Eksamen på Økonomistudiet sommeren 2018
Sandsynlighedsteori og Statistik

2. årsprøve

11. juni, 2018

UDKAST TIL RETTEVEJLEDNING

den 11/6

(3-timers prøve med hjælpemidler)

Dette eksamenssæt består af 4 sider (forsiden inklusiv).

OBS: Bliver du syg under selve eksamen på Peter Bangsvej, ska du kontakte et tilsyn, blive registreret som syg hos denne. Derefeter afleverer du en blank besvarelse i systemet og forlader eksamen. Når du kommer hjem, skal du kontakte din læge og indsende lægeerklæringen til det Samfundsvidenskabelige Fakultet senest en uge efter eksamensdagen.

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 simultane fordeling af X og Y er givet ved

	$Y = 0$	$Y = 1$
$X = -1$	$0,1$	$0,3$
$X = 0$	$0,3$	$0,1$
$X = 1$	$0,1$	$0,1$

1. Beregn middelværdi og varians af den marginale fordeling for Y. Dvs. beregn $E(Y)$ og $V(Y)$.

   $E(Y) = 0 * 0,5 + 1 * 0,5 = 0,5$

   $E(Y^2) = 0^2 * 0,5 + 1^2 * 0,5 = 0,5$

   $V(Y) = E(Y^2) - E(Y)^2 = 0,5 - 0,5^2 = 0,25$

   til brug for senere

   $E(X) = -1 * 0,4 + 0 * 0,4 + 1 * 0,2 = -0,2$

2. Angiv fordelingen af Z hvor $Z = X * Y$. Udregn også $E(Z)$.

   $P(Z = -1) = 0,3$; $P(Z = 0) = 0,6$; $P(Z = 1) = 0,1$

   $E(Z) = -1 * 0,3 + 0 * 0,6 + 1 * 0,1 = -0,2$

3. Beregn kovariansen mellem X og Y, dvs. udregn $COV(X, Y)$.

   $cov(X, Y) = E(XY) - E(X)E(Y) = -0,2 - (-0,2) * 0,5 = -0,1$

4. Er X og Y uafhængige? Begrund svaret.

   uafhængighed vil medføre at kovarians bliver 0. Altså afhængighed.

5. Udregn den betingede middelværdi af Y givet $X = 1$.

   dvs. udregn $E(Y|X = 1)$

   $P(Y = 0|X = 1) = \frac{P(Y=0,X=1)}{P(X=1)} = \frac{0,1}{0,2} = 0,5$

   $P(Y = 1|X = 1) = \frac{P(Y=1,X=1)}{P(X=1)} = \frac{0,1}{0,2} = 0,5$

   $E(Y|X = 1) = 0 * 0,5 + 1 * 0,5 = 0,5$

Opgave 2

I de berømte PISA undersøgelser kan matematikscoren og læsescoren beskrives med to normalfordelinger.

Lad X være matematikscoren og Y læsescoren. Der gælder da at $X \sim N(510, 100^2)$ og $Y \sim N(490, 100^2)$.

Endvidere gælder at korrelationskoefficient er på 0,5.

$$\rho = \frac{cov(X,Y)}{\sqrt{V(X) * V(Y)}} = 0,5.$$

1. Udregn sandsynligheden for at en tilfældig udvalgt elev får en matematikscore på mindst 600.

   $P(X > 600) = 1 - P(X < 600) = 1 - P(\frac{X-510}{100} < \frac{600-510}{100}) =$

   $1 - \Phi(\frac{90}{100}) = 1 - \Phi(0,9) = \Phi(-0,9) = 0,18$

2. Bestem 95% fraktilen for X. Dvs. bestem et k så P(X<=k)=0,95.

   $P(X <= k) = P(\frac{X-510}{100} <= \frac{k-510}{100}) = \Phi(\frac{k-510}{100}) = 0,95$

   $\frac{k-510}{100} = \Phi^{-1}(0,95) = 1,645$

   $k = 1,645 * 100 + 510 = 674,5$

   Undervisningsministeriet udregner nu et PISA-gennemsnit som kaldes Z,

   dvs. $Z = \frac{X+Y}{2}$

3. Angiv fordelingen for Z, og udregn middelværdi og varians

   $E(Z) = E(\frac{X+Y}{2}) = \frac{E(X+Y)}{2} = \frac{E(X)+E(Y)}{2} = \frac{510+490}{2} = 500$

   $V(Z) = V(\frac{X+Y}{2}) = \frac{V(X+Y)}{4} = \frac{V(X)+V(Y)+2cov(X,Y)}{4}$

   $\frac{100^2+100^2+2*0,5*\sqrt{100^2 100^2}}{4} = \frac{100^2+100^2+100^2}{4} = 0,75 * 100^2$

   Z er normalfordelt da den er en liniær kombination af normalfordelinger

4. Udregn P(Z>700)

   $P(Z > 700) = 1 - P(Z < 700) = 1 - P(\frac{Z-500}{(\sqrt{075})100} < \frac{700-500}{(\sqrt{075})100}) =$

   $1 - \Phi(\frac{200}{(\sqrt{075})100}) = 1 - \Phi(2,3) = \Phi(-2,3) = 0,01$

Opgave 3

I nedenstående tabel er vist resultaterne af to opinionsundersøgelser i U.K.

Undersøgelserne er foretaget før og efter den meget berømte Brexit afstemning i juni 2016.

I undersøgelserne er personerne blevet spurgt om, de er er for eller imod Europarlamentet (E.P.)

	før Brexit	efter Brexit
for E.P.	664	731
imod E.P.	1.434	1.131
i alt	2.098	1.862

kilde: European social survey (ESS)

Den statstiske model er:
P(at udtage en person der er for E.P. før Brexit)=p_1
P(at udtage en person der er imod E.P. før Brexit)=$1 - p_1$
P(at udtage en person der er for E.P. efter Brexit)=p_2
P(at udtage en person der er imod E.P. efter Brexit)=$1 - p_2$

1. Argumenter for at du også kan bruge to uafhængige binomialfordelinger til at beskrive ovenstående data

2. Opskriv likelihood-funktionen $L(p_1, p_2)$.Angiv scorefunktionerne samt $\widehat{p}_1$ og $\widehat{p}_2$

   Hvis man tager udgangspunkt i to uafhængige BIN fordelinger så fås:

   $L(p_1, p_2) = \binom{n_1}{x_1} p_1^{x_1} (1 - p_1)^{n_1 - x_1} \binom{n_2}{x_2} p_2^{x_2} (1 - p_2)^{n_2 - x_2}$

   her er $n_1 = 2.098$; $x_1 = 664$; $n_1 - x_1 = 1.434$

   her er $n_2 = 1.862$; $x_2 = 731$; $n_2 - x_2 = 1.131$

   $l(p_1, p_2) = \ln[L(p_1, p_2)]$

   $l(p_1, p_2) = \ln[\binom{n_1}{x_1}] + x_1 \ln(p_1) + (n_1 - x_1) \ln(1 - p_1) +$

   $\ln[\binom{n_2}{x_2}] + x_2 \ln(p_2) + (n_2 - x_2) \ln(1 - p_2)$

   $l'_{p_1} = \frac{x_1}{p_1} - \frac{(n_1 - x_1)}{(1 - p_1)}$

$l'_{p_2} = \frac{x_2}{p_2} - \frac{(n_2 - x_2)}{(1-p_2)}$

$\widehat{p}_1 = \frac{x_1}{n_1} = \frac{664}{2.098} = 0,32$

$\widehat{p}_2 = \frac{x2}{n_2} = \frac{731}{1.862} = 0,39$

3. Angiv Hesse-matricen

   $l''_{p_1} = -\frac{x_1}{p_1^2} - \frac{(n_1 - x_1)}{(1-p_1)^2}$

   $l''_{p_2} = -\frac{x_2}{p_2^2} - \frac{(n_2 - x_2)}{(1-p_2)^2}$

   H=$\begin{bmatrix} -\frac{x_1}{p_1^2} - \frac{(n_1-x_1)}{(1-p_1)^2} & 0 \\ 0 & -\frac{x_2}{p_2^2} - \frac{(n_2-x_2)}{(1-p_2)^2} \end{bmatrix}$

4. Angiv et 95% konfidensinterval for p_1.

   $E[-\frac{x_1}{p_1^2} - \frac{(n_1-x_1)}{(1-p_1)^2}] = -\frac{E(x_1)}{p_1^2} - \frac{(n_1-E(x_1))}{(1-p_1)^2} =$

   $-\frac{n_1 p_1}{p_1^2} - \frac{(n_1 - n_1 p_1)}{(1-p_1)^2} = -\frac{n_1 p_1}{p_1^2} - \frac{n_1(1-p_1)}{(1-p_1)^2} =$

   $-\frac{n_1}{p_1} - \frac{n_1}{(1-p_1)} = \frac{-n_1(1-p_1)-n_1 p_1}{p_1(1-p_1)} = \frac{-n_1}{p_1(1-p_1)}$

   så $V(\widehat{p}_1)$ er approksimativt $=\frac{p_1(1-p_1)}{n_1}$

   $\widehat{p}_1 \pm 1,96\sqrt{\frac{p_1(1-p_1)}{n_1}}$

   giver [0,30 til 0,34]

5. Test $H_0 : p_1 = 0,5$ mod $H_A : p_1 \neq 0,5$. Brug et Wald test

   W=$\frac{0{,}32-{,}5}{\sqrt{\frac{{,}32*(1-0{,}32)}{2098}}} = -17,7$

   Z $=\frac{0{,}32-{,}5}{\sqrt{\frac{{,}5*(1-0{,}5)}{2098}}} = -16,5$

   Man ønsker nu at undersøge om der er sket en ændring efter Brexit.

6. Test $H_0 : p_1 = p_2$ mod $H_A : p_1 \neq p_2$. Brug et likelihood Ratio test

   sæt $p = p_1 = p_2$

   $L(p) = \binom{n_1}{x_1} p^{x_1}(1-p)^{n_1-x_1} \binom{n_2}{x_2} p^{x_2}(1-p)^{n_2-x_2} =$

   $\binom{n_1}{x_1}\binom{n_2}{x_2} p^{x_1+x_2}(1-p)^{n_1+n_2-(x_1+x_2)}$

   da fås at $\widehat{p} = \frac{x_1+x_2}{n_1+n_2}$

   $L(\widehat{p}_1, \widehat{p}_2) =$

$x_1 \ln(\widehat{p}_1) + (n_1 - x_1)\ln(1 - \widehat{p}_1) +$

$x_2 \ln(\widehat{p}_2) + (n_2 - x_2)\ln(1 - \widehat{p}_2) = -2556,9$

$l(\widehat{p}) = \widehat{p}^{x_1+x_2}(1 - \widehat{p})^{n_1+n_2-(x_1+x_2)} = -2569,4$

Dermed bliver teststørrelsen=- 2*(-2569,4-2556,9)=25

DF=1