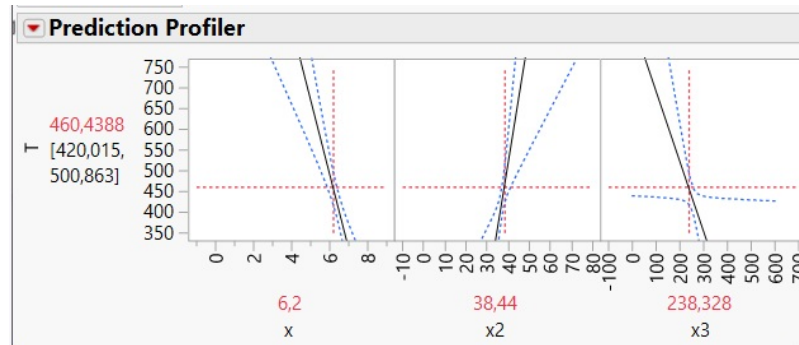
x	T	x2	x3	Studentized Resid T	Predicted T	Lower 95% Mean T	Upper 95% Mean T
0	740	0	0	-1,094539224	770,70175439	696,65941622	844,74409255
0,5	710	0,25	0,125	0,5615462107	688,68937049	638,92422323	738,45451774
1	610	1	1	-0,273772112	621,07585139	581,42093629	660,7307665
1,5	650	2,25	3,375	2,0531565861	566,59184727	527,7344981	605,44919643
2	470	4	8	-1,337807591	523,96800826	483,77394584	564,16207067
2,5	540	6,25	15,625	1,19147889	491,93498452	451,74092211	532,12904693
3	440	9	27	-0,717878586	469,22342621	430,77106308	507,67578935
3,5	420	12,25	42,875	-0,838797886	454,56398349	418,63727882	490,49068816
4	400	16	64	-1,123753994	446,6873065	412,60539012	480,76922288
4,5	450	20,25	91,125	0,1366190753	444,32404541	410,24212903	478,40596175
5	440	25	125	-0,150579153	446,20485036	410,27814569	482,13155503
5,5	480	30,25	166,375	0,7109070448	451,06037152	412,60800838	489,51273465
6	530	36	216	1,7941894139	457,62125903	417,42719662	497,81532144
6,5	470	42,25	274,625	0,133409821	464,61816305	424,42410064	504,81222547
7	420	49	343	-1,250031893	470,78173375	431,92438458	509,63908292
7,5	480	56,25	421,875	0,1274797235	474,84262126	435,18770615	514,49753636
8	450	64	512	-0,672767712	475,53147575	425,76632849	525,296623
8,5	490	72,25	614,125	0,6567235345	471,57894737	397,5366092	545,62128554
6,2	460,4388	38,44	238,328		460,43878225	420,01505899	500,86250551

Man får $T = 460.44$ og 95% konfidensinterval $[420.0; 500.9]$

En anden mulighed er

Rød pil ved Resonse T ► Factor profiling ► Profiler

Der fremkommer to figurer. Tallene ændres i bunden



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"
 - 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. Fit Model, quadratic, beregne R-squarea (adjusted)
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 $\Leftrightarrow H_0: \beta_1 = 0$.
 - b) Konfidensinterval for β_1 :
 - c) Konfidensinterval for y svarende til x = t:
 - d) Prædiktionsinterval for 1 ny observation
- 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)

Test

- 1) Reduktion af model:

Hvis højstegradsleddet har en største P-værdi $> 5\%$

Fjerner leddet, Gentag. Der elimineres én variabel ad gangen
- 2) Koefficienter : Findes i blist (se evt. under multipel regression)

9.9.2 Multipel regression Se eksempel 9.8

Model acceptabel

Se på forklaringsgrad (bør være over 0.7)

Undersøge om der er Outliers (ingen numerisk over 3)

Test

1) Reduktion af model:

 Find det led hvis koefficient har den største P-værdi >5% Fjerner leddet

 Gentag med næste P-værdi Der elimineres én variabel ad gangen

2) Regressionskoefficienter : Findes i blist

3) 95% konfidensinterval for β_1 :

4) Man finder 95% konfidensintervallet for y for given x-værdi

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	

- 1) Tegn en figur, hvor punkterne og en relevant ret linie er indlagt
- 2) Find r^2 og anvend denne, ovenstående figur samt om der er outliers, 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 estimat og et 95% konfidensinterval for hældningen β_1 .
- 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) Giv en vurdering af om en ret linie $m: y = a + bx$ på rimelig måde kan beskrive denne sammenhæng. Undersøg specielt om der er outliers.
- 2) Idet det antages at ovenstående linie er en god model skal man angive en ligning for regressionslinien m .
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 (lineær, potens, eksponential, logaritmisk), 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 trafikekspert 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) Kontroller om forudsætningerne om varianshomogenitet er rimelig opfyldt. Det antages i det følgende at forudsætningerne er rimelig 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 middelveiskositeten 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 regressionslinje 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 afhænger 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	6.37				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 $x_1 = 3$, $x_2 = 4$ og $x_3 = 2$.

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

- 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
- Opskriv regressionsligningen for den i spørgsmål 1 fundne model
- Find den værdi x_m som giver den største y - værdi . Angiv endvidere den til x_m svarende estimerede middelværdi \hat{Y}_m , og et 95% konfidensinterval.

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

- Bestem ved en polynomial regressionsanalyse det polynomium af lavest grad, der giver en tilfredsstillende beskrivelse af Y 's variation.
- Benyt den fundne model til at bestemme 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 eksempelvis 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.

Nullhypotese: $H_0: m = 70$ 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 udskød 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 $= P(w_- \leq 14) = \text{Normal Distribution}\left(\frac{14 - 52,5}{15,93}\right) = 0,0078$,

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 (Wilcoxons teststørrelse).

Tabel over kritiske værdier af Wilcoxons teststørrelse w_α

α 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 . Wilcoxons 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.

Approximation 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 = $P(w_A \leq 25) = \text{Normal Distribution}\left(\frac{25 - 42}{7}\right) = 0.00758$.

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_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset
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_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset	w_n	w_\emptyset
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 $w_n < w_n$ eller $w_n > w_\emptyset$ (ligger udenfor intervallet $[w_n ; w_\emptyset]$)

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 SASJMP til at gøre det på følgende måde:

► Indtast tallene i den første liste, og numrene på medicinen på den anden liste (A = nummer 1 osv)

Cursor i listeoverskrift "tal" ► højre musetast, "Sort" ► Ascending ► ok

Man ser nu at tal listen er sorteret i stigende rækkefølge, og man kan nu udfylde skemaet.

tal	nr
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

tal	nr
1,62	1
1,71	1
1,93	1
1,96	1
2,07	2
2,11	2
2,24	1
2,41	1
2,43	2
2,5	2
2,71	2
2,84	2
2,88	2

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 ordning, og at den bagved liggende fordeling er kontinuert.

Eksempel 10.4 . 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 inddeles 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:

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 SAS.JMP til at gøre det (vist på forrige side):

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 $\chi^2 = \frac{12 \cdot \sum_{i=1}^k n_i (\bar{R}_i - \bar{R})^2}{n(n+1)}$ er χ^2 -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 $= P(\chi^2 > 6.02) = 1 - \text{ChiSquareDistribution}(6.02, 2) = 0.0493 < 0.05$ 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 tvillingpar 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 parametrisk test er opfyldt.
- 2) ved et parametrisk 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-parametrisk 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

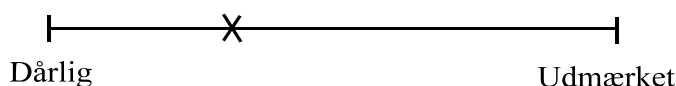
Test, om middeludbyttet er forskelligt ved de 2 metoder.

Da man ikke er sikker på, at forudsætningerne for en parametrisk 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 kontrollkort

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 4.

Tabel 2.1: Kontrollkort.

$\bar{x} - R$ - kontrolkort. Procesvariablen X er normalfordelt $n(\mu, \sigma)$.					
Procestilstand		Centerlinie	Nedre kontrolgrænse	Øvre kontrolgrænse	Estimerer
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

Et R - kort med 3σ - grænser er derfor bestemt ved :

$$NKG_R = (d_2 - 3 \cdot d_3) \cdot \sigma \text{ og } \text{ØKG}_R = (d_2 + 3 \cdot d_3) \cdot \sigma.$$

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 $\text{ØKG}_R = (d_2 + 3 \cdot d_3) \cdot \frac{\bar{R}}{d_2} = D_4 \bar{R}$.

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 :

$$NKG_s = (c_4 - 3 \cdot \sqrt{1 - c_4^2}) \cdot \sigma = B_5 \cdot \sigma \text{ og } \text{ØKG}_s = (c_4 + 3 \cdot \sqrt{1 - c_4^2}) \cdot \sigma = B_6 \cdot \sigma.$$

Idet et estimat for $E(s) \approx \bar{s}$, fås $\sigma \approx \frac{\bar{s}}{c_4}$.

Indsættes dette, fås $NKG_s = \left(1 - \frac{3}{c_4} \sqrt{1 - c_4^2}\right) \cdot \bar{s} = B_5 \cdot \bar{s}$ og $\text{ØKG}_s = \left(1 + \frac{3}{c_4} \sqrt{1 - c_4^2}\right) \cdot \bar{s} = B_6 \cdot \bar{s}$.

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:

Appendix

$$NKG_x = \mu - 3 \cdot \frac{\sigma}{\sqrt{n}} = \mu - A_1 \cdot \sigma \text{ og } \text{ØKG}_x = \mu + 3 \cdot \frac{\sigma}{\sqrt{n}} = \mu + A_1 \cdot \sigma$$

Benyttes et R - kort er $\sigma \approx \frac{\bar{R}}{d_3}$, og dermed

$$NKG_x = \bar{x} - 3 \cdot \frac{\bar{R}}{d_2 \cdot \sqrt{n}} = \mu - A_2 \cdot \bar{R} \text{ og } \text{ØKG}_x = \bar{x} + 3 \cdot \frac{\bar{R}}{d_2 \cdot \sqrt{n}} = \mu + A_2 \cdot \bar{R}$$

Benyttes et s kort er $\sigma \approx \frac{\bar{s}}{c_4}$, og dermed

$$NKG_x = \bar{x} - 3 \cdot \frac{\bar{s}}{c_4 \cdot \sqrt{n}} = \mu - A_3 \cdot \bar{s} \text{ og } \text{ØKG}_x = \bar{x} + 3 \cdot \frac{\bar{s}}{c_4 \cdot \sqrt{n}} = \mu + A_3 \cdot \bar{s}.$$

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 , ..., x _{1n₁}	\bar{x}_1	S ₁
R ₂	x ₂₁ , x ₂₂ , x ₂₃ , ..., x _{2n₂}	\bar{x}_2	S ₂
R ₃	x ₃₁ , x ₃₂ , x ₃₃ , ..., x _{3n₃}	\bar{x}_3	S ₃
R _r	x _{r1} , x _{r2} , x _{r3} , ..., 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 σ².

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_e² er variansen for "forsøgsfejlen" eller på engelsk "error". s_e² 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	$s_R^2 = \frac{SAK_R}{r-1}$	$F_R = \frac{s_R^2}{s_e^2}$	
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 : $\left[x_i - \frac{r_{kon}}{\sqrt{2}}; x_i + \frac{r_{kon}}{\sqrt{2}} \right]$

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 Bartletts test eller Levines test.

b) Bartletts test. Denne test er beregningsmæssigt vanskelig, og har den svaghed, at den er særdeles følsom overfor afvigelse 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 $\chi^2 = \frac{\text{[ovenstående ligning]}}{1 - \frac{1}{N-k} - \sum_{i=1}^k \frac{1}{n_i - 1}}$

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 $d_{ij} = |y_j - \hat{y}_i|$ 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:

		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:

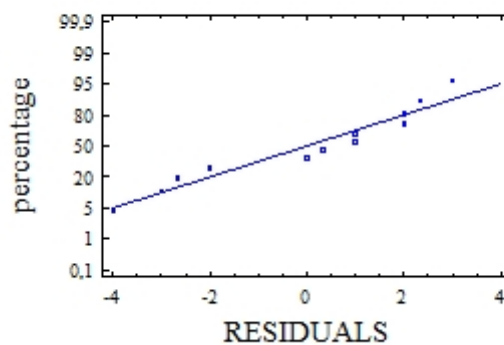
Lad residualerne (fra eksempel 6.1) -2, 0, 2, -3, 2, 1, 2.7, -2.3, -0.3, 1, 3, -4

De ordnes i rækkefølge og man beregner deres komulative 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:

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

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

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

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

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

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

(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 række faktor)

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

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

$$\tilde{\mu}_{ij} = x_{i.} + x_{.j} - x_{..} \text{ (jævnfør betragtningerne i afsnit 2.3.2.2 side 59.)}$$

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:

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

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

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

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

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

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

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

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

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

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

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. Transformerer 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

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.

Appendix

Resultatet kan umiddelbart aflæses: Nedenfor er de relevante resultater angivet i den sædvanlige variansanalysetabel

Variation	SAK=SS	df	MS= s^2	F	P-værdi
Factor: klædefabrikat	0.057665	3	0.019222		
Block: Metoder	0.080445	3	0.026815		
Error	0.008881	9	0.000987		

Her svarer "Error" en sum af "støj"+"vekselvirkning".

Da vi fra 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 $x \Leftrightarrow H_0: Model = 0 \Leftrightarrow H_0: \beta_1 = 0$

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 $t = \frac{\hat{\beta}_1}{s_{\hat{\beta}_1}}$, hvor $s_{\hat{\beta}_1} = \frac{|\hat{\beta}_1| \cdot s_{\text{residual}}}{s_{\text{model}}}$ 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 $t = \frac{\hat{\beta}_1 - a}{s_{\hat{\beta}_1}}$, hvor $s_{\hat{\beta}_1} = \frac{|\hat{\beta}_1| \cdot s_{\text{residual}}}{s_{\text{model}}}$

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 $s_{\hat{\beta}_1} = \frac{|\hat{\beta}_1| \cdot s_{\text{residual}}}{s_{\text{model}}}$.

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, μ_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 6.1.

b) Varianshomogenitet (se appendix 6.1)

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_{21}	x_{22}	\cdot	\cdot	\cdot	x_{2k}	y_2
\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 $\bar{y} = X \cdot \bar{\beta}$

$$\text{hvor } \bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdot & \cdot & \cdot & x_{1k} \\ 1 & x_{21} & x_{22} & \cdot & \cdot & \cdot & x_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{N1} & x_{N2} & \cdot & \cdot & \cdot & x_{Nk} \end{bmatrix}, \bar{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_k \end{bmatrix}.$$

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

$$X^T \cdot X \cdot \hat{\beta} = X^T \cdot \bar{y} \quad (\text{kaldet } \mathbf{normalligningssystemet}). \quad (3)$$

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. $r_i = y_i - \hat{y}_i$.

$$SAK_{\text{residual}} = \sum_{i=1}^n r_i^2.$$

Appendix

Sættes $\bar{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ r_n \end{bmatrix}$ kan vi foretage følgende omskrivning

$$SAK_{\text{residual}} = \sum_{i=1}^n r_i^2 = \bar{r}^T \cdot \bar{r} = (\bar{y} - X\hat{\beta})^T \cdot (\bar{y} - X\hat{\beta}).$$

$SAK_{\text{total}} = (n-1) \cdot s_y^2$ 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 nulhypoteserne

$$H_0: \beta_i = 0 \text{ mod den alternative hypotese } H: \beta_i \neq 0$$

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

$$H_0 \text{ forkastes, hvis } P\text{-værdi} = P(T > |t_i|) < \frac{\alpha}{2}.$$

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 varianterne 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.

$$100(1 - \alpha)\% \text{ konfidensinterval for } \beta_i: \hat{\beta}_i - t_{1-\frac{\alpha}{2}}(N-p) \cdot s(\hat{\beta}_i) \leq \beta_i \leq \hat{\beta}_i + t_{1-\frac{\alpha}{2}}(N-p) \cdot s(\hat{\beta}_i).$$

Konfidensinterval for et til punktet \bar{x}_0 svarende værdi \hat{y}_0 :

$$\text{Lad } \bar{x}_0 = \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \\ \vdots \\ \vdots \end{bmatrix}$$

$$\left[\hat{y}_0 - t_{1-\frac{\alpha}{2}}(N-p) \cdot \sqrt{s_{\text{residual}}^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0}; \hat{y}_0 + t_{1-\frac{\alpha}{2}}(N-p) \cdot \sqrt{s_{\text{residual}}^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0} \right]$$

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

$$\left[\hat{y}_0 - t_{1-\frac{\alpha}{2}}(n-p) \cdot \sqrt{s_{\text{residual}}^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0}; \hat{y}_0 + t_{1-\frac{\alpha}{2}}(n-p) \cdot \sqrt{s_{\text{residual}}^2 \cdot \bar{x}_0^T \cdot (X^T \cdot X)^{-1} \cdot \bar{x}_0} \right]$$

APPENDIX 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 $\chi^2 > \chi^2_{1-\alpha}(k-1)$

eller P -værdi = $P(Y > \chi^2) < \alpha$, hvor Y er χ^2 -fordelt med $k - 1$ frihedsgrader.

FACITLISTE

KAPITEL 1

- 1.1 (1) 0.6915 (2) 0.1088 (3) 112,20 (4) 117.32 6.535
 1.2 69.94 15.91
 1.3 a) P -værdi = 0.48% b) 1.316 [1.186; 1.445] c) 18.42%
 1.4 P - værdi = 0.043
 1.5 P - værdi = 0,063179

KAPITEL 2

- 2.1 (1) $NKG_R = 0$, $\emptyset KG_R = 10.14$, $NKG_{\bar{x}} = 451.4$, $\emptyset KG_{\bar{x}} = 457.8$ (2) 1.03 %.
 2.2 $NKG_s = 0$, $\emptyset KG_s = 1.410$, $NKG_{\bar{x}} = 11.30$, $\emptyset KG_{\bar{x}} = 13.44$
 2.3 (1) - (2) - (3) $\bar{x} = 2929.1$ $\tilde{\sigma} = 131.7$
 2.4 (1) $NKG_R = 0$, $\emptyset KG_R = 1.01$, $NKG_{\bar{x}} = 85.26$, $\emptyset KG_{\bar{x}} = 86.42$
 (2) alarm 5 , 6 , 3
 2.5 (1) 1.33 (2) 0,876% (3) 18,5 0.0066%
 2.6 a) $NKG = 11.32$, $\emptyset KG = 13.51$ b) 1.068
 2.7 $NKG = 0.82$, $\emptyset KG = 19.85$
 2.8 $\emptyset KG = 16.09$ $NKG = 0$
 2.9 a) $NKG = 0$ $\emptyset KG = 15.16$ 7.14 b) 1.2 0.0167%
 2.10 a) - b) alarm 3 , 2, 1
 2.11 a) 1.179 b) -

KAPITEL 3

- 3.1 (1) - (2) 82.6% (3) 81.43%
 3.2 (1) $NKG = 0$, $\emptyset KG = 3.47$ (2) $c = 4$, $n = 300$, (3) 5.24 (4a) 0.70% (4b) 60.67%
 3.3 (1) $n = 230$, $c = 11$, (2) ca 78,7% (3) 78.7%
 3.4 (1) $n = 180$, $c = 8$, (2) - (3) - (4) ca 650.3
 3.5 (1) (82,2) (2) (60, 120,1,3,3) (3) - (4) ca 32,3% (5) 74.17 117.49 102.66 74.83
 3.6 (1) (78,78,3,6,6) (2) 51.3% (3) 86.31

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 = 0.0472 , [0.015 ; 1.87]
 5.5 a) P - værdi = 0.0001, b) 0.2258 C) [11.93 ; 14.96] c) 0.2258
 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 = 0,024 (2) P - værdi = 0.00334
 5.12 P - værdi = 0.233

Facitliste

- 5.13 (1) P - værdi = 0.2563 (2) P - værdi = 0.0177
5.14 (1) P - værdi = 0.0134 (2) P - værdi = 0.0001
5.15 P - værdi = 0.00028

KAPITEL 6

- 6.1 a) - b) P - værdi = 0.0004 c) B2, 3181.5 [3055.4 ; 3307.1]
6.2 P - værdi = 0.0309
6.3 P - værdi = 0.0064
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 P - værdi = 0.0696
6.8 P - værdi = 0.027

KAPITEL 7

- 7.1 (1) nej, P - værdi = 0.0108, (2a) K1 (2b) frit valg (2c) $A_3 K_2$, $A_2 K_1$, $A_1 K_2$, $A_3 K_1$
7.2 (1) ja (2) glas 2 og enten fosfor 1 eller 3.
7.3 (1) - 30.84% (2) limtyper (3) 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 300⁰, tryk 60
7.6 (1) Nej, P - værdi = 0.046 (2) ja, P - værdi = 0.078
7.7 Begge, Vitaminbehandling 1 eller 3.
7.8 (1) Nej, P - værdi = 0.0092 (2) råmateriale 4. (3) kun uger
7.9 Begge har signifikant virkning, vælg vævemets metode 2, og ikke matrialtypen 3.

KAPITEL 8

- 8.1 1) A og B 2) $A_3 B_1$ 3) 45.75 [44.85 ; 48.65]
8.2 a) - b) - c) - d) 3
8.3 a) - b) - c) - d) 3
8.4 a) - b) -
8.5 (a) Fabrikationsmetode og farvestof vekselvirker (b) Fabrikationsmetode 2, Farvestof ikke tilsat
8.6 (a) B, C, D, E har virkning (b) B og E på højt, C og D på lavt, -3.875, [-12.39 ; 4.64]
8.7 (a) A og C vekselvirker (b) A, C på højt, 70.85, [59.48 ; 82.22]
8.8 (a) A, E og F har virkning (b) A højt, E lavt, F højt

KAPITEL 9

- 9.1 (1) - (2) 0.7205 (3) $y = 607.1 - 21.4025x$ (4) -21.4 [-26.66 ; -16.14] (5) [249.1 ; 258.8]
9.2 (1) 70.5% (2) $y = 40.78 + 0.766x$ (3) P - værdi = 0.0024 (4) [0.362 ; 1.169]
(5) [72.51 ; 85.61] (6) [57.95 ; 100.17]
9.3 (1) $r^2 = 0.99$ potens (2) $y = 40.673x^{0.53597}$ (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) 13.08 (e) [12.66 ; 13.52]
9.5 (1) - (2) $r^2 = 0.9805$ (3) $\tilde{\alpha}_0 = 1718$ $\tilde{\beta}_1 = 0.8534$ (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) $P(\text{lack of fit}) = 0.5453$, $y = 6.683 + 0.3327 x$
 (4) P - værdi = $2.9 \cdot 10^{-7}$ (5) [25.77 ; 40.76] (6) [56.9 ; 62.69]
- 9.8 (1) $Y = -0.1438 + 108.959 x$ (2) P - værdi = 0.144 (3) ja
- 9.9 (1) $r^2 = 0.723$ (2) - (3) $Y = 1131.8 + 9.5846 \cdot x_1$ (3) 2329.96 [2267 ; 2393]
- 9.10 (1) $r^2 = 0.9823$ ja (2) ja (3) $Y = 21625 + 34615x_1 + 62881x_2$ (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], [0.107; 0.231] (4) 8.526, [8.406 ; 8.646]
- 9.12 (1) - (2) - (3) $y = 5.9167 + 2.251x_1 + 1.648x_2$ (4) [1.92; 2.57] (5) 56.88, [55.33; 58.44]
- 9.13 (1) - (2) $Y = -1144 + 0.7197x - 0.0379x^2$, $Y = 2.037$, [1.96 ; 2.114]
- 9.14 (1) - (2) $Y = 68.720 + 0.5664x - 0.007625x^2$, (3) 37.14, 79.237 [77.147 ; 81.327]
- 9.15 (1) $Y = 613.56 - 355.11x + 60.4679x^2 - 3.3623x^3$, (2) $y = 53,48$, [47.81; 59.15]

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 29
 Aceptable Quality Level AQL 30
 additiv model 88, 90
 adjusted R^2 150
 aftagerens risiko 30
 aliasrelationer 109
 Appendix 174
 Average Outgoing Quality, AOQ 32
 Average Outgoing Quality Limit, AOQL 32
 Average Sample Number, ASN 34
 Average Total Inspection, ATI 32

B

bagatelgrænse 44
 Bartlett's test 65, 178
 behandlinger 41
 binomialfordeling
 kontrolkort 21
 test
 een variabel 3
 en faktor på to niveauer 51
 en faktor på mere end to niveauer 74
 faktorforsøg 92
 blokforsøg, randomiseret 50, 71, 92, 114
 boxplot 15

C

c - kontrolkort 22
 Control Chart Sigma 19

D

definitionsrelation 109
 delforsøg 41
 dimensionering 44, 48, 55
 dobbelt stikprøveplan 32
 tabel 37

E

ekstrapolation 135
 én faktor ad gangen 81
 enkelt regressionsanalyse 128
 med gentagelser 143
 formler 187
 uden gentagelser 146

formler 185
 eksponentiel model 140
 enkelt stikprøveplan 29
 tabel 36
 ensidet variansanalyse 63
 beregninger-formler 64, 177

F

facitliste 193
 faktorer 41
 1 faktor på 2 niveauer
 normalfordelte 45
 binomialfordelte 51
 Poissonfordelte 53
 1 faktor på mere end 2 niveauer 62
 faktorforsøg
 2^k , fuldstændigt 106
 2^k , partielt 106, 110, 111
 fejl af type I 44, 48
 fejl af type II 44, 48
 F - test 48, 66
 forklaringsgrad 132, 135
 adjusted 150
 formler til beregning af
 test af differens mellem
 μ_1 og μ_2 i normalfordelte prøver 45
 p_1 og p_2 i binomialfordelte prøver 51
 μ_1 og μ_2 i Poissonfordelte prøver 53
 ensidet variansanalyse 177
 tosidet variansanalyse 180
 enkelt regressionsanalyse 185, 187
 multipel regressionsanalyse 189
 fortegnsmatrix 110
 forudsætninger for
 variensanalyse 64
 regressionsanalyse 137
 flere end 2 faktorer 102
 fuldstændig
 faktorstruktur 82
 2^k faktorstruktur 106
 fuldstændig randomiseret blokforsøg 71, 93

G

gentagelser 41

godkendelseskontrol 28
 godkendelsestal c 29
 grænser for kontrolkort 9

H

hovedvirkning 88, 90
 hypotesetest
 middelværdi :1 normalfordelt stikprøve 2
 middelværdi:2 normalford. stikprøver 45
 varians: 2 normalfordelte variable 48

I,J

ikke-parametrisk test 53, 165
 ISO-9000 7

K

kapabilitet 16
 kapabilitetsindeks 18
 konfidensinterval
 1 normalfordelt variabel 3
 differens, 2 normalfordelte variable 54
 differens, 2 binomialfordelte variable 57
 differens, 2 Poissonfordelte variable 57
 for C_p 20
 i ensidet variansanalyse 66, 175
 i tosidet variansanalyse 84, 88, 181
 LSD 67
 konfunderet partielt 2^k - faktorforsøg 113
 kontrolkortanalyse 11
 kontrol
 af binomialfordelt variabel 21
 af fejlprocent 21
 af normalfordelt variabel 10, 68
 af Poissonfordelt variabel 22
 løbende 16
 rektificerende 32
 varianshomogenitet 65, 69
 kontrolkort 9, 15
 kontrolgrænser 9
 kvalitetsstyring 7
 kvalitet
 tilfredsstillende 30
 utilfredsstillende 30
 korrelationskoefficient 135
 Kruskal- Wallis test 170
 kvalitativ faktor 41
 kvantitativ faktor 41

L

leverandørens risiko 30
 Levines test 65, 178
 Limiting Quality LQ 30
 lineær model 129
 logaritme model 140
 LSD (Least Signifikant Difference) 6
 løbende kontrol 16

M

mindste kvadraters metode 130
 MR=Moving Range 15
 multipel regression 146, 189

N

nedre kontrolkortgrænse 9
 niveau for faktor 41
 normalfordeling
 kontrolkort 10, 15
 test, 1 variabel 2
 2 variable
 af middelværdi 45
 af varians 48
 np - kontrolkort 21

O

OC-Kurve 29
 opgaver
 kapitel 1 5
 kapitel 2 24
 kapitel 3 38
 kapitel 5 58
 kapitel 6 78
 kapitel 7 98
 kapitel 8 125
 kapitel 9 158
 kapitel 10 172
 outliers 132
 oversigter over
 centrale formler i kapitel 5 54
 centrale formler i kapitel 6 76
 fremgangsmåde: tosidet variansanalyse 97
 fremgangsmåde: partielt 2^k faktorforsøg 124
 fremgangsmåde: regressionsanalyse 156
 Test af parametre for binomialfordeling 57
 Test af parametre for Poissonfordeling 57

P

partiel 2^k- faktorforsøg 112
parvise observationer 50
planlægning af forsøg 40, 42, 81, 112
Poissonfordeling
 kontrolkort 22
 test
 1 variabel 4
 2 variable 53
 mere end 2 variable 75
polynomial regressionsanalyse 150
poolet estimat for varians 65
potens model 140
proces i statistisk kontrol 8
proceskontrol 7
proces ude af statistisk kontrol 8
procesvariablen 10, 21
prædestinationsinterval 138

Q

Quality Control 7
Quality management 7

R

randomisering 42
randomiseret forsøg 65
randomiseret blokforsøg 71, 92
rangtal 166
rangtest
 1 variabel 166
 2 variable 53, 168
 mere end 2 variable 170, 192
regressionsanalyse
 enkelt
 med gentagelser 143
 formler 187
 uden gentagelser 131
 formler 185
 transformation 140
 forudsætninger 137
 multipel 146
 polynomial 150
regressionskoefficienter 129
regressionsligning 129
regressionslinie 129
rektificerende kontrol 32

repræsentativ 41
risikopunkter 30
R - kontrolkort 10
residual 132
 studentized 138
romersk kvadratforsøg 95
rækkefaktor 83

S

SAK = SS 84
Satterthwaites test 45
screeningsforsøg 104, 106
seks sigma proces 18
sekventiel forsøgsstrategi 123
Shewart-kontrolkort 9
s - kontrolkort 10, 14
specifikationsgrænser 17
SPC 7
SS = Sum of squares 84
statistisk gyldigt forsøg 42
statistisk uafhængige 42
stikprøveplan
 enkelt 29
 dobbelt 32
 ækvivalente 34
statistisk godkendelseskontrol 28
statistisk proceskontrol 7
stikord 196
stikprøvestørrelse n 9, 29
studentized residualer 136
søjlefaktor 83

T

Tabeller
 Bestemmelse af enkelt stikprøveplan 36
 Bestemmelse af dobbelt stikprøveplan 37
 Wilcoxons rangtest for 1 variabel 166
 Wilcoxons rangtest for 2 variable 168
tilfredsstillende kvalitet 30
tolerancegrænser 17
tosidet variansanalyse 81, 85
transformation 140, 188
Trinvis variansanalyse 93

U

uafhængige statistiske variable 41
ubalancerede forsøg 93

utilfredsstillende kvalitet 30

V

variansanalyse

ensidet 63

beregninger, formler 64

tosidet 81

beregninger, formler 83

med vekselvirkning 85

additiv 88, 90

varianshomogenitet 65 ,144

vekselvirkning 83

W

Wilcoxons rangtest for 1 variabel 166

Wilcoxons rangtest for 2 variable 168

X

x - streg kontrolkort 10

Z

Z-fordeling 1

Æ

ækvivalente stikprøveplaner 34

Ø

øvre kontrolgrænse 9