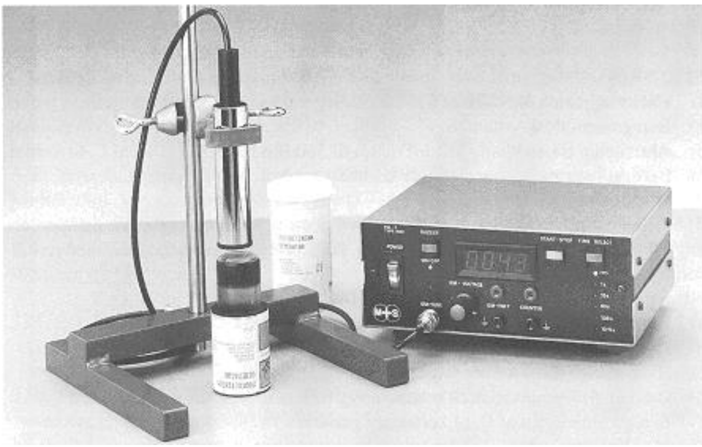
ISBN 9788770668699

Projekt 4.4 Linearisering af data fra radioaktivt henfald

Bemærk, at i det følgende er værktøjet TINspire anvendt. Det kan lige så godt laves i et andet værktøj.
En vigtig metode til at få overblik over data er at transformere dem, således at der fremkommer en lineær sammenhæng. Ordet transformation betyder omformning.
Vi vil tage udgangspunkt i forsøgsdata hvor radioaktivt henfald undersøges idet det vides at radioaktivt henfald er et eksempel på eksponentiel udvikling:

$$y = b \cdot a^x$$

Men først lidt teori fra fysikken:



Billedet viser en Protactinium-generator. Den består af en beholder, der indeholder to væsker: dels en vandig opløsning af uranylnitrat i koncentreret saltsyre, dels en organisk væske, som har mindre massefylde end vand. De to væsker er ikke blandbare, så når beholderen har stået et stykke tid, vil den organiske væske samle sig øverst i beholderen. I modsætning til Uran og Thorium er Protactinium opløselig i den organiske væske.
Beholderen med de radioaktive stoffer rystes og vendes om, lige inden forsøget startes. Derved isoleres Pa-234 i det øverste lag i beholderen. Den øvrige α- og β-stråling er ikke kraftig nok til at trænge igennem væsken og beholderen og op til Geiger-Müller røret.

Vi vil studere β-henfaldet med en geiger-müller tæller, der tæller i ti sekunders intervaller, og dernæst bestemme sammenhængen.

ISBN 9788770668699

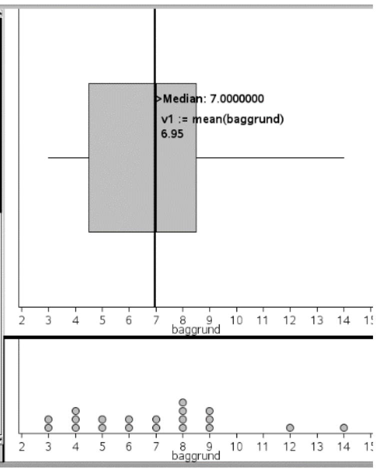
L&R Uddannelse EGMONT

Fremstil jeres egne data, hvor den uafhængige variabel er *tid* vores afhængige variabel er henfald. Eller hent elevdata fra et tilsvarende forsøg:

	A tid	B henfald
◆		
1	10	328
2	20	430
3	30	453
4	40	484
5	50	458
6	60	454
7	70	413
8	80	414
9	90	400
10	100	305
11	110	268
12	120	272
13	130	248
14	140	226
15	150	206
16	160	175
17	170	156
18	180	159
19	190	133
20	200	137
21	210	124
22	220	110
23	230	107
24	240	90
25	250	87
26	260	91
27	270	62
28	280	68
29	290	57
30	300	55
31	310	42
32	320	50
33	330	39
34	340	43
35	350	40
36	360	35
37	370	35
38	380	33
39	390	39
40	400	24
41	410	29
42	420	24
43	430	18
44	440	15
45	450	13
46	460	22
47	470	15
48	480	18
49	490	19
50	500	18
51	510	15
52	520	14
53	530	16
54	540	19
55	550	14
56		
57		
58		
59		
60		
61		
62		
63		

Dataene kan hentes i et regnearkformat **her** Inden vi kommer alt for godt i gang med vores undersøgelse skal vi huske at korrigere for baggrundsstrålingen. Her får vi igen hjælp af gruppen idet de 20 gange har målt baggrundsstrålingen over 10 sekunder:

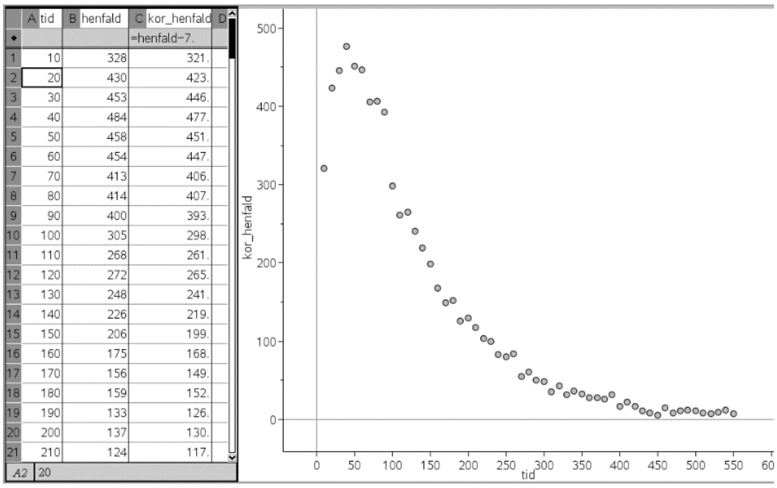
	A baggrund	B	C	D	E	F
◆			=OneVar('baggru			
1	5	Titel	Statistik med é...			
2	4	$\bar{x}$	6.95			
3	7	Σx	139.			
4	6	Σx²	1125.			
5	9	sx := s_{n-1}x	2.89237			
6	12	σx := σ_nx	2.81913			
7	5	n	20.			
8	14	MinX	3.			
9	7	Q_1X	4.5			
10	3	MedianX	7.			
11	4	Q_3X	8.5			
12	8	MaxX	14.			
13	3	SSX := $\Sigma(x-\bar{x})^2$	158.95			
14	9					
15	8					
16	8					
17	6					
18	8					
19	9					
20	4					
21						
C15						



Medianen er med 7.0 næsten sammenfaldende med middeltallet på 6.95. Vi bruger medianen til at korrigere for baggrundsstrålingen.

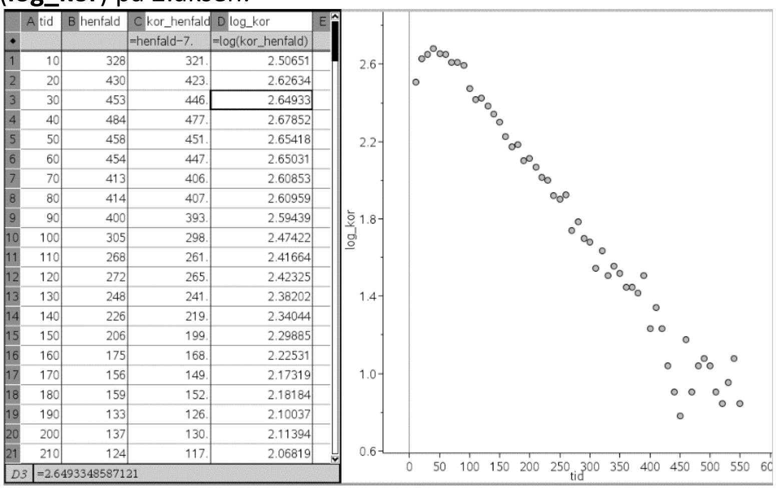
Det giver os en ny variabel over korrigerede henfaldsdata (**kor_henfald**) som vi kan bruge til at lave en graf over:

ISBN 9788770668699



	A tid	B henfald	C kor_henfald
◆			=henfald-7.
1	10	328	321.
2	20	430	423.
3	30	453	446.
4	40	484	477.
5	50	458	451.
6	60	454	447.
7	70	413	406.
8	80	414	407.
9	90	400	393.
10	100	305	298.
11	110	268	261.
12	120	272	265.
13	130	248	241.
14	140	226	219.
15	150	206	199.
16	160	175	168.
17	170	156	149.
18	180	159	152.
19	190	133	126.
20	200	137	130.
21	210	124	117.
A2	20		

Idet vi forventede at sammenhængen mellem **tid** og **korrigeret henfald** var eksponentielt virker grafen noget skuffende. Vi ser indledningsvis en stigning i β-henfaldet (frem til 40 sekunder), hvilket strider mod vores forestilling om eksponentielt henfald. Ellers ser forløbet af grafen rimelig ud.
Her vil vi altså gerne kunne komme med gode argumenter for at ekskludere de første skuffende data. Begrundelsen for det rimelige i at ekskludere data hænger sammen med den måde vore generator fungerer. Efter at have rystet beholderen skal vi vente på det radioaktive stof Pa-234 isoleres i det øverste lag i beholderen. Generatoren skal altså i en vis forstand startes op.
For at vurdere hvor mange data vi skal ekskludere vil vi nu transformere vores korrigerede henfaldsdata med logaritme funktionen. Samtidigt frembringer vi en graf hvor vi har afsat **tid** på 1.aksen og de log-transformerede korrigerede henfaldsdata (**log_kor**) på 2.aksen:



	A tid	B henfald	C kor_henfald	D log_kor
◆			=henfald-7.	=log(kor_henfald)
1	10	328	321.	2.50651
2	20	430	423.	2.62634
3	30	453	446.	2.64933
4	40	484	477.	2.67852
5	50	458	451.	2.65418
6	60	454	447.	2.65031
7	70	413	406.	2.60853
8	80	414	407.	2.60959
9	90	400	393.	2.59439
10	100	305	298.	2.47422
11	110	268	261.	2.41664
12	120	272	265.	2.42325
13	130	248	241.	2.38202
14	140	226	219.	2.34044
15	150	206	199.	2.29885
16	160	175	168.	2.22531
17	170	156	149.	2.17319
18	180	159	152.	2.18184
19	190	133	126.	2.10037
20	200	137	130.	2.11394
21	210	124	117.	2.06819
D3	=2.6493348587121			

Lidt matematisk teori, som også er gennemgået i kap 4, afsnit 5 om linearisering
Ifølge logaritmeregneregler kan vi vise at der for eksponentielle sammenhænge mellem en uafhængige variabel; *x* og en afhængige variabel; *y* gælder en lineær sammenhæng mellem *x* og $\log(y)$:

$$y = b \cdot a^x$$
$$\log(y) = \log(b \cdot a^x)$$
$$\log(y) = \log(b) + \log(a^x)$$
$$\log(y) = \log(b) + \mathrm{x} \cdot \log(a)$$
$$\log(y) = \log(a) \cdot \mathrm{x} + \log(b)$$

Hvis vi definerer $\alpha = \log(a)$ og $\beta = \log(b)$ ser vi at der gælder det påståede:

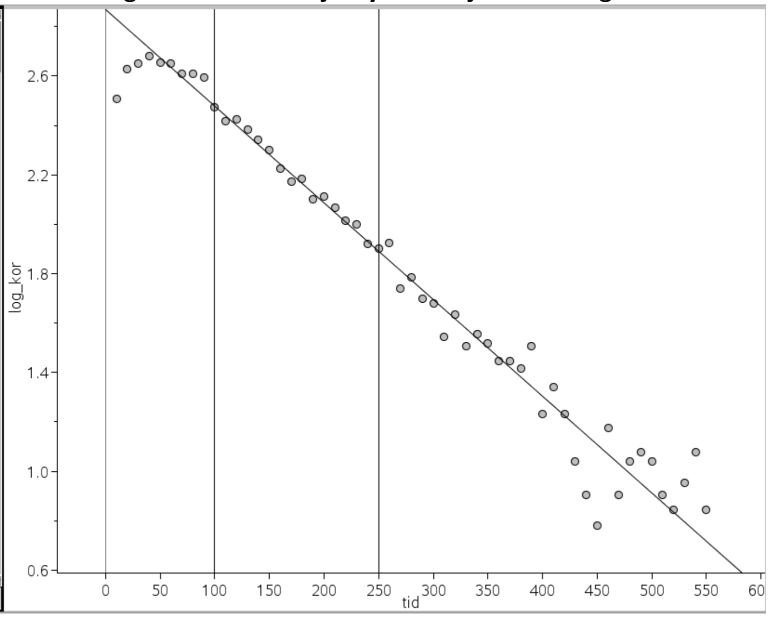
$$\log(y) = \alpha \cdot \mathrm{x} + \beta$$

Af definitionen af logaritmefunktionen med grundtal 10 gælder desuden $a = 10^{\alpha}$ og $b = 10^{\beta}$.

ISBN 9788770668699

Nu kan vi komme lidt længere i vores undersøgelse ved at **tilføje flytbar linje** til vores graf:

	C tid	D log_kor
◆		=log(kor_henfald)
1	10	2.50651
2	20	2.62634
3	30	2.64933
4	40	2.67852
5	50	2.65418
6	60	2.65031
7	70	2.60853
8	80	2.60959
9	90	2.59439
10	100	2.47422
11	110	2.41664
12	120	2.42325
13	130	2.38202
14	140	2.34044
15	150	2.29885
16	160	2.22531
17	170	2.17319
18	180	2.18184
19	190	2.10037
20	200	2.11394
21	210	2.06819
D9	=2.5943925503754	



Linje fitter vi til vores data således at den tilnærmelsesvis beskriver en lineær sammenhæng mellem tid og de logaritmisk transformerede data. Vi skal i denne del af vores undersøgelse ikke lave mindste kvadraters linje idet vi først skal have ekskluderet data der forringer modellen.
Frem til 100 sekunder ser vi at generatorens opstart ”smudser” vores forventede sammenhæng til idet der her en systematik: De første fem datapunkter alle ligger under linjen og de følgende fire alle ligger over. Vi kan på den baggrund ekskludere de 9 første data.
Vi bemærker desuden at der i intervallet fra 100 til 250 sekunder gælder at vores data sig fordeler tilfældigt og snævert omkring den rette linje. Efter 250 sekunder fordeler datapunkterne sig stadig tilfældigt omkring den rette linje men med en tindens til større trompetformede afvigelser, altså desto længere tid der går desto større residualer. Grunden til dette skyldes baggrundsstrålingen. I starten udgør baggrundsstrålingens udsving relativt lidt i forhold til generatorens aktivitet hvorimod dens tilfældige udsving får større relativ betydning efterhånden som aktiviteten i generatorens aktivitet henfalder.
Vi ekskluderer de ni første data. Desuden udelukker vi dataene efter 250 sekunder når vi efterfølgende bestemmer dén eksponentielle vækstmodel, der bedst forklarer sammenhængen mellem **tid** og **korrigeret henfald**. Det gør vi ved at markere og kopiere data fra 100 sekunder frem til 250 sekunder (inklusiv de dertil knyttede **log_kor** og **kor_hen** data). Dataene sættes ind i nye kolonner:

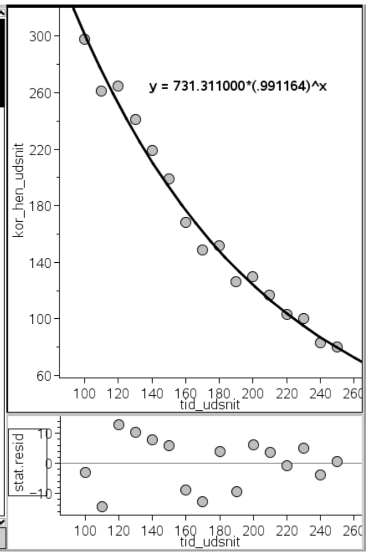
E tid_udsnit	F log_kor_udsnit	G kor_hen_udsnit
100	2.47422	298.
110	2.41664	261.
120	2.42325	265.
130	2.38202	241.
140	2.34044	219.
150	2.29885	199.
160	2.22531	168.
170	2.17319	149.
180	2.18184	152.
190	2.10037	126.
200	2.11394	130.
210	2.06819	117.
220	2.01284	103.
230	2.	100.
240	1.91908	83.
250	1.90309	80.

ISBN 9788770668699

Ved eksponentiel regression bestemmer vi sammenhængen mellem **tid_udsnit** og **kor_hen_udsnit** (udsnit af de korrigerede henfald):

	F tid_udsnit	G kor_hen_udsnit	H	I	J	K
◆				=ExpReg('tid		
1	100	298.	Titel	Eksponent...	rest	1024.17
2	110	261.	RegEqn	a*b^x	total	514977.
3	120	265.	a	731.311		
4	130	241.	b	.991164		
5	140	219.	r²	.987813	forklaring	.998011
6	150	199.	r	-.993888		
7	160	168.	Resid	{-3.064939...		
8	170	149.	ResidTra...	{-.0102325...		
9	180	152.				
10	190	126.				
11	200	130.				
12	210	117.				
13	220	103.				
14	230	100.				
15	240	83.				
16	250	80.				
17						
18						
19						
20						
21						

J6



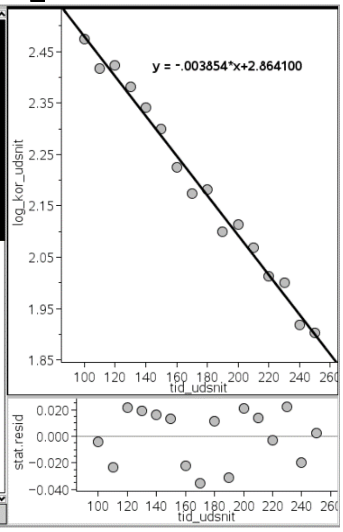
Der er ingen systematik i residualerne som er i størrelsesordenen ±15. Den logaritmisk udregnede forklaringsgrad er 98.78 % og mens vores egen udregnede forklaringsgrad med 99.80 % er lidt højere (udregnet på udsnittet af det korrigerede henfald). den eksponentielle vækst model er givet ved:

$$y = 731.311 \cdot 0.99116^x$$

Til sammenligning bestemmer vi ved lineær regression sammenhængen mellem **tid_udsnit** og **log_kor_udsnit** (udsnit af de logaritmisk korrigerede henfald). Bagefter viser vi hvorledes ud fra denne kan komme frem til den bedste eksponentielle sammenhæng mellem **tid_udsnit** og **kor_hen_udsnit**:

	F tid_udsnit	G log_kor_udsnit	H	I	J	K
◆				=LinRegMx('ti		
1	100	2.47422	Titel	Lineær regre...	rest	.006232
2	110	2.41664	RegEqn	m*x+b	total	.511354
3	120	2.42325	m	-.003854		
4	130	2.38202	b	2.8641		
5	140	2.34044	r²	.987813	forklaring	.987813
6	150	2.29885	r	-.993888		
7	160	2.22531	Resid	{-.004443918..		
8	170	2.17319				
9	180	2.18184				
10	190	2.10037				
11	200	2.11394				
12	210	2.06819				
13	220	2.01284				
14	230	2.				
15	240	1.91908				
16	250	1.90309				
17						
18						
19						
20						
21						

K10



Som før er der ingen systematik i residualerne som er i størrelsesordenen ±0.025. Den udregnede forklaringsgrad er 98.78 % som netop (**!**) svarer til den forklaringsgrad vi fik ved eksponentiel regression. Den lineære model er givet ved:

$$\log(y) = -0.00354 \cdot \mathrm{x} + 2.8641$$

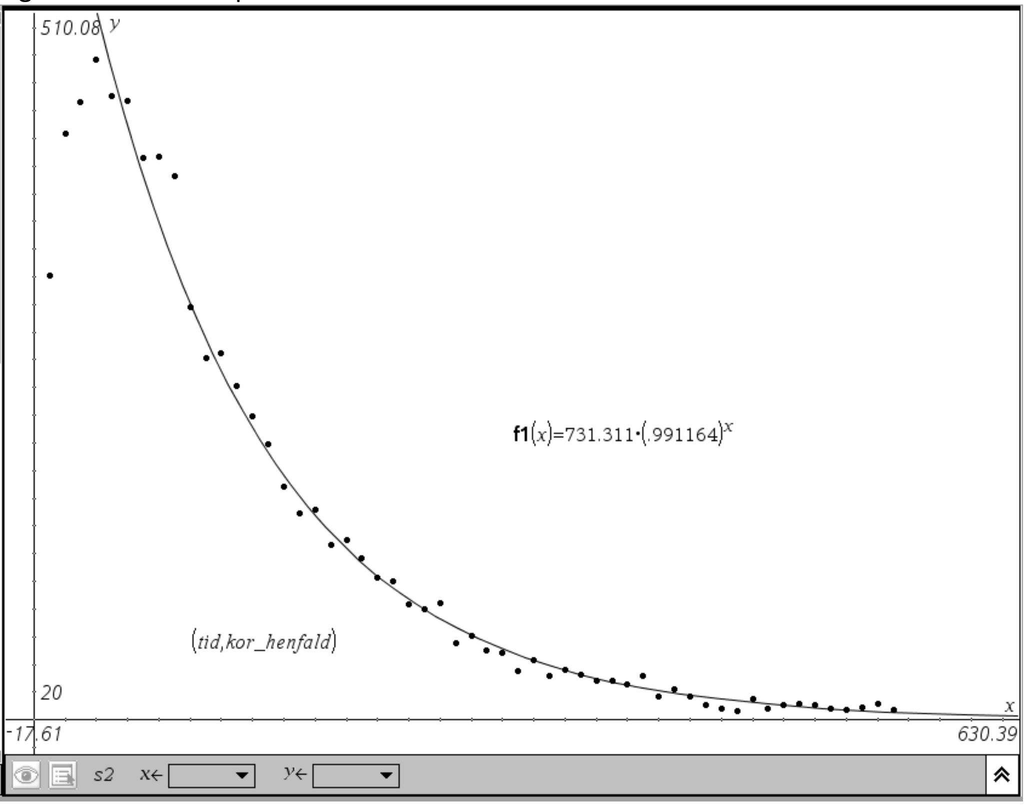
Her er $\alpha = \log(a) = -0.00354$ og $\beta = \log(b) = 2.8641$ hvormed vi ved "tilbage" transformation kommer frem til fremskrivningsfaktoren $a = 10^{\alpha} = 10^{-0.00354} = 0.99116$ og begyndelsesværdien $b = 10^{2.8641} = 731.311$, altså præcis ligesom før!

ISBN 9788770668699

Afslutningsvis tegner vi grafen for den eksponentielle vækstmodel sammen med vores rådata:

	A tid	B kor_henfald
◆		=henfald−7.
1	10	321.
2	20	423.
3	30	446.
4	40	477.
5	50	451.
6	60	447.
7	70	406.
8	80	407.
9	90	393.
10	100	298.
11	110	261.
12	120	265.
13	130	241.
14	140	219.
15	150	199.
16	160	168.
17	170	149.
18	180	152.
19	190	126.
20	200	130.
21	210	117.

B3 =446.



Det er ok.