Eksamen på Økonomistudiet sommeren 2020
Sandsynlighedsteori og Statistik
2. årsprøve
13. august, 2020
(3-timers prøve med hjælpemidler)
UDKAST TIL RETTEVEJLEDNING

Syg under eksamen:

- Bliver du syg under selve eksamen på Peter Bangs Vej, skal du
- kontakte en eksamensvagt for at få hjælp til registreringen i systemet som syg og til at aflevere blankt.
- forlade eksamen
- kontakte din læge og indsende lægeerklæring til Det Samfundsvidenskabelige Fakultet senest 5 dage efter eksamensdagen.

Pas på, du ikke begår eksamenssnyd!

Det er eksamenssnyd, hvis du under prøven

- Bruger hjælpemidler, der ikke er tilladt
- Kommunikerer med andre eller på anden måde modtager hjælp fra andre
- Kopierer andres tekster uden at sætte citationstegn eller kildehenvise, så det ser ud som om det er din egen tekst
- Bruger andres ideer eller tanker uden at kildehenvise, så det ser ud som om det er din egen ide eller dine egne tanker
- Eller hvis du på anden måde overtræder de regler, der gælder for prøven

Opgaven består af tre delopgaver, som alle skal besvares. De tre opgaver kan regnes uafhængigt af hinanden.

Opgave 1 og 2 indgår tilsammen med samme vægt som opgave 3.

Opgave 1

Den international organisation OHW har vurderet, at sandsynligheden for at en tilfældig udvalgt person fra land A har COVID-19 er 0,004. Tilsvarende er det vurderet, at sandsynligheden for at en tilfældig udvalgt person fra land B har COVID-19 er 0,002. OHW udvælger tilfældigt 1.000 personer fra hvert land.

OHW antager at antallet af personer, der har COVID-19 blandt de 1.000 udtrukne kan beskrives med en binomialfordeling.

.

1) Giv en vurdering af den model OHW anvender.

Brug nu OHW's model.

.

Forudsætning for Bin er: hvert forsøg ja/nej, samme p og uafhængighed. Der er et problem med uafhængighed, man smittes jo. Kan også diskutere om p er konstant, (super-spredere)

.

2) Hvad er sandsynligheden for, at der er præcis 5 personer, der bliver testet positiv for COVID-19 i land A?

.

$P(X = 5|X \sim bin(1.000; 0,004) = 0,156607$

.

3) Det er indrapporteret, at et land har 5 tilfælde af COVID-19, hvad er sandsynligheden for at det indrapporterede land er A?

.

$P(X = 5) = P(X = 5|A)P(A) + P(X = 5|B)P(B) = 0,156607 * \frac{1}{2} + 0,036017 * \frac{1}{2}$

.

$P(A|X=5) = \frac{P(A \cap X=5)}{P(X=5)} = \frac{\frac{P(X=5 \cap A)P(A)}{P(A)}}{P(X=5)} = \frac{P(X=5|A)P(A)}{P(X=5)} = \frac{0,156607 * \frac{1}{2}}{0,156607 * \frac{1}{2} + 0,036017 * \frac{1}{2}} = 0,81$

Opgave 2

1. Lad (Y, X) være bivariat Normafordelte

   med $E(X) = E(Y) = 1$ og $V(X) = V(Y) = 1$ samt $COV(X, Y) = 0,2$.

   1. Udregn $P(X < 0)$

.

$P(X < 0) = P(\frac{X-1}{1} < \frac{0-1}{1}) = P(X - 1 < -1) = \Phi(-1) = 0,15$

2. Angiv den betingede fordeling af Y givet X=0

.

Vil være normalfordelt med

$E(Y|X = 0) = \mu_Y + \frac{\sigma_{YX}}{\sigma_X^2}(0 - \mu_X) = 1 + \frac{0,2}{1}(0 - 1) = 0,8$

.

$V(Y|X = 0) = \sigma_Y^2 + \frac{\sigma_{YX}^2}{\sigma_X^2} = 1 - \frac{0,2^2}{1} = 0,96$

.

Lad $Z_1 = 2Y + X$ og $Z_2 = Y + 3X$. Dette kan også skrives som

.

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} Y \\ X \end{pmatrix}$$

.

3. Hvad er middelværdierne af Z_1 og Z_2?

.

$E(Z_1) = E(2Y + X) = 2E(Y) + E(X) = 3$

.

$E(Z_2) = E(Y + 3X) = E(Y) + 3E(X) = 1 + 3 = 4$

.

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

4. Hvad er varianserne for Z_1 og Z_2?

.

$V(Z_1) = V(2Y + X) = V(2Y) + V(X) + 2cov(2Y, X) =$

.

$4V(Y) + V(X) + 4cov(Y, X) = 4 * 1 + 1 + 4 * 0,2 = 5,8$

.

$V(Z_2) = V(Y + 3X) = V(Y) + V(3X) + 2cov(Y, 3X) =$

.

$V(Y) + 9V(X) + 6cov(Y, X) = 1 + 9 + 6 * 0,2 = 11,2$

5. Hvad er fordelingen af $(Z_1,\ Z_2)$?

.

bivarit normalfordeling. med

$cov(Z_1, Z_2) = cov(2Y + X, Y + 3X) =$

.

$cov(2Y, Y) + cov(2Y, 3X) + cov(X, Y) + cov(X, 3X) =$

.

$2V(Y) + 6cov(Y, X) + cov(X, Y) + 3V(X) =$

.

2+6*0,2+0,2+3=6,4

.

kan også gøres med matrice beregning:

.

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0,2 \\ 0,2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 2,2 & 1,4 \\ 1,6 & 3,2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} =$$

.

$$\begin{pmatrix} 5,8 & 6,4 \\ 6,4 & 11,2 \end{pmatrix}$$

6. Hvad er korrelationskoefficienten mellem Z_1 og Z_2?

.

$\rho = \frac{cov(Z_1, Z_2)}{\sqrt{V(Z_1)V(Z_2)}} = \frac{6,4}{\sqrt{5,8*11,2}} = 0,79$

Opgave 3

En undersøgelse blandt 19 forsøgspersoner har målt sammenhængen mellem kondition og daglig motion.

I det følgende er X udtryk for personens daglige motion, mens Y er udtryk for personens kondi.

Nedenstående tabel viser et uddrag af målingerne.

person nr.	kondi (y)	motion (x)
1	11,25	6,90
2	8,40	5,65
3	9,80	6,53
19	11,20	6,65

Der er opstillet følgende model

$Y_i \sim N(\beta X_i; 1,16^2)$ i=1,....,19 uafhængige

så $E(Y_i) = \beta X_i$ og $V(Y) = 1,6^2 = 2,56$

Vi betragter X'er som givne størrelser.

.

tæthedsfunktionen er dermed

.

$\varphi(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{1,6} \exp(-\frac{\frac{1}{2}(y-\beta x_i)^2}{1,6^2})$

.

Det kan oplyses at:

$\sum_{i=1}^{19} y_i x_i = 1203,2$ og $\sum_{i=1}^{19} x_i^2 = 743,0$

1) Opskriv likelihoodfunktionen $L(\beta)$.

.

$L(\beta) = \Pi_{i=1}^{19} \frac{1}{\sqrt{2\pi}} \frac{1}{1,16} \exp(-\frac{1}{2} \frac{(y_i-\beta x_i)^2}{1,6^2}) =$

.

$(\frac{1}{\sqrt{2\pi}})^{19} (\frac{1}{1,16})^{19} \exp(-\frac{1}{2} \sum_{i=1}^{19} \frac{(y_i-\beta x_i)^2}{1,6^2})$

2) angiv log-likelihoodfunktionen og score-funktionen

.

$l(\beta) = ln[L(\beta)] = 19(\frac{1}{\sqrt{2\pi}}) + 19(\frac{1}{1{,}16}) - \frac{1}{2}\sum_{i=1}^{19}\frac{(y_i-\beta x_i)^2}{1{,}6^2}$

.

$l'(\beta) = \frac{1}{1{,}6^2}\sum_{i=1}^{19}(y_i - \beta x_i)x_i = \frac{1}{1{,}6^2}[\sum_{i=1}^{19} y_i x_i - \beta\sum_{i=1}^{19} x_i^2]$

.

3) vis at $\widehat{\beta} = \frac{\sum_{i=1}^{19} y_i x_i}{\sum_{i=1}^{19} x_i^2} = \frac{1.203{,}6}{743{,}0} = 1,62$

.

Brug: $\frac{1}{1{,}6^2}[\sum_{i=1}^{19} y_i x_i - \beta\sum_{i=1}^{19} x_i^2] = 0$

4) angiv hesse-matricen.

.

$l'(\beta) = \frac{1}{1{,}6^2}[\sum_{i=1}^{19} y_i x_i - \beta\sum_{i=1}^{19} x_i^2]$

.

$l''(\beta) = -\frac{1}{1{,}6^2}\sum_{i=1}^{19} x_i^2$

.

Det oplyses endvidere at:

$\sum_{i=1}^{19} y_i^2 = 1.994,4$ og at $\sum_{i=1}^{19}(y_i - \widehat{\beta}x_i)^2 = 46,2$

.

5) test hypotesen at $\beta = 0$ mod $\beta \neq 0$ med et likelihood-ratio test.

.

$l(\beta) = 19(\frac{1}{\sqrt{2\pi}}) + 19(\frac{1}{1{,}16}) - \frac{1}{2}\sum_{i=1}^{19}\frac{(y_i-\beta x_i)^2}{1{,}6^2}$

.

$l(\widehat{\beta}) = 19(\frac{1}{\sqrt{2\pi}}) + 19(\frac{1}{1{,}16}) - \frac{1}{2}\frac{1}{1{,}6^2} * 46,2$

.

$l(0) = 19(\frac{1}{\sqrt{2\pi}}) + 19(\frac{1}{1{,}16}) - \frac{1}{2}\sum_{i=1}^{19}\frac{(y_i)^2}{1{,}6^2} =$

.

$19(\frac{1}{\sqrt{2\pi}}) + 19(\frac{1}{1{,}16}) - \frac{1}{2}\frac{1}{1{,}6^2} * 1.994,4$

.

-2 $* [l(0) - l(\widehat{\beta})] = \frac{1994{,}4-46{,}2}{1{,}6^2} = 761,0$

.

som er χ^2 med DF=1

6) Kommenter resultatet af testen.

.

Klar sammenhæng mellem motion og kondi

7) test hypotesen at $\beta = 1,0$ mod $\beta \neq 1,0$ med et wald test.

.

V($\widehat{\beta}$) kan beregnes som $\frac{1{,}6^2}{\sum_{i=1}^{19} x_i^2} = \frac{1{,}6^2}{743{,}0} = (0,0587)^2$

.

test=$\frac{1{,}62-1{,}0}{0{,}0587} = 10$

8) Angiv et 95% appr. konfidensinterval for β.

.

1,96*0,0587=0,11   [1,50 - 1,74]

9) kommenter dette interval og sammenlign med svarene i 5) og 6)

Intervallet indeholder ikke 0

intervallet indeholder ikke 1