



BØRNE- OG
UNDERVISNINGSMINISTERIET
STYRELSEN FOR
UNDERVISNING OG KVALITET

Matematik stx

B-niveau

Vejledende enkeltopgaver om
lineær regression

Oktober 2021

Eksempler på visse eksamensopgaver i matematik stx B.

Efter vejledningsjusteringen i oktober 2021 vil datasæt, der skal importeres til eksaminandens matematiske værktøjsprogram ved de skriftlige prøver i matematik B, alene indeholde heltal. Der kan fortsat være tabeller med kommatotal i opgaverne til behandling med eksaminandens matematiske værktøjsprogram, men så er talmaterialet så lille, at det kan tastes ind i hånden. Ligeledes skal en vurdering af modellens anvendelighed fremover ikke nødvendigvis ske ved hjælp af residualspredningen.

I det følgende ses nogle eksempler på nye versioner af opgaver om brug og vurdering af en model.

Opgaverne har fået samme nummer her som opgaverne i tidligere samlinger af vejledende enkeltopgaver, som de bygger på.

Der henvises i øvrigt til vejledningen til læreplanerne for matematik stx B.

Med venlig hilsen

Kim Bertelsen

Fagkonsulent i matematik stx og hf

4.D2.1

Tabellen viser sammenhørende data for længden af en bestemt type fisk (målt i cm) og indhold af kviksølv i fisken (målt i ppm).

Længde(cm)	26	30,4	33	39,5	44	47,3	50,8	54,2
Kviksølv(ppm)	0,28	0,51	0,78	1,2	1,4	1,5	1,8	2,1

I en model antages det, at sammenhængen mellem længden af en fisk og dens indhold af kviksølv kan beskrives ved en lineær funktion

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

hvor x betegner længden af fisken (målt i cm), og $f(x)$ betegner indholdet af kviksølv i fisken (målt i ppm).

- a) Bestem tallene a og b ved lineær regression, og tegn residualplottet.
- b) Hvilken længde har en fisk med et kviksølvindhold på 52 ppm ifølge modellen?

En fisk med længde 59 cm har et kviksølvindhold på 3,6 ppm.

- c) Bestem $f(59)$, og kommentér resultatet.

4.D2.3



Foto: www.colourbox.dk

Old Faithful geyseren i Yellowstone National Park kommer jævnligt i udbrud. Tabellen viser sammenhørende værdier for varigheden af et udbrud og ventetiden til det næste udbrud.

Til opgaven
hører et bilag

Varigheden af et udbrud (sekunder)	96	100	...	304	306
Ventetiden til det næste udbrud (minutter)	52	47	...	76	96

(Hele tabellen med alle 270 datapunkter findes i bilaget "OldFaithful.xlsx")

I en model antages det, at sammenhængen mellem varigheden af et udbrud og ventetiden til det næste udbrud kan beskrives ved en lineær funktion

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

hvor x betegner varigheden af udbruddet (målt i sekunder), og $f(x)$ betegner ventetiden til det næste udbrud (målt i minutter).

- Bestem tallene a og b ved regression.
- Tegn et residualplot, og bestem residualet i punktet (221,81).
- Hvad vil ventetiden til det næste udbrud være ifølge modellen, hvis et udbrud har varet 218 sekunder?

Kilde: Azzalini, A. and Bowman, A. W. (1990). A look at some data on the Old Faithful geyser. *Applied Statistics*, 39, 357–365

4.D2.4



Foto: www.colourbox.dk

Nedenstående tabel viser udviklingen i kornproduktionen i et bestemt land i perioden 1961 – 2016 .

År	1961	1962	...	2019	2020
Kornproduktion	66	75	...	303	305

Til opgaven
hører et bilag.

(Hele tabellen med alle 60 datapunkter findes i bilaget Korndata.xlsx)

I en model kan udviklingen i kornproduktionen beskrives ved

$$f(t) = a \cdot t + b,$$

hvor $f(t)$ betegner kornproduktionen (målt i mio. ton) i landet til tidspunktet t (målt i år efter 1961).

- Benyt lineær regression på alle tabellens data, og tegn residualplottet.
- Benyt residualplottet til at vurdere modellens anvendelighed til at beskrive udviklingen.

4.D2.5



Foto: www.colourbox.dk

Nedenstående tabel viser udviklingen i arbejdsstyrken i et bestemt land i perioden 1990 - 2020.

År	1990	1995	2000	2005	2010	2015	2020
Arbejdsstyrke (mio.)	13,72	14,99	16,26	17,41	18,25	19,49	20,45

I en model kan udviklingen i arbejdsstyrken beskrives ved

$$f(t) = a \cdot t + b,$$

hvor $f(t)$ betegner arbejdsstyrken (målt i mio. personer) i landet til tidspunktet t (målt i år efter 1990).

- Benyt lineær regression på tabellens data, og tegn residualplottet.
- Benyt residualplottet til at vurdere modellens anvendelighed til at beskrive udviklingen.

4.D2.6



Foto: www.colourbox.dk

I et studie har man undersøgt sammenhængen mellem alder og kolesteroltal. Tabellen herunder viser de målte data fra studiet.

Alder (år)	29	34	...	71	71
Kolesteroltal (mg/dL)	204	182	...	302	265

(Hele tabellen med alle 100 datapunkter findes i bilaget "Kolesterol.xlsx")

Til opgaven
hører et bilag

I en model antages det, at sammenhængen mellem alder og kolesteroltal kan beskrives ved en lineær funktion

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

hvor $f(x)$ betegner kolesteroltallet (målt i mg/dL) for en person, der er x år gammel.

- Benyt lineær regression til at redegøre for, at $a = 1,811$ og $b = 150,5$.
- Tegn punktplottet og regressionslinjen i det samme koordinatsystem.

Det højeste kolesteroltal i undersøgelsen var 417.

- Hvor mange procent var den pågældende persons kolesteroltal større end det forventede ifølge modellen?

4.D2.7

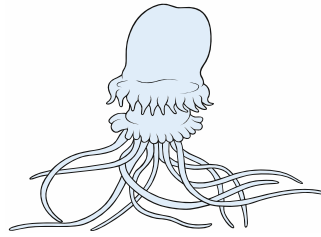


Foto: www.colourbox.dk

Tabellen viser sammenhørende værdier fra et bestemt havområde for en vandmands bredde og længden på dens længste tentakel (føletråd).

Bredde på vandmand (målt i mm)	12	15	14	...	19	20
Længde af længste tentakel (målt i mm)	14	16	17	...	22	22

(Hele tabellen med alle 24 datapunkter findes i bilaget *VandMænd.xlsx*)

Til opgaven
hører et bilag

I en model kan sammenhængen mellem vandmandens bredde og længden af den længste tentakel beskrives ved

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

hvor $f(x)$ betegner længden af en vandmands længste tentakel (målt i mm), og x betegner bredden af en vandmand (målt i mm).

- Benyt tabellens data til at bestemme konstanterne a og b .
- Tegn et residualplot, og benyt dette til at vurdere modellens anvendelighed.
- Bestem residualet hørende til $x = 12$, og forklar betydningen af tallet.