

## Projekt 8.6 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 uranyl nitrat 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  $\alpha$ - og  $\beta$ -stråling er ikke kraftig nok til at trænge igennem væsken og beholderen og op til Geiger-Müller røret.

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

# Hvad er matematik? 1

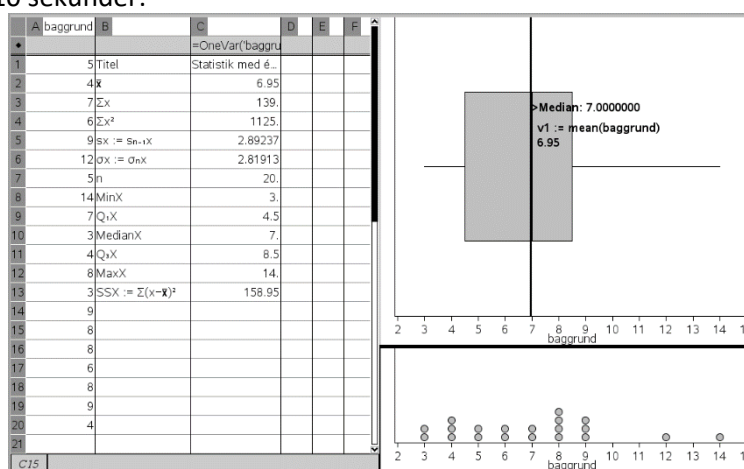
ISBN 9788770668279

Projekter: Kapitel 8. Projekt 8.6 Linearisering af data fra radioaktivt henfald

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

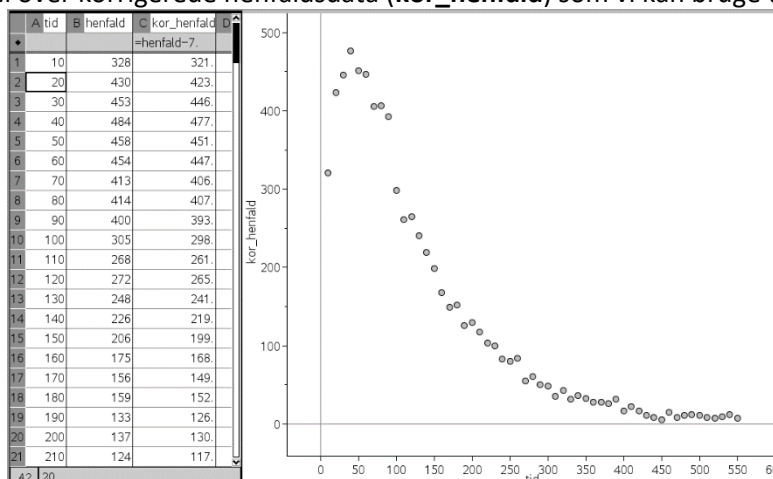
	A tid	B henfald		A tid	B henfald		A tid	B henfald
1	10	328	22	220	110	43	430	18
2	20	430	23	230	107	44	440	15
3	30	453	24	240	90	45	450	13
4	40	484	25	250	87	46	460	22
5	50	458	26	260	91	47	470	15
6	60	454	27	270	62	48	480	18
7	70	413	28	280	68	49	490	19
8	80	414	29	290	57	50	500	18
9	90	400	30	300	55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

Dataene kan hentes i et regnearkformat [her](#). Inden vi kommer alt for godt i gang med vores undersøgelse skal vi huske at korrigerer for baggrundstrålingen. Her får vi igen hjælp af gruppen idet de 20 gange har målt baggrundstrålingen over 10 sekunder:



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

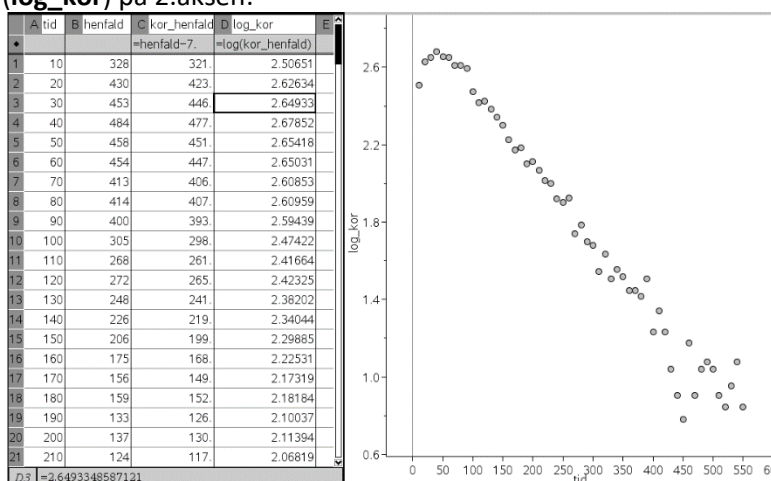
Det giver os en ny variabel over korrigerede henfaldsdata (**kor\_henfald**) som vi kan bruge til at lave en graf over:



Idet vi forventede at sammenhængen mellem **tid** og **korrigeret henfald** var eksponentielt virker grafen noget skuffende. Vi ser indledningsvis en stigning i  $\beta$ -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 hvis 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:



**Lidt matematisk teori, som også er gennemgået i kap 8, 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  $\log(y)$ :

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

$$\log(y) = \log(b \cdot a^x)$$

$$\log(y) = \log(b) + \log(a^x)$$

$$\log(y) = \log(b) + x \cdot \log(a)$$

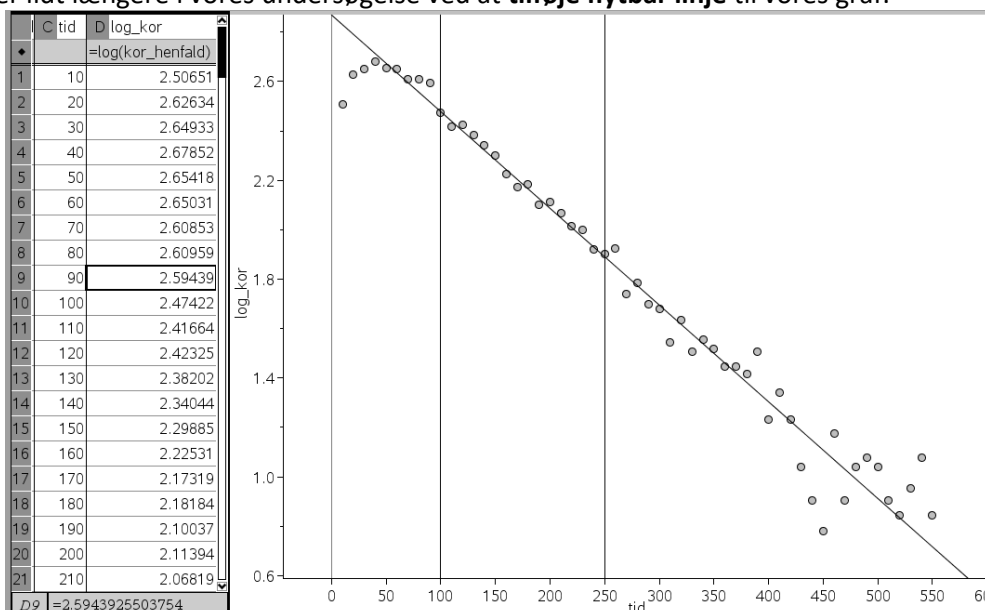
$$\log(y) = \log(a) \cdot x + \log(b)$$

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

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

Af definitionen af logaritme funktionen med grundtal 10 gælder desuden  $a = 10^\alpha$  og  $b = 10^\beta$ .

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



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.

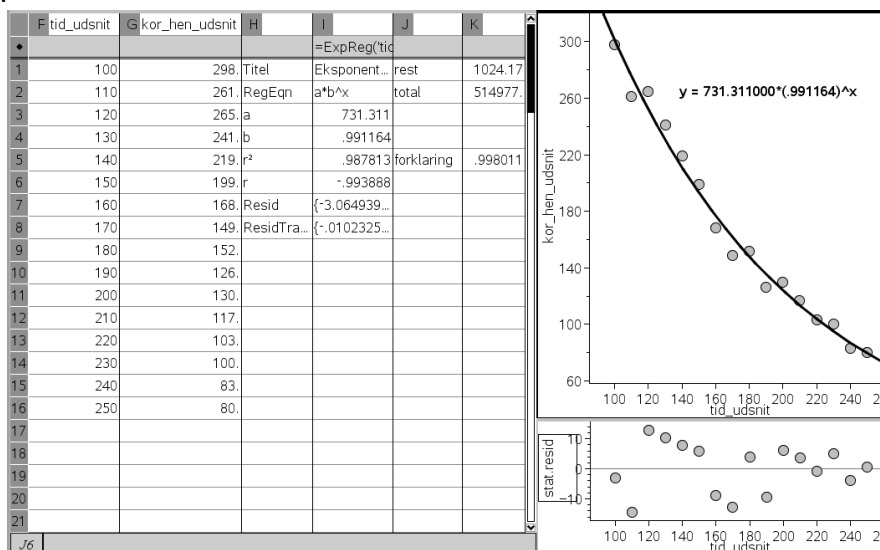
Frem til 100 sekunder ser vi at generatorens opstart "smudser" vores forventede sammenhæng til idet der her en systematik: De første fem datapunkter alle ligger under linjen og de følgende fire alle ligger over. Vi kan på den baggrund ekskludere de 9 første data.

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 tendens 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 den 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.

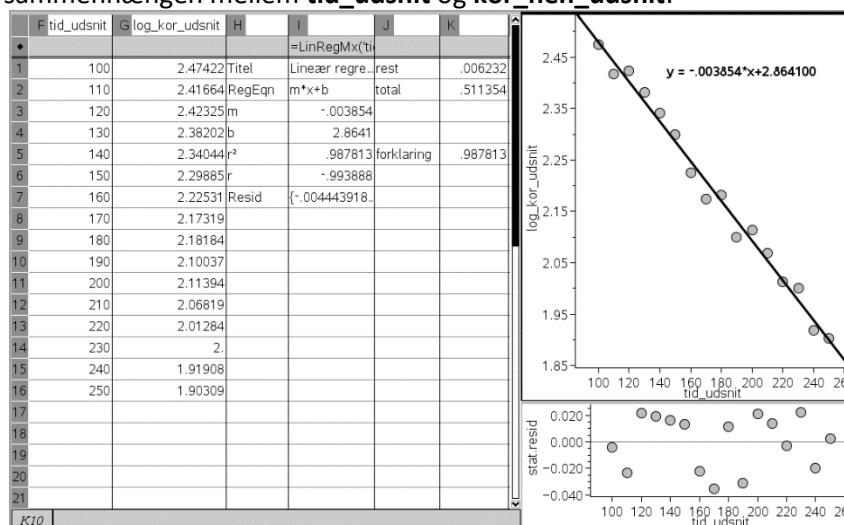
Ved eksponentiel regression bestemmer vi sammenhængen mellem **tid\_udsnit** og **kor\_hen\_udsnit** (udsnit af de korrigerede henfald):



Der er ingen systematik i residualerne som er i størrelsesordenen  $\pm 15$ . 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 sammenhængen mellem **tid\_udsnit** og **kor\_hen\_udsnit**:

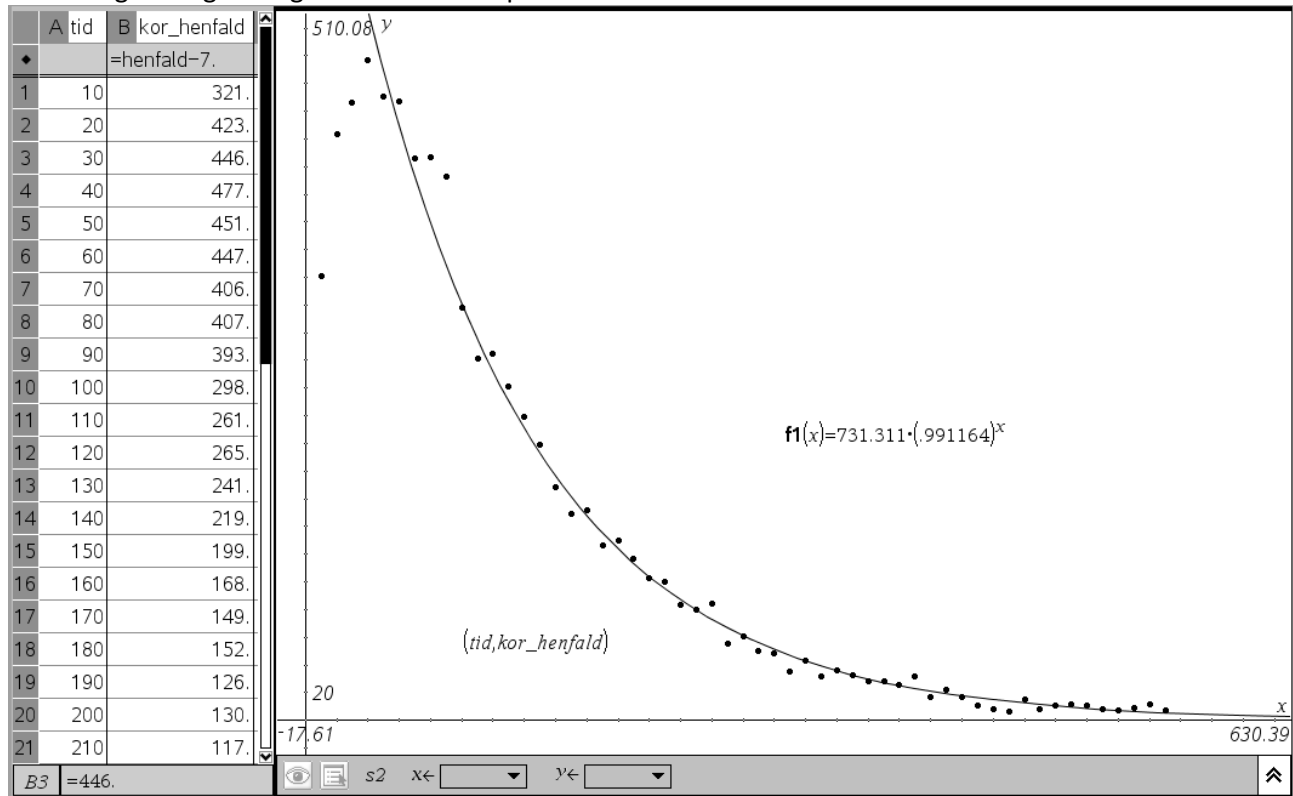


Som før er der ingen systematik i residualerne som er i størrelsesordenen  $\pm 0.025$ . Den udregnede forklaringsgrad er 98.78 % som netop (!) svarer til den forklaringsgrad vi fik ved eksponentiel regression. Den lineære model er givet ved:

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

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!

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



Det er ok.