

## Projekt 9.6 Box's helikopterekspirement – en indføring i simuleringsteknik med brug af TI Nspire

---

### Indhold

1. Modellering af fald med papirhelikopter: Et eksempel på lineær regression.....	2
Empiri.....	2
Helikopterekspirementet.....	2
2. Lineær regression .....	4
Mindste kvadraters metode .....	6
Residualplottet.....	8
En forenklet lineær model .....	9
3. Lineær regression som hypotesetest .....	11
Er der overhovedet en sammenhæng?.....	11
Når der nu er en sammenhæng, hvorfor er den så lineær? .....	15
Usikkerhedsvurdering: 95% konfidensintervaller for hældningen .....	15
Afsluttende bemærkninger: Hvad sker der, hvis der kun er én faldtid for hver faldhøjde? .....	19

Projektet er et oplæg til læreren og rummer et undervisningsforløb i lineær regression, der detaljeret gennemgår alle aspekter heraf og samtidig giver en indføring i, hvordan et værktøjsprogram kan anvendes i undersøgelsen af et datamateriale, bl.a. med en udførlig gennemgang af simuleringsteknikken, der ofte kaldes for "omrøring".

I afsnit 3 inddrages den bekræftende statistik med hypotesetest. Afsnittet indeholder en gennemgang af simuleringstest med omrøring. Endvidere en diskussion af fordele og ulemper ved at reducere et datamateriale og gå over til at se på middelværdier dermed inddrage det kanoniske t-test. Dette bliver foldet ud i en gennemgang af lineær regressionstest med inddragelse af konfidensintervaller.

Projektet bygger på et berømt eksperiment designet af den amerikanske statistiker George Box, der anvendte målinger på papirhelikopteres faldtid til at illustrerer væsentlige statistiske pointer. Vi har valgt at illustrere nogle af de mange muligheder med en konkret analyse af et datasæt, men tanken er selvfølgelig at klassen beslutter sig for hvilke aspekter af den lineære regressionsmodel projektet skal omfatte og designer eksperimentet og dataanalysen derefter.

## 1. Modellering af fald med papirhelikopter: Et eksempel på lineær regression

Modellering af lineære fænomener har en lang historie bag sig. Hovedmodellen er den lineære regressionsmodel baseret på mindste kvadraters metode, men den kan også udvides til en *hypotesetest* af den lineære afhængighed samt en analyse af konfidensintervaller for hældningen. I dette eksempel vil vi forsøge at gå hele vejen og undervejs omtale nogle faldgruber.

### Empiri

Selve eksperimentet er en klassiker, der stammer fra den indflydelsesrige amerikanske statistiker George Box (død i 2013), som havde nære forbindelser til statistikens fader Ronald Fisher. Det er en del af et større eksperiment om variabel kontrol og du kan høre George Box selv fortælle om det på videoklipet

[http://www.statisticsviews.com/details/video/4916021/A-Helicopter-Experiment-by-George-E\\_P\\_-Box-Sren-Bisgaard-and-Conrad-A\\_-Fung.html](http://www.statisticsviews.com/details/video/4916021/A-Helicopter-Experiment-by-George-E_P_-Box-Sren-Bisgaard-and-Conrad-A_-Fung.html)



På denne webside kan du også finde referencer til nogle af Box originale artikler, hvor han beskriver eksperimentet, ikke mindst

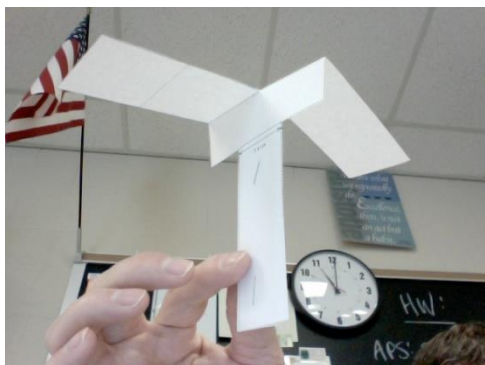
George Box: Teaching Engineers Experimental Design with a Paper Helicopter (1991)

Vi vil her udelukkende koncentrere os om sammenhængen mellem to variable, faldhøjden og faldtiden for en papirhelikopter, der slippes fra en aftalt højde.

*Bemærkning:* I fysikundervisningen er der tradition for et lignende eksperiment med kaffefiltre, der slippes fra forskellige højder. Det kan selvfølgelig sagtens erstatte helikopterstyrtet.

### Helikopterekperimentet

Men her tager vi altså udgangspunkt i en papirhelikopter, der klippes ud og foldes efter et fast design som vist og til sidst forsynes med en papirklips i halen. Eksperimentet kan nu udføres med støtte fra et målebånd og et stopur. Men man kan selvfølgelig også videofilme helikopterstyrtet og bruge videofilmene som udgangspunkt for dataanalysen.



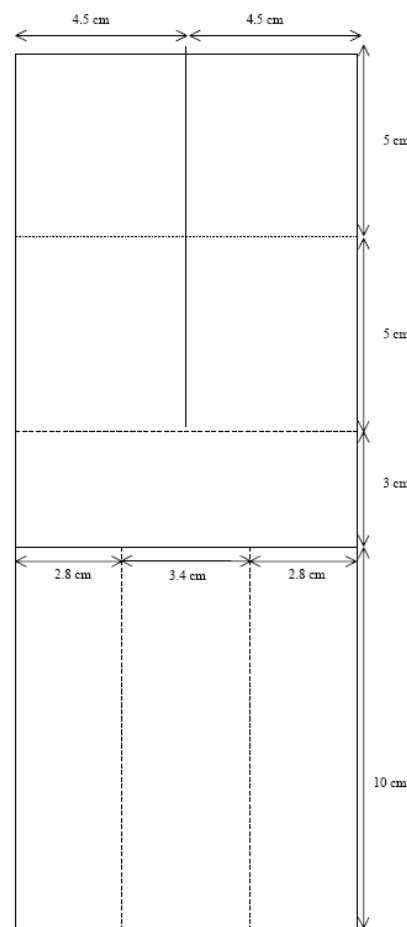
For at udføre eksperimentet skal klassen selvfølgelig selv bruge en skabelon, som du kan hente [her](#), og klippe og samle deres egen helikopter i stift papir eller karton. Det kan fx gøres gruppevis, men også enkeltvis. Det afgørende er at de bruger samme type papir og samme type klips, så alle helikopterne statistisk set er ens.

Herefter kan der udføres forskellige former for data-analyse. For at give en ide om hvad der kan gøres med sådanne data illustrerer vi dem med forsøgsdata fra en Amerikansk skole:

*Actual Sample Data* (Height is measured in meters, time in seconds)

Trial	Height: 1.77	Height: 2.29	Height: 2.83	Height: 3.36	Height: 4.39
1	1.49	1.77	2.36	2.60	3.38
2	1.12	1.62	2.04	2.65	3.42
3	1.39	1.66	2.27	2.80	3.91
4	1.19	1.64	2.30	2.58	3.13
5	1.35	1.28	1.75	2.31	2.63
6	1.32	1.20	1.97	2.33	2.58
7	1.28	1.63	2.26	2.73	2.86
8	0.73	1.83	2.20	2.74	2.89
9	1.32	1.72	2.27	2.56	3.27
10	1.53	1.64	2.29	3.04	2.99

PAPER HELICOPTER DESIGN  
Original Design by George Box

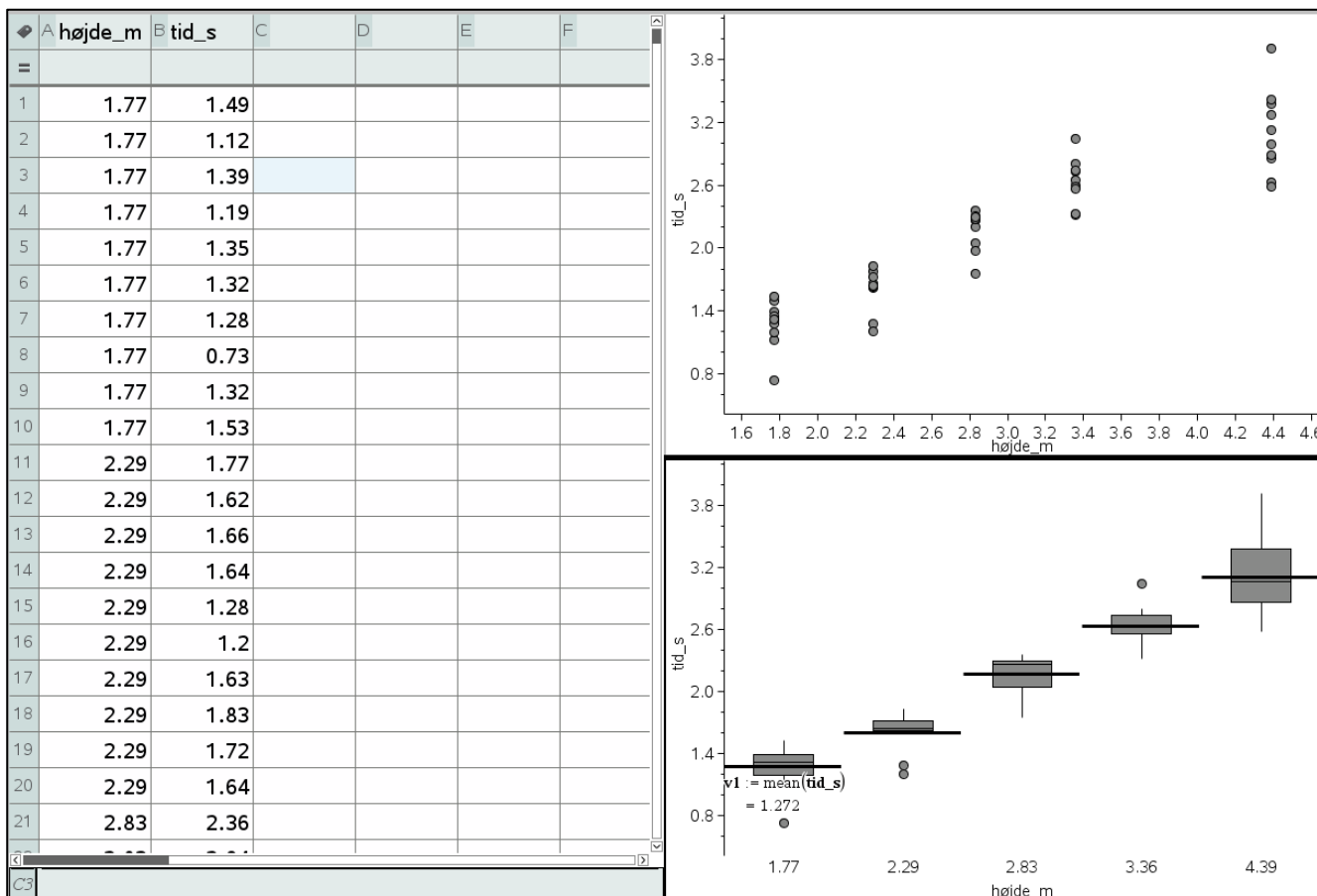


Som det ses, har man fastlagt fem bestemte højder inde i en idrætshal ud fra forskellige højder for tilskuerspladserne samt en balkon i den anden ende af hallen. Derefter har man delt klassen op i grupper, så man for hver af de fem højder har opmålt 10 helikopterstyrt. Tildelingen af højder til de enkelte grupper foregår ved lodtrækning (randomisering), for at undgå at eleverne selv optræder som konfunderende variable. Det er altså disse data vi vil analysere i det følgende, men det er selvfølgelig altid sjovere at analysere klassens egne data.

## 2. Lineær regression

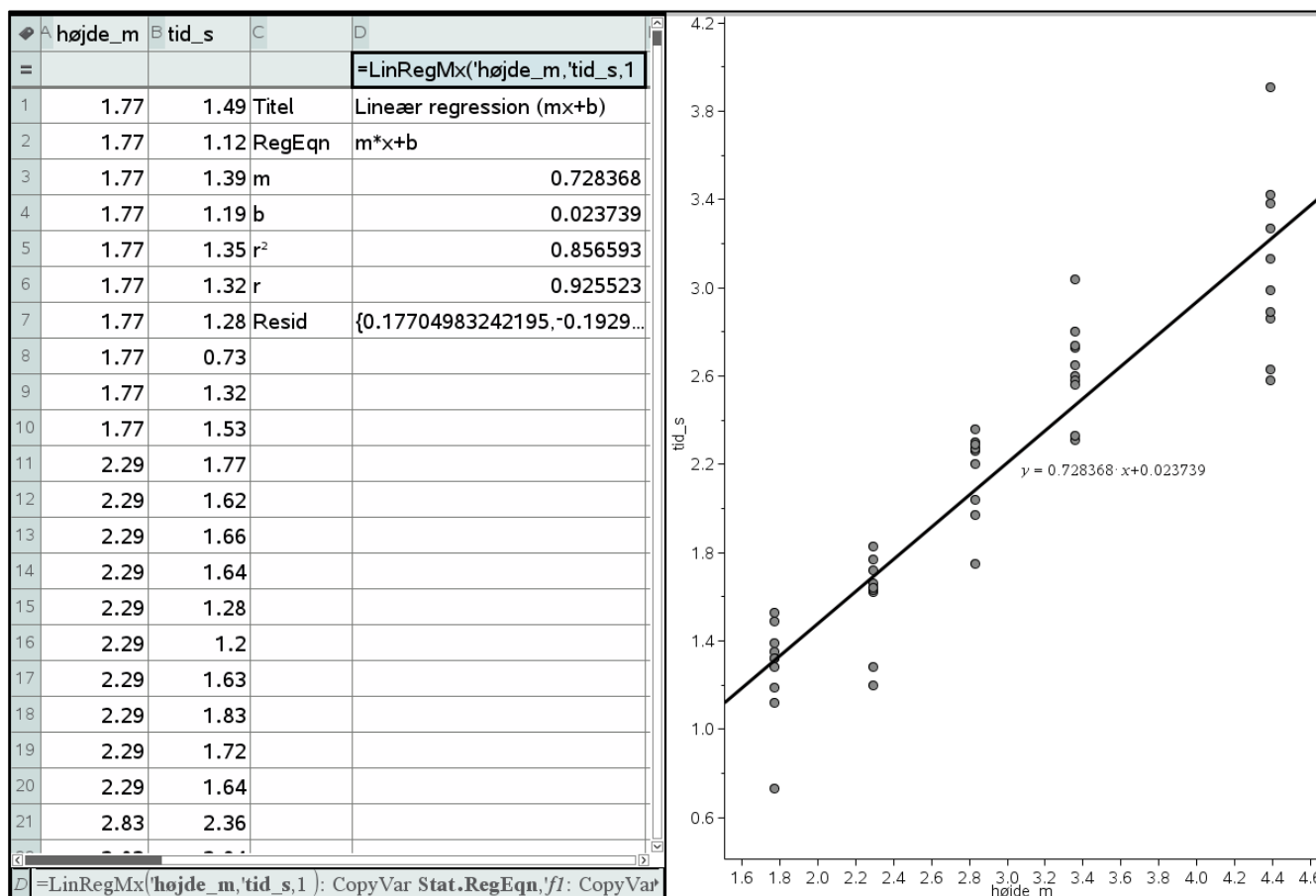
I det følgende benyttes TI-Nspire CAS. Men analysen kan selvfølgelig tilpasses andre CAS-værktøjer.

Dataene tages ind i et **Lister og Regneark**-værksted i TI-Nspire CAS og vi afbilder dem som et punktplot i **Diagrammer og Statistik**:



Vi har valgt faldhøjden som *den uafhængige variabel*, da den er fastlagt på forhånd, mens faldtiden er *den afhængige variabel*, som vi måler gentagne gange. Som det ses, er der en betydelig spredning i faldtiderne, men samtidigt en klar voksende tendens: Jo større faldhøjden er, jo større er faldtiden alt andet lige, selv om man godt kan finde individuelle målinger, der bryder med tendensen (fx udvalgte faldtider fra 2.29 meter, der er mindre end udvalgte faldtider fra 1.77 meter). Vi kan fremhæve strukturen af faldtidernes fordeling ved at tvinge faldhøjden til at blive afbildet som en kategorisk variabel. Det giver fem grupper af faldtider, der som vist kan afbildes som boksplot. Vi har også tilføjet middelfaldtiderne til boksplottene. Igen er den voksende tendens klar og det synes ikke urimeligt at benytte en lineær model. Man skal dog passe lidt på med det kategoriske plot, idet der i virkeligheden er længere afstand mellem de to sidste faldhøjder end mellem de øvrige. Man kan også med en vis rimelighed se de enkelte grupper af faldtider som rimeligt ens fordelte, om end det sidste boksplot synes at have en klar større spredning. Man kunne derfor overveje at droppe den sidste måling fra balkonen, som der var mindre styr over end de øvrige målinger (bl.a. fordi læreren opholdt sig i gymnastiksalen og derfor ikke kunne overvåge kvaliteten af målingerne fra balkonen i den anden ende af hallen). Men i den følgende analyse beholder vi alle data!

Vi udfører derfor nu en lineær regression på data:



Den lineære model får da ligningen

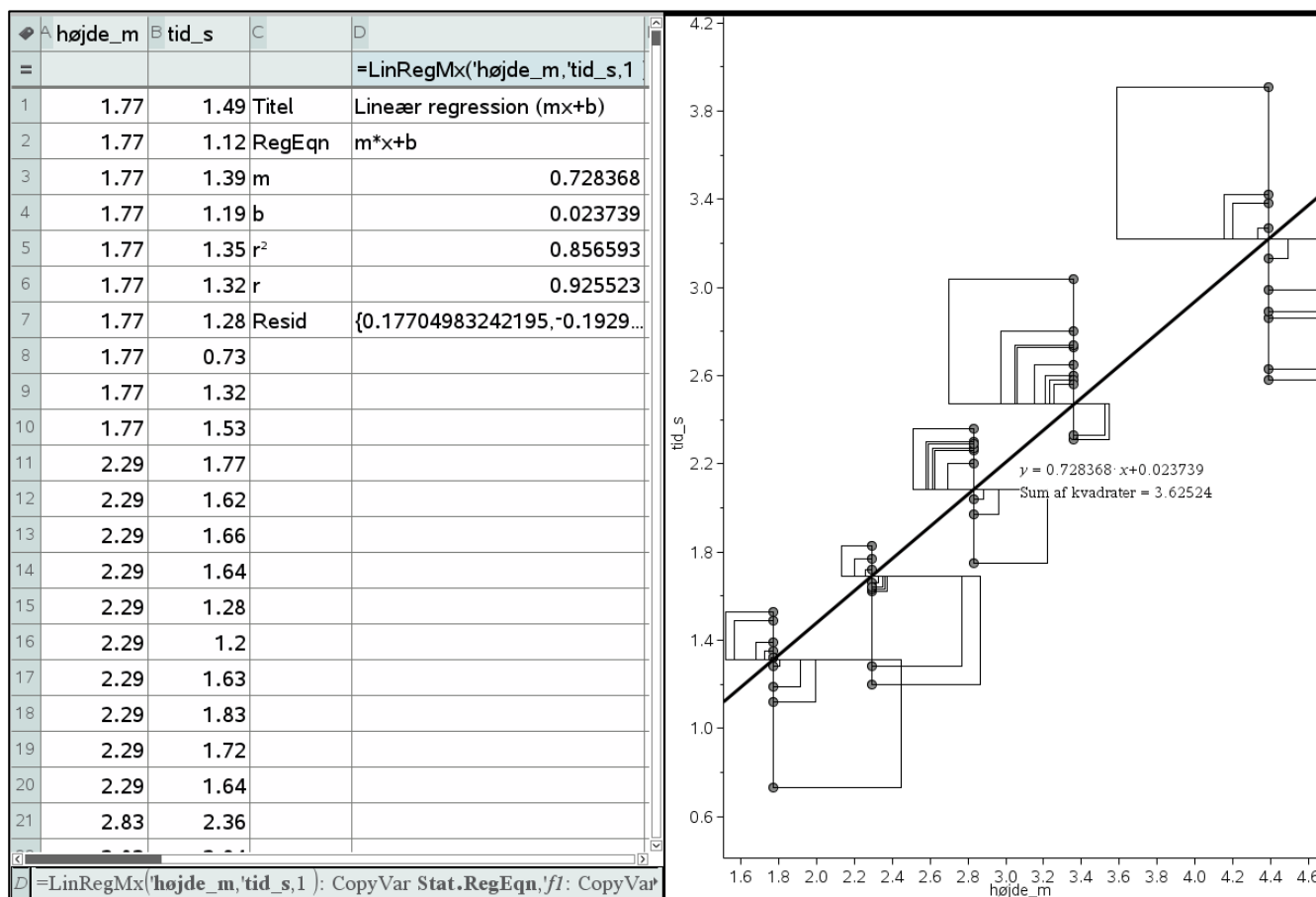
$$y = 0.728368 \cdot x + 0.023739$$

Her er det især hældningen, som er interessant. Den kan tolkes som en *faldkonstant*. Betydningen af faldkonstanten er den tid det tager en helikopter at falde 1 meter. Eller i matematik-jargon: For hver meter vi øger faldhøjden øges faldtiden alt andet lige med 0.73 sekunder.

Men hvor kommer selve linjen fra?

### Mindste kvadraters metode

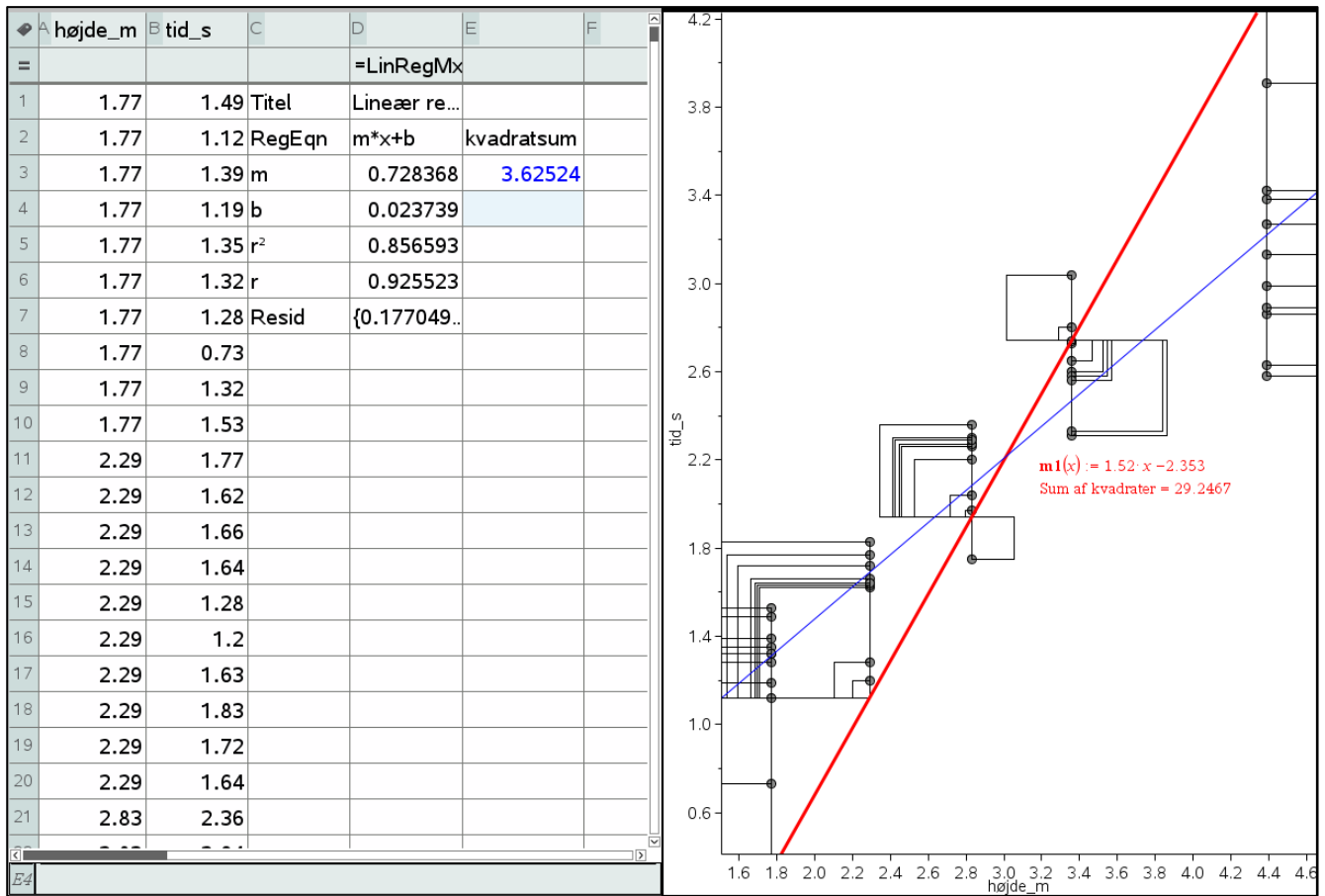
Vi højreklikker derfor på den lineære regressionslinje og vælger **Vis residuelle kvadrater**.



Som det ses bliver hvert datapunkt nu tilknyttet et kvadrat, hvis side netop er den lodrette afstand fra datapunktet til regressionslinjen. Ydermere får vi at vide, at summen af kvadraterne (dvs. summen af deres arealer) er givet ved 3.62524, et tal, som vi noterer i regnearket til senere brug.

Ideen er nu, at den lineære regressionsmodel er fastlagt ved mindste kvadraters metode, dvs. ved at gøre summen af disse kvadrater så lille som overhovedet muligt. Vi kan illustrere ideen ved at tilføje endnu en ret linje (vælg **Undersøg Data ▶ Tilføj Flytbare linjer**). Også denne linje kan vi nu højreklikke på for at få vist de residuelle kvadrater.

For at tydeliggøre billedet har vi nu farvet regressionslinjen blå og den flytbare linje rød. Vi har endvidere skjult de residuelle kvadrater for regressionslinjen, for at lade de residuelle kvadrater for den flytbare linje træde tydeligere frem!



Ved at trække i den flytbare linje kan man nu ved selvsyn konstatere at kvadratsummen for den flytbare linje altid ligger over kvadratsummen for regressionsmodellen og at jo tættere man lægger den flytbare linje opad regressionslinjen jo tættere kommer kvadratsummen for den flytbare linje til kvadratsummen for regressionslinjen.

Det er netop i overensstemmelse med mindste kvadraters metode, der siger at regressionslinjen udvælges blandt alle mulige linjer, så kvadratsummen er mindst mulig, dvs. så regressionslinjen i en vis forstand ligger tættest muligt på datapunkterne. Derfor kaldes regressionslinjen også ofte for *bedste rette linje*.

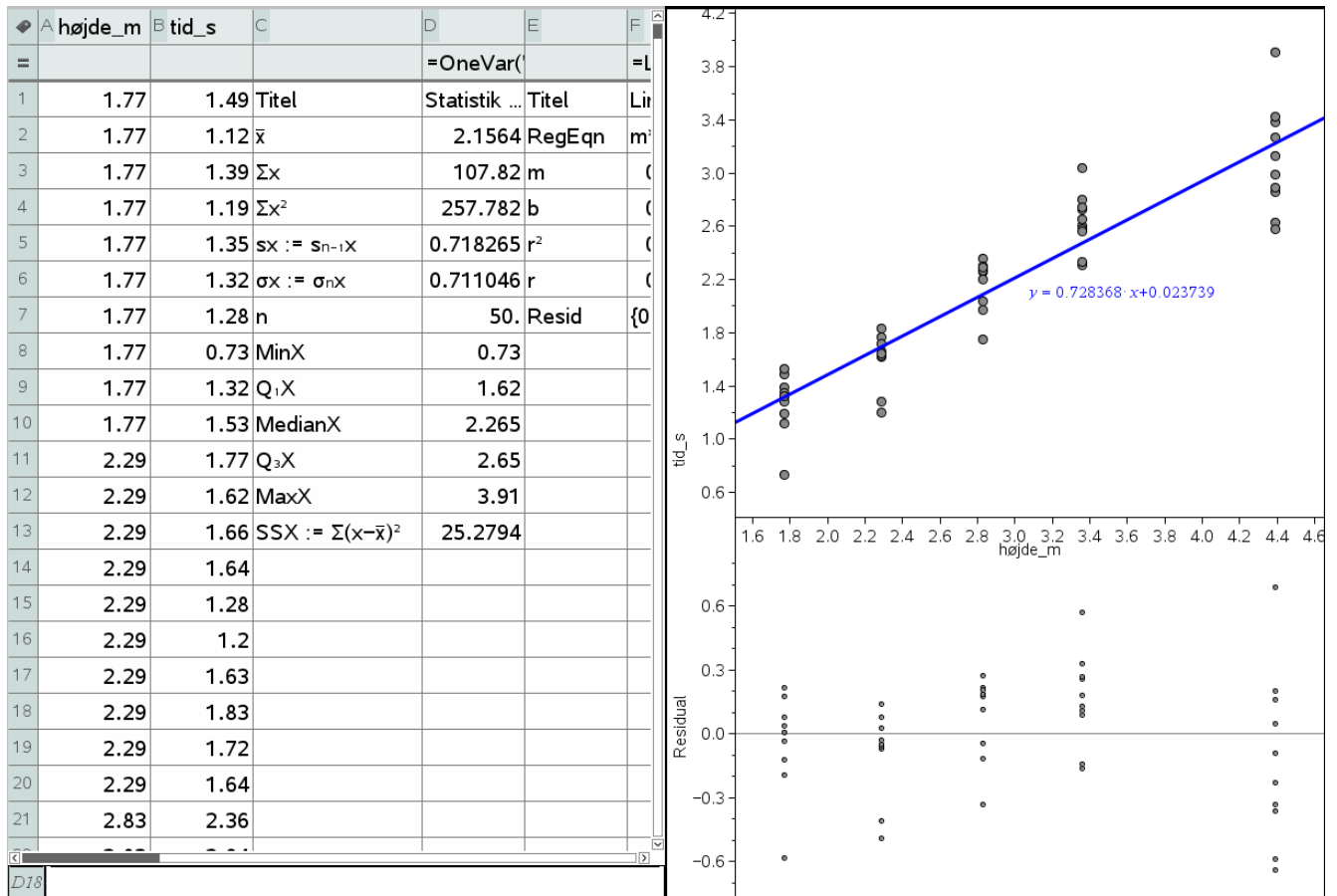
Vi ved nu hvor regressionslinjen stammer fra. Men det siger jo ikke noget om hvor god modellen er! Vi vil derfor nu konstruere et mål for modellens kvalitet, dvs. overensstemmelsen mellem datapunkterne og den lineære regressionsmodel. Den generelle ide er at vi opfatter data som en sum af modelværdien plus en afvigelse, det såkaldte residual:

$$\text{DATA} = \text{LINEÆR MODEL} + \text{RESIDUAL}$$

Den lineære model skal vælges så residualerne er så små som muligt, dvs. ved mindste kvadraters metode (dvs. det er faktisk summen af residualerne kvadrater, der er så lille som mulig). I vores tilfælde er kvadratsummen for den lineære regressionsmodel altså givet ved 3.62524.

### Residualplottet

Vi vælger nu : **Vis residual plot:**



Som det ses er der lidt slingren i residualerne, der typisk ligger mellem -0.5 og 0.5. Hvis den lineære model er en god model skal residualerne fordele sig fuldstændigt tilfældigt. Igen er den mest tydelige tendens nok den større spredning for den største faldhøjde på 4.39 meter. Men ellers er der næppe systematiske tendenser i residualplottet, så der er ikke noget der tyder på at vi ville vinde noget ved at skifte til fx en kvadratisk model eller en eksponentiel model.

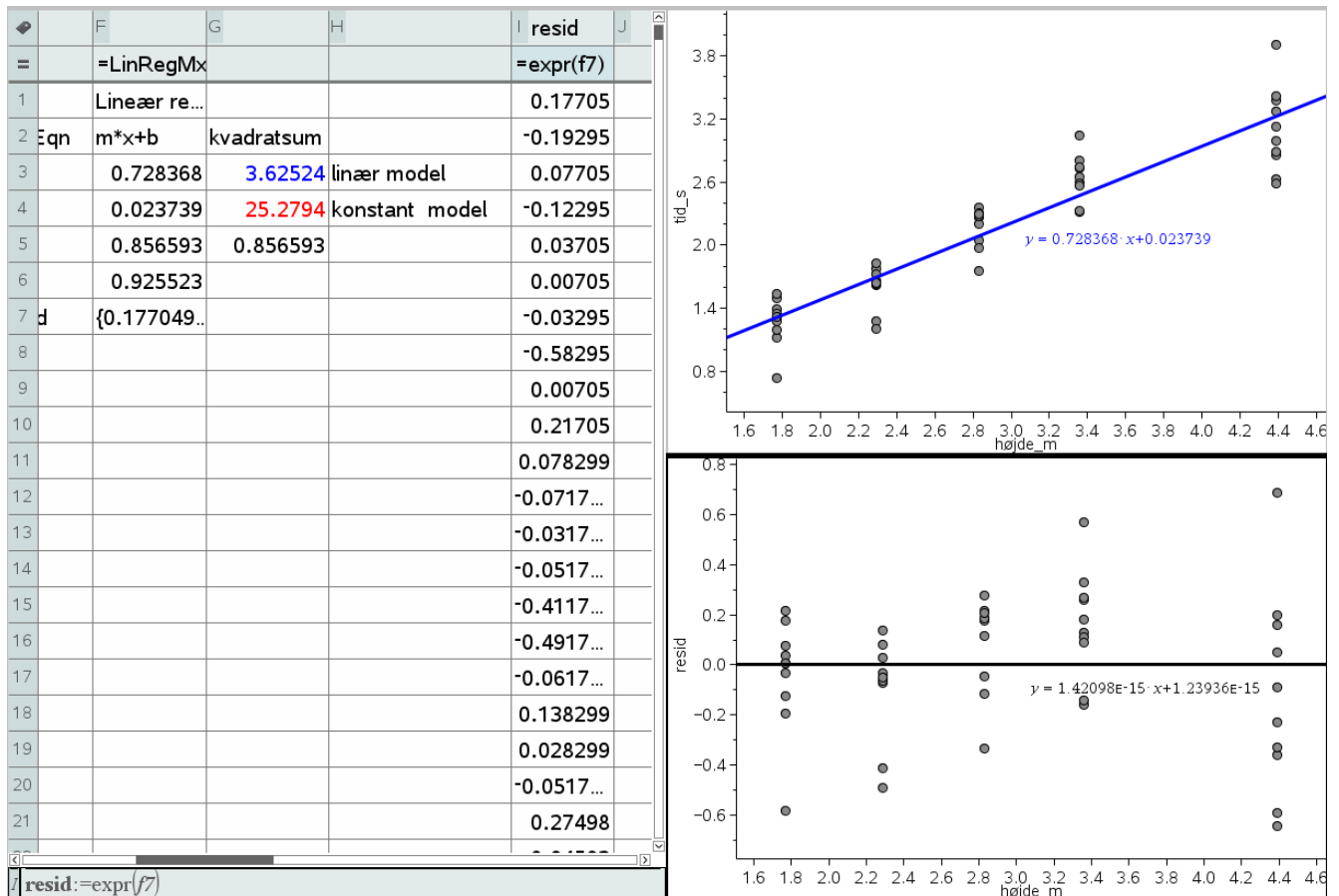
Til gengæld kan vi se at konstantleddet i den lineære regressionsmodel 0.023739 er så lille i forhold til residualerne at det ikke ville være urimeligt at sætte den til nul, dvs. at skifte til en proportionalmodel:

$$Faldtid = 0.728 \frac{s}{m} \cdot Faldhøjde$$

Men faldkonstanten findes stadigvæk mest præcist ud fra den lineære regressionsmodel!

Det automatisk genererede residualplot er rent illustrativt, dvs. vi kan ikke regne på det. For at kunne regne på residualerne er det nødvendigt at trække dem ud af modellen. De står pakket ind i gåseøjne i cellen F7. For at trække dem ud som en liste bruger vi derfor kommandoen **expr()**, der netop laver en tekststreng om til et matematisk udtryk. Derved kan vi som vist oprette et residualplot som et rigtigt punktplot i et separat **Diagrammer og Statistik**-vindue:





En af de interessante beregninger vi kan udføre på residualerne er nu at vi som vist kan udføre en lineær regression på dem! Ikke overraskende får vi resultatet

$$y = 1.421 \cdot 10^{-15} \cdot x + 1.239 \cdot 10^{-15}$$

Dvs. når vi tager højde for afrundingsfejl indenfor programmets beregningsområde er der reelt tale om  $y = 0$ . Vi har altså fjernet enhver lineær tendens ved at trække den lineære regressionsmodel ud af data. Og der er ikke andet.

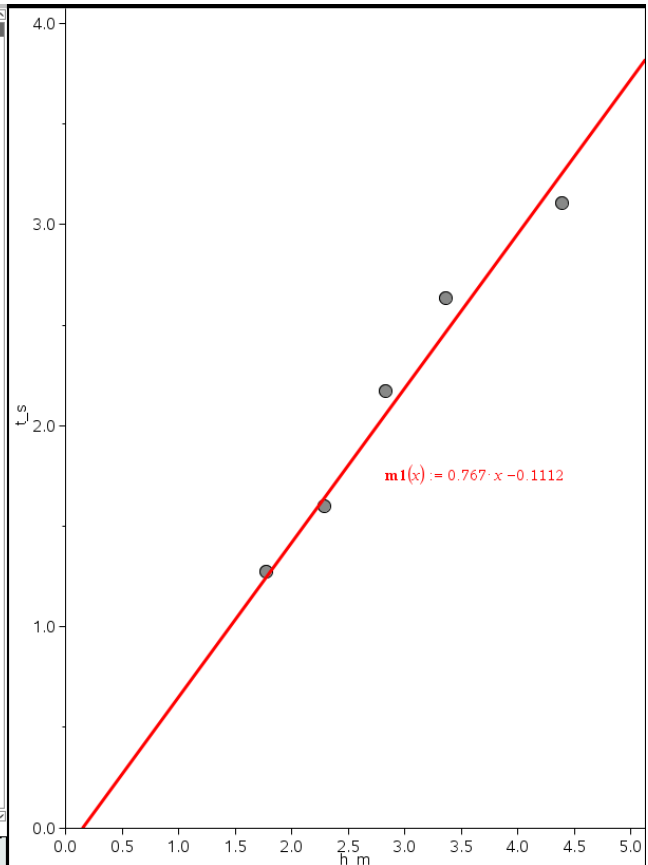
### En forenklet lineær model

Hvis man underviser i lineære modeller på et begynderniveau kan man vinde meget ved selv at lade eleverne finde den lineære model ved hjælp af flytbare linjer. Det er da nærliggende at forenkles datasættet ved at erstatte faldtiderne for en given højde med deres middelværdi, så man slipper for det meget grummede punktplot. Det nemmeste er da at udregne middelværdierne direkte fra de rå data, dvs. indføre to reducerede variable **h\_m**, der udregner middelværdierne for grupperne af faldhøjder og **t\_s**, der udregner middelværdierne for grupperne af faldtider. Der er så kun fem datapunkter i det reducerede datasæt og man kan med rimelighed selv trække en ret linje gennem datasværmen:

Projekter: Kapitel 9 Regression. Projekt 9.6 Box's helikopterekspériment – en indføring i simuleringsteknik med brug af TI Nspire

A	B	C	D	E	F
højde_m	tid_s	h_m	t_s		
1	1.77	1.49	1.77	1.272	
2	1.77	1.12	2.29	1.599	
3	1.77	1.39	2.83	2.171	
4	1.77	1.19	3.36	2.634	
5	1.77	1.35	4.39	3.106	
6	1.77	1.32			
7	1.77	1.28			
8	1.77	0.73			
9	1.77	1.32			
10	1.77	1.53			
11	2.29	1.77			
12	2.29	1.62			
13	2.29	1.66			
14	2.29	1.64			
15	2.29	1.28			
16	2.29	1.2			
17	2.29	1.63			
18	2.29	1.83			
19	2.29	1.72			
20	2.29	1.64			
21	2.83	2.36			

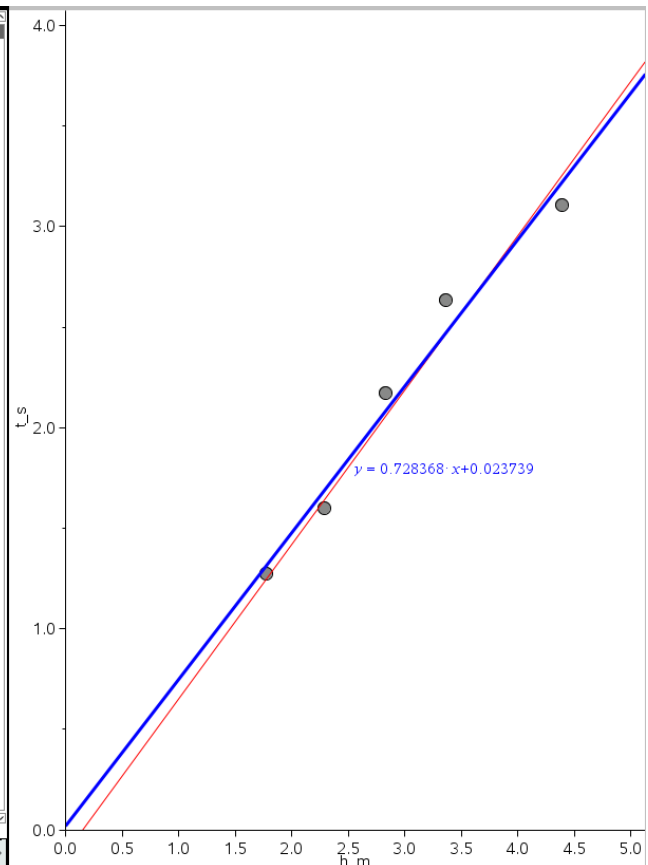
C7 =mean(a1:a10)



Men kunne da også få den ide at anvende lineær regression på det reducerede datasæt.

højde_m	tid_s	h_m	t_s	E	F
1	1.77	1.49	1.77	1.272	Titel
2	1.77	1.12	2.29	1.599	RegEqn
3	1.77	1.39	2.83	2.171	m
4	1.77	1.19	3.36	2.634	b
5	1.77	1.35	4.39	3.106	r <sup>2</sup>
6	1.77	1.32			r
7	1.77	1.28			Resid
8	1.77	0.73			
9	1.77	1.32			
10	1.77	1.53			
11	2.29	1.77			
12	2.29	1.62			
13	2.29	1.66			
14	2.29	1.64			
15	2.29	1.28			
16	2.29	1.2			
17	2.29	1.63			
18	2.29	1.83			
19	2.29	1.72			
20	2.29	1.64			
21	2.83	2.36			

F =LinRegMx('h\_m','t\_s,1'): CopyVar Stat.RegEqn,'f1: CopyVar Stat.,'



Men det skal man være varsom med! Den bedste rette linje får godt nok præcis den samme ligning som før, idet vi har komprimeret datasættet ved at trykke data sammen til deres middelværdier, hvilket netop fører til den mindste kvadratsum.

### 3. Lineær regression som hypotesetest

#### Er der overhovedet en sammenhæng?

Der er flere spørgsmål vi kan tage op under hypotesetesten. Hvis vi nu betragter forsøgsdataene som en stikprøve fra en stor population af data hvor ikke bare en enkelt klasse udfører forsøget, men rigtigt mange klasser udfører forsøget, kan vi så sige noget om hvorvidt den observerede lineære sammenhæng er statistisk signifikant?

Først må vi gøre os klart at der ikke findes nogen simpel sammenhæng mellem forklaringsgraden og den statistiske signifikans. En lineær sammenhæng kan sagtens være statistisk signifikant selv om forklaringsgraden er forbløffende lav, fx 5%. Der findes historisk berømte eksempler på sådanne situationer, fx Vietnam-lotteriet, se projekt 9.6 i C-bogen. Forklaringsgraden kan *ikke* udnyttes i hypotesetest! Vi må altså gå den slagte vej, sådan som vi fx kender den fra  $\chi^2$ -testen.

Vi må da først og fremmest opstille en *nulhypoteses* og vælge et *signifikansniveau*:

*Nulhypotesen*  $H_0$ : Der er ingen sammenhæng mellem faldhøjde og faldtid. Faldtiden er altså uafhængig af faldhøjden. De forskellige faldtider skyldes altså alene tilfældige fluktuationer.

Som signifikansniveau vælger vi som sædvanligt 5% og skal nu teste nulhypotesen. Det gør vi først rent eksperimentelt ved hjælp af en omrøringstest, dvs. vi omrører det givne rå datasæt for at bryde en eventuel sammenhæng mellem den observerede faldhøjde og den observerede faldtid.

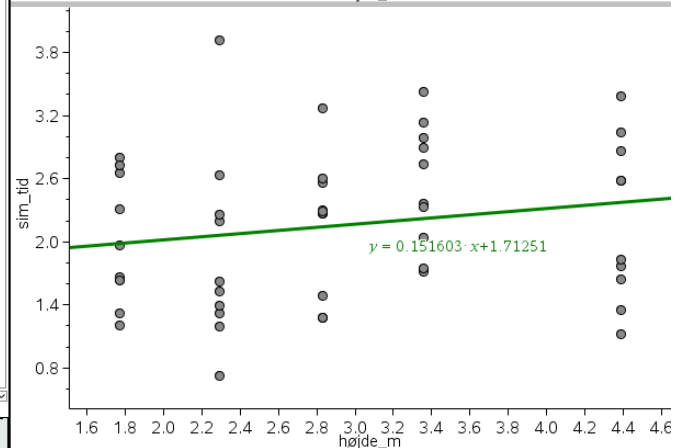
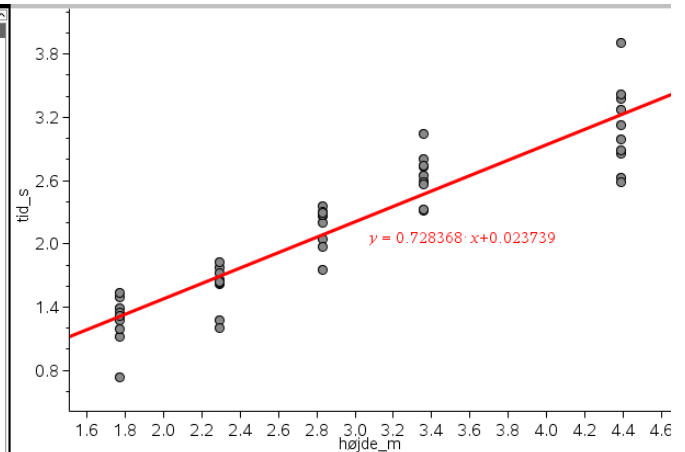
Vi indfører altså igen lister for de observerede faldhøjder **højde\_m** og de observerede faldtider **tid\_s**, men tilføjer denne gang også den omrørte tid:

```
sim_tid := randsamp(tid_s, 50, 1)
```

Her har vi trukket en tilfældig stikprøve på 50 elementer uden tilbagelægning, dvs. den simulerede tid, **sim\_tid**, er netop en tilfældig permutation af den observerede tid, **tid\_s**!

Vi kan så oprette punktplot for såvel den observerede tid som funktion af faldhøjden som den simulerede tid som funktion af faldhøjden:

A	B	C	D	E	F	G
	højde_m	tid_s	sim_tid			
=		=randsam		=LinRegM		=LinRegM
1	1.77	1.49	2.31	Titel	Lineær r...	Lineær r...
2	1.77	1.12	1.64	RegE...	m*x+b	RegE...m*x+b
3	1.77	1.39	2.65	m	0.728368	m
4	1.77	1.19	2.8	b	0.023739	b
5	1.77	1.35	1.66	r <sup>2</sup>	0.856593	r <sup>2</sup>
6	1.77	1.32	1.63	r	0.925523	r
7	1.77	1.28	1.2	Resid...	{0.17704...	Resi... {0.32915...
8	1.77	0.73	2.73			
9	1.77	1.32	1.32			
10	1.77	1.53	1.97			
11	2.29	1.77	3.91			
12	2.29	1.62	1.53			
13	2.29	1.66	0.73			
14	2.29	1.64	2.63			
15	2.29	1.28	2.2			
16	2.29	1.2	1.19			
17	2.29	1.63	1.32			
18	2.29	1.83	1.39			
19	2.29	1.72	1.62			
20	2.29	1.64	2.26			
21	2.83	2.36	1.28			
C sim_tid:=randsamp(tid_s,50,1)						

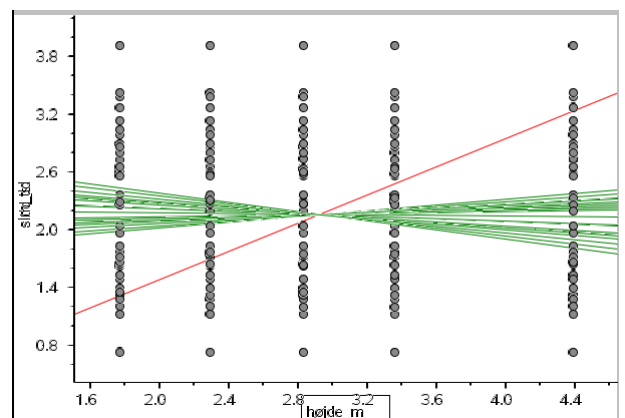


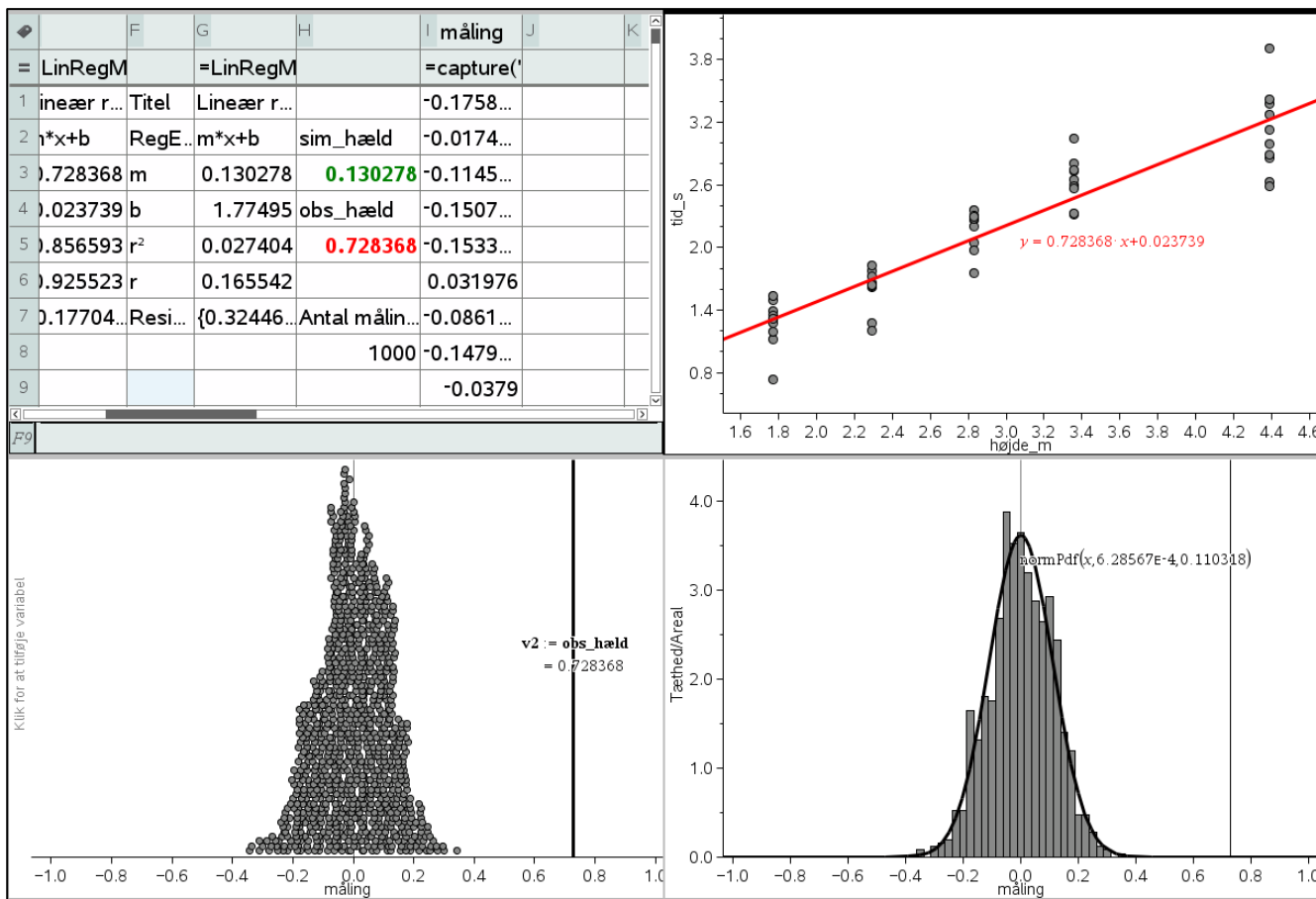
Taster vi nu **Ctrl/Cmd R** inde i regnearket vil vi få genberegnet simulationen igen og igen. Vi kan så se hvor nemt det er at frembringe en hældning, der er lige så stor som den observerede, dvs. vi kan skønne over p-værdien. De simulerede hældninger vil være såvel positive som negative og i middel vil de være 0 i overensstemmelse med nulhypotesen.

I en proto-test prøver man netop 20 gange (fordi signifikansniveauet er 5% = 1/20). Hvis der slet ikke forekommer nogen ekstreme hældninger skønner vi derfor at p-værdien er under 5% og nulhypotesen falder.

Her har vi samlet billederne af de første 20 simuleringer i en fælles graf:

Vi er ikke i nærheden af den røde observerede hældning, så proto-testen antyder kraftigt at p-værdien er et godt stykke under 5% og dermed at nulhypotesen må afvises.





Hvis vi vil udføre en mere præcis uafhængighedstest kan vi overføre værdien af den observerede hældning til variabelen **obs\_hæld** (det sker i rød celle H5 med celleformlen:  $obs\_hæld:=e3$ ) og tilsvarende den simulerede hældning til variabelen **sim\_hæld** (det sker i grøn celle H3 med celleformlen:  $obs\_hæld:=g3$ ).

Vi kan så udføre en automatisk datafangst af den simulerede hældning, hvor vi gentager fangsten 1000 gange (vi holder øje med antallet af målinger i celle H8 med celleformlen  $=count(måling)$ ).

Målingerne afsættes i to **Diagrammer og statistik**-vinduer, til venstre som prikdiagram og til højre som histogram overlejret med en normalfordeling. Som det ses er middelhældningen meget tæt på 0. I begge diagrammer har vi også indsat den observerede hældning.

Ingen af de 1000 simuleringer kommer i nærheden af den observerede hældning så skønnet over p-værdien ligger et godt stykke under 1 promille. Nulhypotesen må altså forkastes!

Vi kan godt danne os et skøn over p-værdiens størrelse ved at udnytte normalfordelingsapproximationen og finde arealet til højre for den observerede hældning. Da nulhypotesen ikke favoriserer et bestemt fortegn kunne de ekstreme hældninger lige så godt være negative og vi skal derfor gange det fundne areal med 2. Arealet findes ved at højreklikke på normalfordelingskurven og vælge **Skraver under funktion**. Arealet er dog under 0.0000, så vi kan ikke finde decimalerne på den måde. I stedet må vi udføre en beregning i et **Note**-værksted:

Skøn over p-værdi:

$$2 \cdot \text{normCdf}(0.728367, \infty, 6.28567 \cdot 10^{-4}, 0.110348) \rightarrow 4.2764412 \text{E-}11$$

Vi slutter af med en såkaldt *kanonisk test*, idet vi udfører det indbyggede lineære regressionstest:

	I måling	J	K
=	=capture('sim_hæld		=LinRegtTest('højde_m','tid
1	-0.175816	Titel	Lineær regressions t-tes...
2	-0.017465	Alternativ Hyp...	$\beta$ & $\rho \neq 0$
3	-0.114572	RegEqn	$a+b \cdot x$
4	-0.150787	t	16.932548
5	-0.153328	PVal	7.062071E-22
6	0.031976	df	48.
7	-0.086101	a	0.02373928
8	-0.147999	b	0.72836773
9	-0.0379	s	0.2748197
K5	=7.0620709504747E-22		

Ifølge nulhypotesen er såvel hældningen  $\beta$  som korrelationskoefficienten  $\rho$  lig med 0. Ifølge den alternative hypotese er de begge forskellige fra 0. Igen fås en astronomisk lille p-værdi og nulhypotesen må på det bestemteste afvises!

Læg mærke til antallet af frihedsgrader som er 48, dvs. 2 mindre end de 50 datapunkter. Det skyldes de to bånd der lægges på residualerne.

### Når der nu er en sammenhæng, hvorfor er den så lineær?

Vi har nu afvist nulhypotesen om uafhængighed og må derfor vælge den alternative hypotese: Faldtiden er afhængig af faldhøjden og afhængigheden er tydeligvis voksende. Men hvorfor lige netop en lineær afhængighed, hvorfor ikke en kvadratisk eller en eksponentiel eller ...? Det er straks sværere at svare på, men vi er nødt til at forsøge at begrunde den lineære model! Som støtte for den lineære model kan vi nu fremføre to typer argumenter:

Det ene kommer fra *den deskriptive statistik*: Hvis der er andet på spil end den lineære sammenhæng burde data vise en tydelig krumning hvilket også burde slå igennem på residualplottet i form af fx en tydelig hængebug. Men det har vi ikke set, så krumningen er ikke specielt åbenlys.

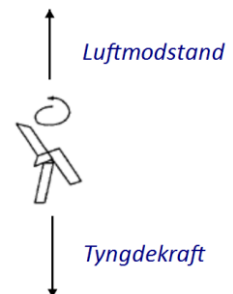
Det andet kommer fra en *bagvedliggende teori*. Vi kan håbe på en simpel fysisk forklaring på hvorfor vi bør forvente en lineær sammenhæng. Og det kan man faktisk godt give:

Faldet deles i to faser:

Den indledende fase: Helikopteren begynder at falde og rotere. I løbet af tidsrummet  $t_0$  falder det dybden  $s_0$ , hvorefter faldet stabiliseres, fordi luftmodstanden er vokset i størrelse, indtil den netop matcher tyngdekraften. Herefter falder helikopteren jævnt med farten  $v$ , der netop er afpasset så luftmodstanden er lige så stor som tyngdekraften! Vi forventer altså en sammenhæng af typen

$$s - s_0 = v \cdot (t - t_0)$$

$$t = t_0 + \frac{1}{v} \cdot (s - s_0) = \frac{1}{v} \cdot s + \left( t_0 - \frac{1}{v} \cdot s_0 \right)$$



Den forventede sammenhæng mellem faldhøjden  $s$  og faldtiden  $t$  er netop lineær. Læg mærke til at denne simple forklaring *ikke* favoriserer en proportional sammenhæng.

Som alternativ hypotese vælger vi altså en model på formen

$$\begin{aligned} \text{Data} &= \text{Lineær model} && + \text{Residual} \\ \text{Faldtid} &= a \cdot \text{Faldhøjde} + b && + \varepsilon \end{aligned}$$

Denne opskrivning betyder, at den systematiske sammenhæng ligger i den lineære model og residualerne anses for at variere fuldstændigt tilfældigt med middelværdi 0. I videregående statistik vil man ofte antage at residualerne for de enkelt faldhøjder følger en normalfordeling med middelværdi 0 og en fælles ukendt spredning  $\sigma$ .

### Usikkerhedsvurdering: 95% konfidensintervaller for hældningen

Det fører os til den sidste problemstilling. Når først vi *har* accepteret den lineære model

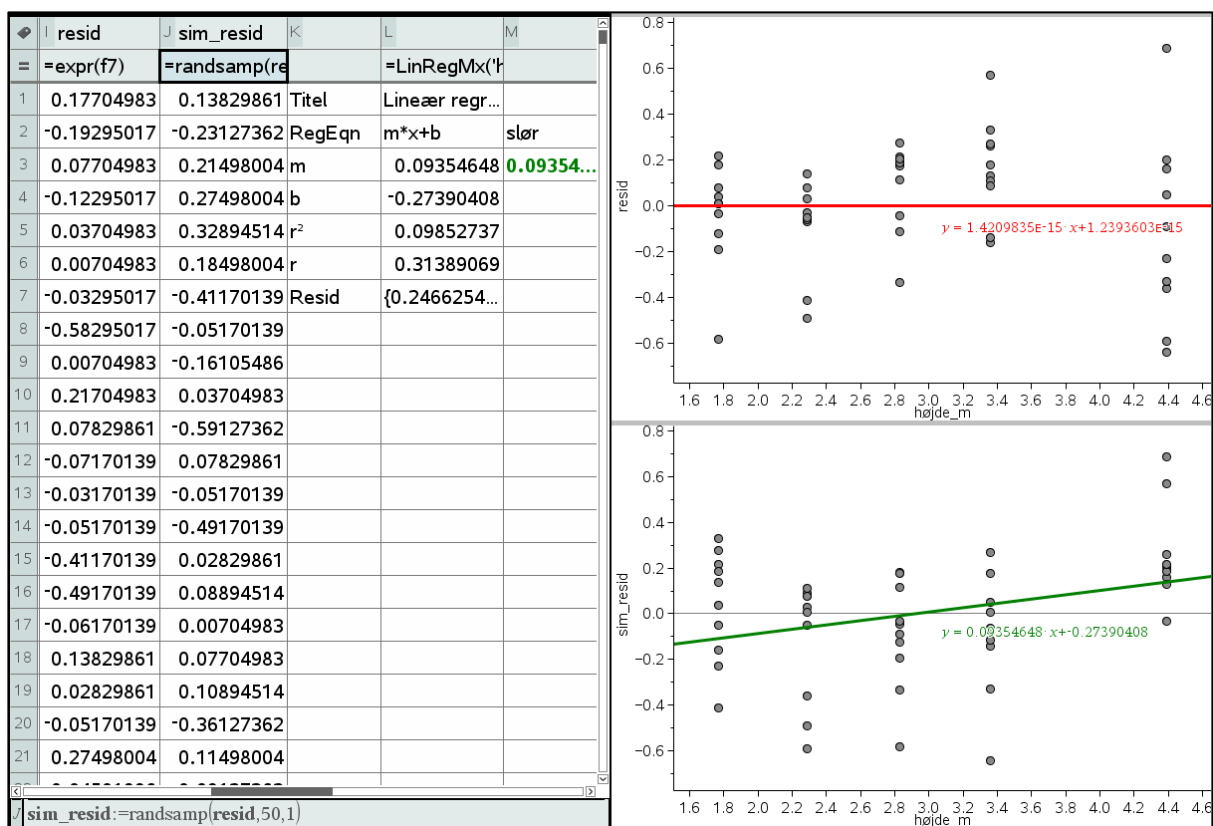
$$y = 0.728368 \cdot x + 0.023739$$

hvor sikre kan vi så være på hældningen? Det er klart at jo mere residualerne spreder sig jo større usikkerhed er der på hældningen. Vi kan nu eksperimentelt fastlægge et 95% konfidensinterval for hældningen på

følgende måde: Vi accepterer den lineære model som udtryk for den systematiske sammenhæng, men residualerne svinger tilfældigt. Gentages forsøget vil vi derfor få nogle andre residualer og dermed en anden hældning for regressionsligningen.

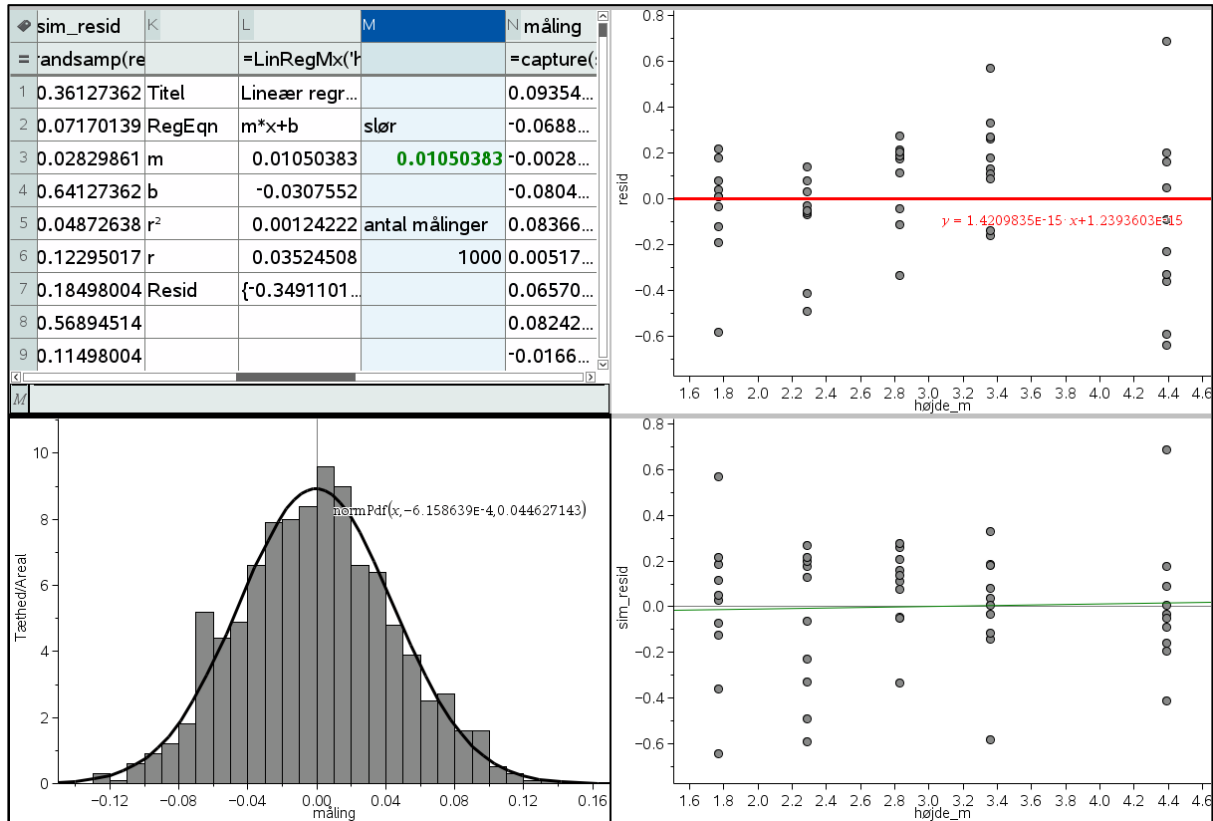
Vi kan simulere det slør, der kommer fra de tilfældige variationer i residualerne ved at omrøre residualerne! Vi trækker derfor residualerne ud som en selvstændig liste og foretager en omrøring i residualerne ved hjælp af **randsamp**-kommandoen lige som før (se figur næste side):

Ser vi nu på de simulerede residualer, så har de faktisk en hældning, men i middel er den selvfølgelig 0, når vi gentager simuleringen uendeligt mange gange. Vi overfører nu den simulerede hældning til variabelen slør ved hjælp af cellekommandoen `slør:=L3`.

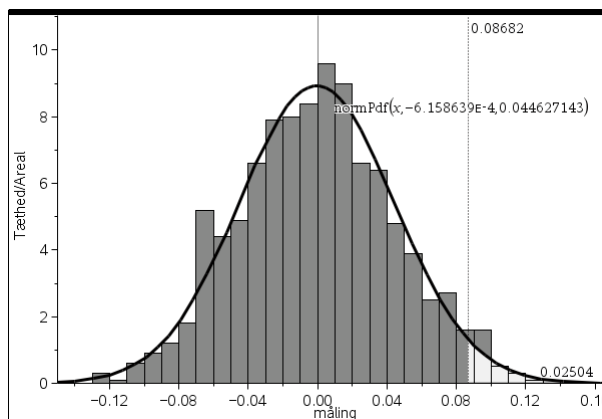


Herefter kan vi lave automatisk dataopsamling på sløret ligesom før. Vi gentager simuleringen af residualerne 1000 gange:



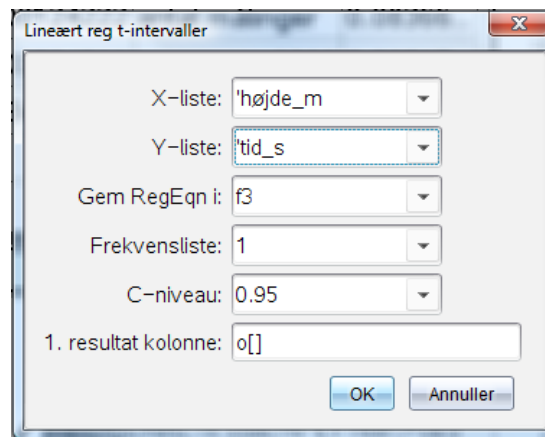
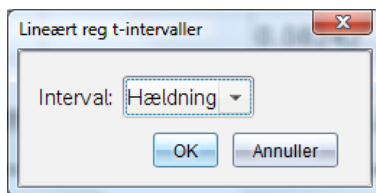
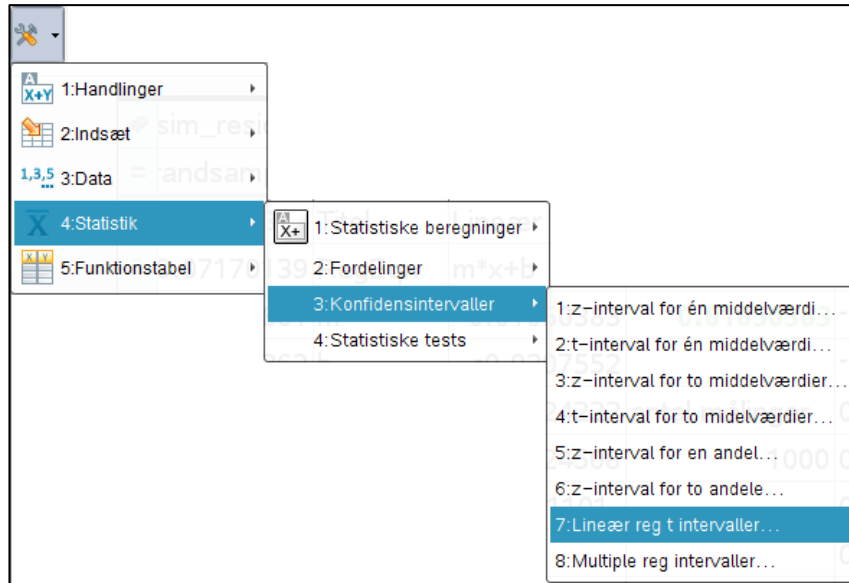


De 1000 simulerede residualhældninger er da med god tilnærmelse normalfordelte med middelværdi meget tæt på 0 og en spredning meget tæt på 0.0446. Vi skal da have fundet de 95% midterste simulerede hældninger. Det gøres nemmest ved at højreklikke på normalfordelingen, vælge **Skriver under funktion** og justere arealet så vi lægge halvdelen, dvs. 2.5% ude i den højre hale:



Det passer godt med tommelfingerreglen der siger, at 95%-konfidensintervallet svarer til to spredninger! Usikkerheden på hældningen skønnes altså til at være på 0.087, dvs. faldkonstanten (den tid det tager helikopteren at styrte 1 meter) bør opgives som  $0.728 \pm 0.087$  svarende til usikkerhedsintervallet  $[0.641; 0.815]$ .

Igen kan vi supplere med en *kanonisk* beregning af 95%-konfidensintervallet:

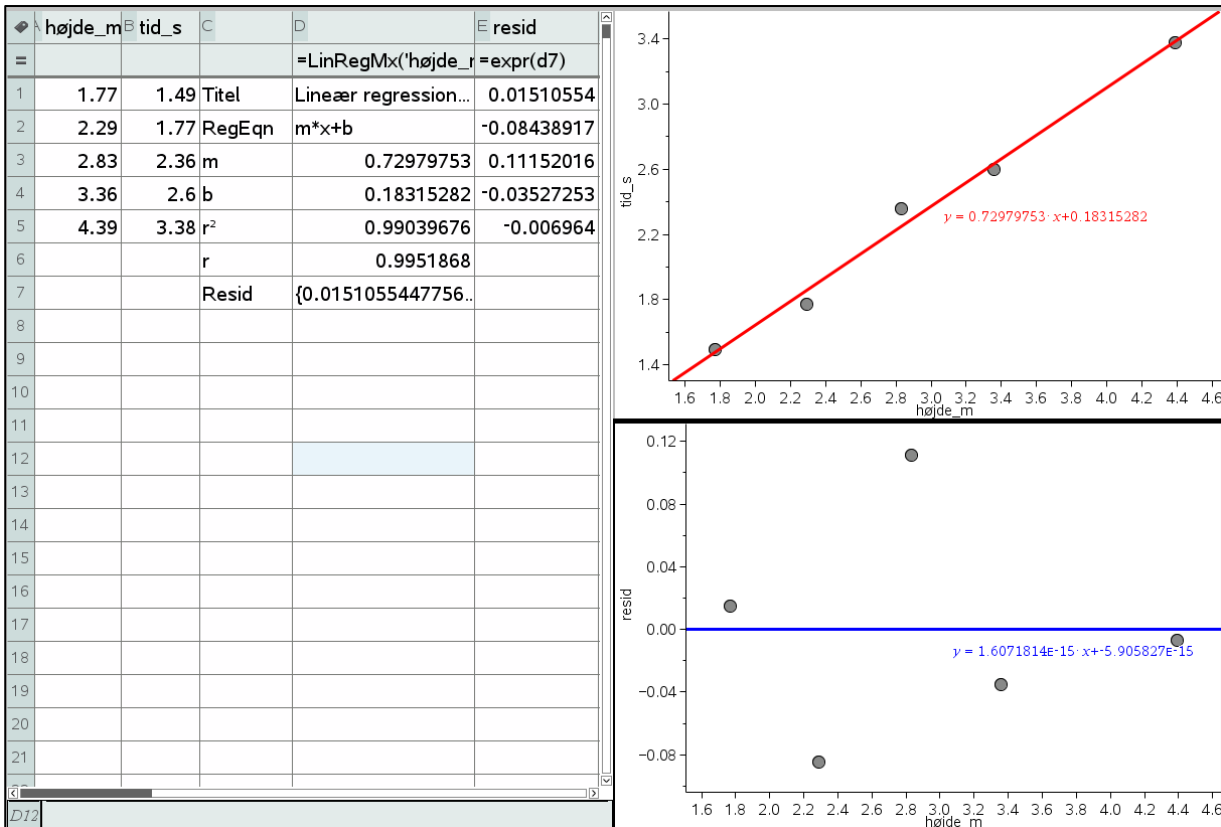


N	måling	O	P	Q
=	=capture(s		=LinRegtIntervals('højde_m','tid	
1	0.09354...	Titel	Lineært Reg t-interval	
2	-0.0688...	RegEqn	a+b*x	
3	-0.0028...	CLower	0.64187859	
4	-0.0804...	CUpper	0.81485687	
5	0.08366...	b	0.72836773	
6	0.00517...	ME	0.08648914	
7	0.06570...	df	48.	
8	0.08242...	s	0.2748197	
9	-0.0166...	SESlope...	0.04301584	
P6	=0.086489135977502			

Vi får da vide at usikkerheden, dvs. Margin of Error eller bare ME, er 0.0865 i fin overensstemmelse med det eksperimentelle resultat. Tilsvarende er usikkerhedsintervallet givet ved  $[C_{Lower}; C_{Upper}] = [0.642; 0.815]$  igen i fin overensstemmelse med det eksperimentelle resultat.

**Afsluttende bemærkninger: Hvad sker der, hvis der kun er én faldtid for hver faldhøjde?**

På forhånd er det klart, at vi ikke uden videre kan tillade os at udføre hypotesetesten på de komprimerede data, hvor vi erstatter de ti faldtider indenfor en given faldhøjde med deres middelværdi. Men hvad nu hvis vi slet ikke har foretaget 10 tilfældige målinger pr faldhøjde, men kun en, dvs. for hver af de fem faldhøjder har vi udført netop et eksperiment, fx det første i den oprindelige serie. Så har vi kun fem målepunkter at lege med. Det gør selvfølgelig kvaliteten af målingerne langt dårligere:



I princippet kan man dog stadigvæk udføre hypotesetest. Godt nok kan man ikke uden videre vurdere antagelsen om at y-værdierne har den samme spredning for de forskellige x-værdier, for der er jo kun en y-værdi for hver x-værdi. Men man kan stadigvæk vurdere residualerne samlet, og bruge deres spredning som skøn over den fælles spredning, under forudsætning af at man altså tror på regressionsmodellen og dermed også tror på at y-værdierne er fordelt med en middelværdi, der varierer lineært med x og en fælles spredning. Men det er klart et tyndere grundlag at arbejde videre med end hvis man har gentagne forsøg for de forskellige x-værdier.

Samtidigt har man stadigvæk mistet 2 frihedsgrader og har derfor kun tre reelle frihedsgrader tilbage til at teste modellen på. Det er et tyndt grundlag. Man vil da også opleve, at der ikke længere er nogen særlig god overensstemmelse mellem de eksperimentelle tests udført som i det foregående på den ene side og de kanoniske t-test på den anden side.

Når man tilrettelægger sit forsøg, er det derfor absolut anbefalelsesværdigt at man gentager målingerne flere gange for de forskellige x-værdier, der indgår!