

Opgave 1, brain weight

(Matthews & Farewell: Using and Understanding Medical Statistics, 2nd. ed.)

For 20 musekuld er der i tabellen nedenfor anført oplysning om

- kuldstørrelsen (fra 3 til 12 mus i et kuld)
- gennemsnitlig kropsvægt i kullet (g)
- gennemsnitlig hjernevægt i kullet (g)

<i>kuldstørrelse</i> litter	<i>kropsvægt</i> body	<i>hjernevægt</i> brain
3	9.447	0.444
3	9.780	0.436
4	9.155	0.417
4	9.613	0.429
5	8.850	0.425
5	9.610	0.434
6	8.298	0.404
6	8.543	0.439
7	7.400	0.409
7	8.335	0.429
8	7.040	0.414
8	7.235	0.409
9	6.600	0.387
9	7.260	0.433
10	6.305	0.410
10	6.655	0.405
11	7.183	0.435
11	6.133	0.407
12	5.450	0.368
12	6.050	0.401

1. Data er indlagt på filen **brain.txt** med tre kolonner, svarende til de tre variable: **litter**, **body** og **brain**. Disse variabelnavne er anført i øverste linie af datafilen. Indlæs data.

Vi ønsker at udtale os om hjernevægtens afhængighed af kuldstørrelsen.

Efterhånden som I gennemgår nedenstående spørgsmål, skal I udfylde tabellen på næste side!



2. Vurder løseligt udfra en tegning, om forudsætningerne for at foretage lineær regressionsanalyse (med hjernevægt som respons og kuld størrelsen som forklarende variabel) ser ud til at være opfyldt. I hvilken retning går en evt. sammenhæng?
3. Bestem parameterestimerer for afskæring og hældning i en lineær regressionsanalyse af hjernevægt, med kuld størrelse som forklarende variabel. Forklar resultatet med ord.
4. Lav en tilsvarende analyse af hjernevægtens afhængighed af kropsvægten.
5. Er der signifikant korrelation mellem kuld størrelse og kropsvægt?
6. Lav en regressionsanalyse af kropsvægt som funktion af kuld størrelse og giv en biologisk fortolkning af resultatet.
7. Foretag en multipel regressionsanalyse med hjernevægt som responsvariabel og såvel kuld størrelse som kropsvægt som forklarende variable.
8. Hvilken biologisk fortolkning har forskellene?



Respons: brain		koefficient til		residual spredning	% forklaret variation
Kovariater i model		litter $\hat{\beta}_1$ (se($\hat{\beta}_1$))	body $\hat{\beta}_2$ (se($\hat{\beta}_2$))	(Root MSE) s	R^2
litter alene		-0.004 (0.001)	—	0.015	0.39
body alene		—	0.34 (0.017)	0.013	0.56
litter og body		0.007 (0.003)	0.024 (0.007)	0.012	0.65

Root MSE bruges til at lave prediktionsgrænser

2 Vi kan ikke sige at R^2 er den andel der forklares af modellen
fordi: vi har ikke et tilfældigt sample (vi har 2 kuld af hver størrelse)

Opgave 2, biomasse

Nedenstående tabel angiver sammenhørende observationer af det kumulerede antal solskinstimer og biomassen af sojabønner over en 8-ugers periode efter spiring. Biomassen er målt som den gennemsnitlige tørvægt i gram af fire uafhængige planter.

<i>nr.</i>	<i>soltimer</i>	<i>biomasse</i>
1	29.7	16.6
2	68.4	49.1
3	120.7	121.7
4	217.2	219.6
5	313.5	375.5
6	419.1	570.8
7	535.9	648.2
8	641.5	755.6

1. Ser det ud til, at der er en lineær sammenhæng mellem de to variable?



Under antagelse om en lineær regressionsmodel med normalfordelte fejl, ønskes følgende spørgsmål besvaret:

2. Giv et estimat for hældningen, med tilhørende 95% sikkerhedsinterval. Undersøg, om hældningen kan antages at være 1.
3. Undersøg om interceptet kan antages at være 0. Hvad bliver hældningsestimatet under denne hypotese? Hvad sker der med spredningsestimatet for hældningen ved overgang fra modellen med intercept til modellen uden intercept (dvs. intercept=0)?
4. Bestem et 95% sikkerhedsinterval for den estimerede biomasseproduktion når det kumulerede antal solskinstimer når op på 200, for modellen med hhv. *uden* intercept. Forklar forskellen.



Log-log:
Hældningsestimatet skal bruges som potens for at angive hvor meget y stiger