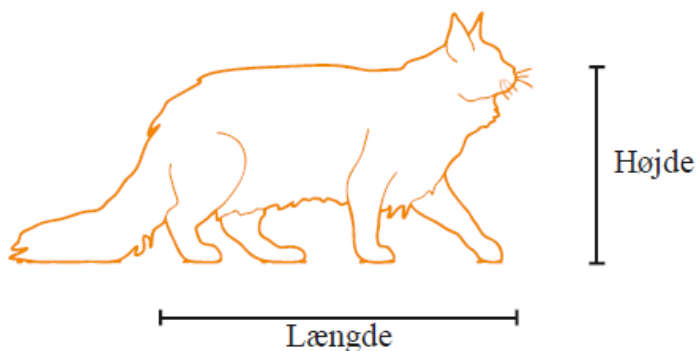


Tidligere eksamensopgaver i statistik

KBJ, december 2023

Is Ma

Opgave 1



Tabellen viser sammenhørende værdier af nogle kattes længde og højde.

Længde (mm)	480	484	...	832	826
Højde (mm)	250	331	...	465	438

(Hele tabellen med alle 95 datapunkter findes i bilaget "Kattedimensioner.xlsx")

I en model kan sammenhængen beskrives ved

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

hvor $f(x)$ betegner højden (målt i mm), og x betegner længden (målt i mm).

a) Bestem tallene a og b ved lineær regression.

Opgave 2

Tabellen viser udviklingen i antallet af trafikulykker, hvor føreren havde drukket spiritus, i perioden 1997–2018.

Antal år efter 1997	0	1	...	20	21
Antal spiritus-trafikulykker	656	583	...	151	156

(Hele tabellen med alle 22 datapunkter findes i bilaget "Trafikulykker.xlsx")

I en model antages det, at udviklingen i antallet af spiritus-trafikulykker kan beskrives ved funktionen

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

hvor $f(x)$ betegner antallet af spiritus-trafikulykker til tidspunktet x (målt i antal år efter 1997).

a) Benyt tabellens data til at bestemme a og b .

Det oplyses, at udviklingen i antallet af alle typer trafikulykker i samme periode kan beskrives ved modellen

$$g(x) = -64,6x + 1735,$$

hvor $g(x)$ betegner antallet af alle typer trafikulykker til tidspunktet x (målt i antal år efter 1997).

b) Bestem tallet $\frac{f(21)}{g(21)}$, og forklar, hvad dette tal betyder.

Opgave 3

Figuren viser et udklip af en avisartikel.

Ny undersøgelse: 67% af danskerne lider af kroniske sygdomme

Undersøgelsen er baseret på svar fra 200 tilfældigt udvalgte danskere.

- a) Opstil på baggrund af undersøgelsen et 95%-konfidensinterval for andelen af alle danskere, der lider af kroniske sygdomme.

En tidligere undersøgelse viste, at 58% af alle danskere lider af kroniske sygdomme.

- b) Afgør på baggrund af 95%-konfidensintervallet, om andelen af alle danskere, der lider af kroniske sygdomme, har ændret sig signifikant.

Opgave 4



Grafik: freepik.com

På en ø har man udtaget en stikprøve på 900 mariehøns. I stikprøven var 3% af mariehønsene syvprikkede.

- a) Bestem et 95% konfidensinterval for øens andel af syvprikkede mariehøns.

I en meget stor undersøgelse på øen har man tidligere fundet, at 2% af øens mariehøns var syvprikkede.

- b) Afgør ud fra stikprøven, om øens andel af syvprikkede mariehøns har ændret sig signifikant.

Opgave 5

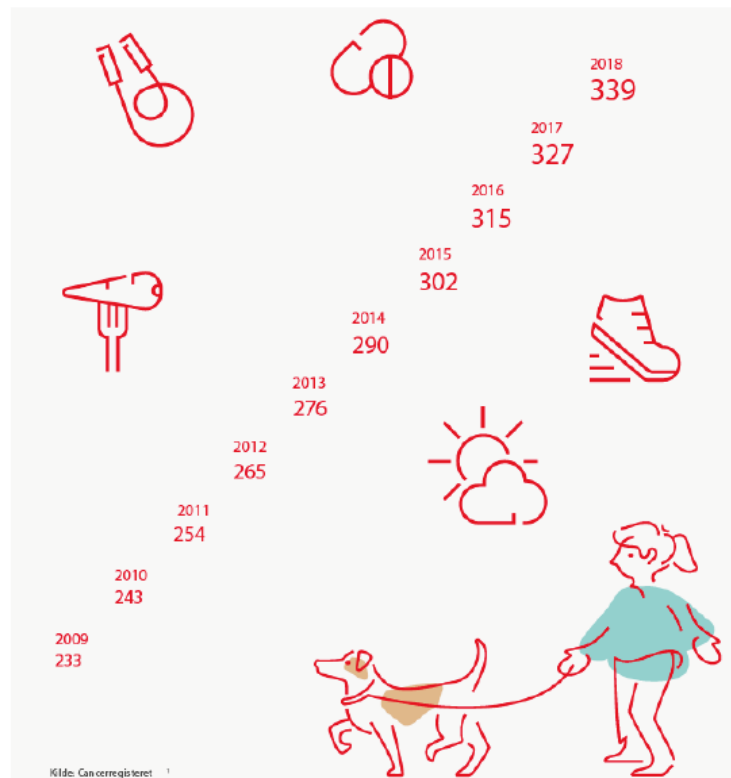
Figuren viser for perioden 2009-2018 antallet af personer (målt i tusinde), der lever efter at have fået en kræftdiagnose.

Udviklingen kan beskrives ved modellen

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

hvor t er antal år efter 2009, og $f(t)$ er antallet af personer (målt i tusinde), der lever efter at have fået en kræftdiagnose.

- Bestem tallene a og b ved regression.
Benyt figurens data.
- Tegn residualplottet, og vurder modellens anvendelighed.
- Bestem residualet til tidspunktet $t = 4$, og forklar, hvad dette tal betyder.



Billedkilde: "Kræft i Danmark", 2020

Opgave 6

I en meningsmåling fra 2018 blandt 1008 tilfældigt udvalgte danske vælgere svarede 60 %, at de ville stemme nej til euroen.

- Bestem ud fra denne meningsmåling et 95 % konfidensinterval for andelen af danske vælgere, der ville stemme nej til euroen.

I 2000 var der en dansk folkeafstemning om euroen. Ved denne folkeafstemning stemte 53,2 % af vælgerne nej til euroen.

- Afgør ved at bruge 95 % konfidensintervallet, om meningsmålingen viste en signifikant ændring i de danske vælgeres holdning til euroen.

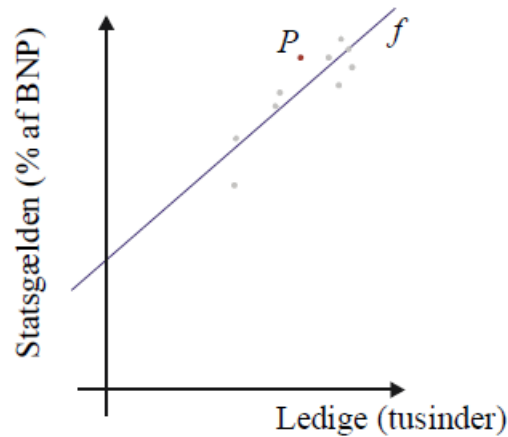


Billedkilde: euskolen

Opgave 7 (KUN FORMELSAMLING)

Over en årrække har man indsamlet data for antallet af ledige i Danmark og lavet en model for sammenhængen mellem ledigheden og statsgældens størrelse i forhold til bruttonationalproduktet (BNP).

Figuren nedenfor viser sammenhørende datapunkter for ledigheden og statsgældens størrelse i forhold til BNP sammen med en lineær regressionsmodel f baseret på disse data.



Det oplyses, at forskriften for funktionen f er givet ved

$$f(x) = 0,2x + 17,$$

hvor $f(x)$ betegner statsgældens størrelse (målt i % af BNP), og x betegner antal ledige (målt i tusinde).

Det på figuren fremhævede datapunkt P har koordinaterne $(130, 51)$.

- a) Bestem residualen for punktet P , og forklar betydningen af dette tal.

Opgave 8



Kilde: mst.dk

Nogle elever har efter en storm samlet en stikprøve på 800 sten på en bestemt strand. I elevernes stikprøve var 4% af stenene af typen *rhombeporfyr*.

- a) Bestem et 95% konfidensinterval for andelen af *rhombeporfyr* blandt strandens sten.

Fra en meget stor undersøgelse før stormen har geologer bestemt andelen af *rhombeporfyr* blandt strandens sten til at udgøre 3,5%.

- b) Afgør ud fra stikprøven, om der er statistisk belæg for, at andelen af *rhombeporfyr* blandt strandens sten har ændret sig efter stormen.

Opgave 9

Tabellen viser sammenhørende værdier for rumfanget af en mands hoved og vægten af hans hjerne.

Rumfang (målt i cm^3)	4512	3738	...	3458	3890
Vægt (målt i gram)	1530	1297	...	1225	1310

(Hele tabellen med alle 57 datapunkter findes i bilaget *Hovedrumfang.xlsx*)

I en model kan sammenhængen beskrives ved

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

hvor $f(x)$ betegner vægt (målt i gram), og x betegner rumfang (målt i cm^3).

- a) Benyt regression på tabellens data til at bestemme tallene a og b .