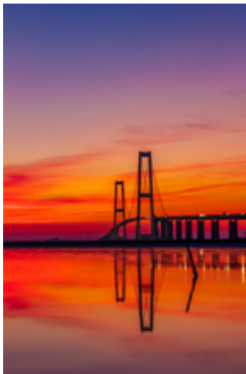


Lineær regression – Analyse af residualerne og hældningskoefficienten

Opgave 1:

Siden årsskiftet 07-08 har det på visse veje været lovligt at køre med ekstra lange lastbiler, de såkaldte modulvogntog.



Colourbox.com

Følgende data viser antallet af disse vogntog, der passerer Storebæltsbroen hver måned i perioden januar 2008-april 2008. Månederne regnes fra januar 2008.

Måned	1	2	...	112
Antal	2386	2530	...	7977

Resten af tabellens data kan findes i bilag, 'Ekstra_lange_lastbiler'.

Man antager, at sammenhængen følger en lineær model:

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

Hvor y er antallet af ekstra lange lastbiler der passerer broen pr. måned, og x er måneden, hvor januar 2008 er måned nummer 1.

- Tegn et punkplot for de givne data.
- Bestem tallene a og b ved at benytte alle tabellens data, og forklar betydningen af tallet a .
- Undersøg, om residualerne med god tilnærmelse kan siges at være normalfordelte.

Opgave 2:



Colourbox.com

Følgende data viser antallet af motorcykler der er kørt over Storebæltsbroen i perioden 1999-2017.

År	1999	1993	...	2017
Antal i tusinde	40,1	44,5	...	140,1

Resten af tabellens data kan findes i bilag, 'Motorcykler_storebaelt'.

I en model antager man, at der er en lineær sammenhængen

$$f(x) = a \cdot x + b$$

Hvor $f(x)$ er antallet af motorcykler over broen, og x er antallet af år efter år 1999.

- Tegn et punktplot for de givne data.
- Bestem tallene a og b ved at benytte alle tabellens data, og forklar betydningen af tallet a .
- Undersøg, om residualerne med god tilnærmelse kan siges at være normalfordelte.

Opgave 3:

Tabellen viser sammenhørende værdier for arbejdsløshedsraten og inflationsraten i USA.

Tabellen viser sammenhørende værdier for arbejdsløshedsraten og inflati

Arbejdsløshedsraten	4	4,7	...	4,9	4,4
Inflationsraten	3,4	2,8	...	1,3	2,1

(Hele datasættet ligger i bilaget philip.xlsx)

I en model antages det, at sammenhængen mellem arbejdsløshedsraten og inflationsraten i USA kan beskrives ved

$$f(x) = a \cdot x + b$$

Hvor $f(x)$ betegner inflationsraten (målt i %), og x betegner arbejdsløshedsraten (målt i %).

- Benyt tabellens data til at bestemme a og b .
- Bestem et 95% konfidensinterval for a , og benyt dette til at afgøre om sammenhængen mellem arbejdsløshedsraten og inflationsraten er aftagende.

Opgave 4:

Et væksthushar hver anden dag i en lang periode plantet små bambusplanter. De har på en bestemt dag målt alle planternes højde. Tabellen viser sammenhængen mellem, hvor mange dage det er siden planten blev plantet og plantens højde i cm.

Dage siden plantning	0	2	...	112
Højde i cm	31,6	71,3	...	322,2

Resten af tabellens data kan findes i bilag, 'Bambus'.

- Lav lineær regression på de givne data.
- Lav et residualplot.
- Undersøg, om residualerne med god tilnærmelse kan siges at være normalfordelt, og afgør om den lineære model kan bruges til beskrive sammenhængen mellem bambusplanter vækst og tiden.
- Bestem et 95%-konfidensinterval for hældningskoefficienten i modellen.

Opgave 5:



I en sportshal har nogle elever målt sammenhørende værdier af hastigheden og højden over gulvet for en basketball, der springer op efter at have ramt gulvet. Tabellen viser elevernes målinger.

Højde (m)	0,36	0,45	0,55	0,63	0,76	0,85	0,91
Hastighed (m/s)	3,25	2,94	2,64	2,34	1,73	1,11	0,49

I en model er der en lineær sammenhæng mellem h og v^2 givet ved

$$v^2 = a \cdot h + b,$$

Hvor v betegner boldens hastighed (målt i meter pr. sekund), og h betegner boldens højde.

- Benyt tabellens data til at bestemme tallene a og b .
- Benyt modellen til at bestemme boldens hastighed, da den var 0,20 meter over gulvet.

Opgave 6:

Det gråsprængte kæmpeegern er en stor gnaver, der lever på Sri Lanka.



iStockphoto.com/Utopia_88

I en stor undersøgelse har man indsamlet data over sammenhængen mellem alder i uger og vægten i gram for en population af kæmpeegern.

Tabellens data kan findes i bilag, 'Kæmpeegern'.

- a) Undersøg om sammenhængen mellem alder og vægt kan beskrives ved en lineær model ved at undersøge om residualerne er normalfordelte.

Opgave 7:

Følgende data viser antallet af nyregistrerede biler for perioden 1992-2017.

År	1992	1993	...	2017
Antal	84518	82084	...	221482

Resten af tabellens data kan findes i bilag 'nyregistrerede_biler'.

I en model antager man, at der er en lineær sammenhæng

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

Hvor y er antallet af nyregistrerede biler, og x er antallet af år efter 1992.

- Tegn et punkplot for de givne data.
- Bestem tallene a og b ved at benytte alle tabellens data, og forklar betydningen af tallet a .
- Benyt residualplottet til at vurdere den lineære models anvendelighed til at beskrive sammenhængen.
- Undersøg, om residualerne med god tilnærmelse kan sides at være normalfordelte.
- Bestem et 95%-konfidensinterval for hældningskoefficienten i modellen.