Eksamen på Økonomistudiet sommeren 2018
Sandsynlighedsteori og Statistik
2. årsprøve
16. august, 2018
Udkast til rettevejledning
(3-timers prøve med hjælpemidler)

Dette eksamenssæt består af 5 sider (forsiden inklusiv).

OBS: Bliver du syg under selve eksamen på Peter Bangsvej, skal du kontakte et tilsyn, blive registreret som syg hos denne. Derefter 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

X og Y er uafhængige stokastiske variable fordelt på 0 og 1.
Med følgende sandsynligheder:
P(X=0)=0,2 og P(X=1)=0,8
P(Y=0)=0,5 og P(Y=1)=0,5

1. Angiv den simultane fordeling for (X,Y).

	Y=0	Y=1	
X=0	0,1	0,1	0,2
X=1	0,4	0,4	0,8
	0,5	0,5	1,0

Sæt $Z_1 = 3X + Y$ og $Z_2 = X - Y$

2. Udregn $E(Z_1)$ og $V(Z_1)$

$E(X) = 0 * 0,2 + 1 * 0,8 = 0,8$

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

$E(X^2) = 0^2 * 0,2 + 1^2 * 0,8 = 0,8$

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

$V(X) = E(X^2) - [E(X)]^2 = 0,8 - 0,64 = 0,16$

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

$E(3X + Y) = 3E(X) + E(Y) = 3 * 0,8 + 0,5 = 2,9$

$V(3X + Y) = 9V(X) + V(Y) = 9 * 0,16 + 0,25 = 1,69$

3. Beregn kovariansen mellem Z_1 og Z_2, dvs. udregn $COV(Z_1, Z_2)$.

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

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

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

$3V(X) + 0 + 0 - V(Y) = 3 * 0,16 - 0,25 = 0,23 \neq 0$

4. Er Z_1 og Z_2 uafhængige? begrund svaret.

5. Udregn den betingede middelværdi af X givet $Z_1 = 0$. Dvs. $E(X|Z_1 = 0)$

Opgave 2

Firmaet BLOMST producerer urtepotter. Vægten (i gram) af disse urtepotter kan beskrives med en normalfordeling.

Lad X angive vægten af en urtepotte. Der gælder da, at $X \sim N(40, 5^2)$.

1. Udregn sandsynligheden for at en tilfældig urtepotte vejer mindre end 35 gram.

   $P(X < 35) = P(\frac{X-40}{5} < \frac{35-40}{5}) = \Phi(-1) = 0,16$

2. Bestem 5% fraktilen for X. Dvs. bestem et k så P(X<=k)=0,05.

   $P(X < k) = P(\frac{X-40}{5} < \frac{k-40}{5}) = \Phi(\frac{k-40}{5}) = 0,05$

   $\frac{k-40}{5} = \Phi^{-1}(0,05) = -1,645$

   $k = -1,645 * 5 + 40 = 31,2$

   Firmaet pakker urtepotterne i kasser med 10 urtepotter i hver kasse. Du kan antage, at kassen uden urtepotter vejer præcist 200 gram.

3. Angiv fordelingen for vægten af en kasse med 10 urtepotter. Præciser de antagelser, du har gjort dig.

   $Y = 200 + \sum_{i=1}^{10} X_i$

   $E(Y) = 200 + 10 * 40 = 600$

   $V(Y) = \sum_{i=1}^{10} V(X_i) = 10 * 25 = 250$

   $Y \sim N(600; 15,8^2)$. Husk uafhængighed

Opgave 3

Firmaet BLOMST producerer også mindre kroge, som kan bruges til at hænge urtepotterne op med. Produktionen foregår to steder A og B.

BLOMST ved fra mange års erfaringer at vægten af krogene kan beskrives med en normalfordeling, hvor variansen er 1.

Man ønsker at vægten af krogene skal være 2,5 gram. Dette er der dog usikkerhed om ligesom det også er usikkert om de to produktionssteder rammer den samme vægt.

Der udføres en række forsøg hvor der på produktionssted A produceres 10 kroge og på produktionssted B produceres 20 kroge.

Resultaterne ses i nedenstående tabel:

Tabel for produktion af kroge

	antal	Sum	gennemsnit $(\overline{X})$	$\sum_i (X_i - \overline{X})^2$
A	10	24,3	2,4	3,2
B	20	56,1	2,8	11,9
A og B	30	80,4	2,7	16,1

kilde: BLOMST

Den statistiske model BLOMST bruger er:

$X_i \sim N(\mu_1, 1^2)$ i=1,2,...,10 $Y_j \sim N(\mu_2, 1^2)$ j=1,2,...,20

Alle stokastiske variable er uafhængige.

Her angiver $\mu_1$og μ_2 middelværdierne.

1. Opskriv likelihood-funktionen $L(\mu_1, \mu_2)$ og log-likelihood. Angiv score-funktionerne.

   $L(\mu_1, \mu_2) =$

   $\prod_{1=1}^{10} \frac{1}{\sqrt{2\pi}} \exp(\frac{-(x_i-\mu_1)^2}{2}) \prod_{1=1}^{20} \frac{1}{\sqrt{2\pi}} \exp(\frac{-(y_i-\mu_2)^2}{2}) =$

   $(\frac{1}{\sqrt{2\pi}})^{10} \exp(-\frac{1}{2}\sum_{i=1}^{10}(x_i-\mu_1)^2)(\frac{1}{\sqrt{2\pi}})^{20} \exp(-\frac{1}{2}\sum_{j=1}^{20}(y_j-\mu_2)^2)$

   .

   $l(\mu_1, \mu_2) =$

   $10\ln(\frac{1}{\sqrt{2\pi}}) - \frac{1}{2}\sum_{i=1}^{10}(x_i-\mu_1)^2 + 20\ln(\frac{1}{\sqrt{2\pi}}) - \frac{1}{2}\sum_{j=1}^{20}(y_j-\mu_2)^2$

$\mathrm{l}'_{\mu_1} = \sum_{i=1}^{10}(x_i - \mu_1) = \sum_{i=1}^{10} x_i - 10\mu_1$

$\mathrm{l}'_{\mu_2} = \sum_{i=1}^{20}(y_j - \mu_2) = \sum_{j=1}^{20} y_j - 20\mu_2$

2. Angiv MLE for μ_1 og μ_2.

   $\sum_{i=1}^{10}(x_i - \mu_1) = 0$ giver $\widehat{\mu}_1 = \frac{1}{10}\sum_{i=1}^{10} x_i = 2,4$

   $\sum_{j=1}^{20}(y_i - \mu_2) = 0$ giver $\widehat{\mu}_2 = \frac{1}{20}\sum_{j=1}^{20} x_i = 2,8$

3. Sæt $\widehat{\mu}_1$ og $\widehat{\mu}_2$ ind i log-likelihood funktionen.

   $l(\widehat{\mu}_1, \widehat{\mu}_2) =$

   $30\ln(\frac{1}{\sqrt{2\pi}}) - \frac{1}{2}\sum_{i=1}^{10}(x_i - \widehat{\mu}_1)^2 - \frac{1}{2}\sum_{i=1}^{10}(x_i - \widehat{\mu}_1)^2$

   $30\ln(\frac{1}{\sqrt{2\pi}}) - \frac{1}{2}(3,2 + 11,9)$

4. Angiv Hesse-matricen.

   $$\begin{bmatrix} -10 & 0 \\ 0 & -20 \end{bmatrix}$$

5. Angiv den aymptotiske fordeling for $\widehat{\mu}_1$.

   $\widehat{\mu}_1 \sim N(\mu_1, \frac{1}{10})$

6. Angiv et 95% konfidensinterval for μ_1.

   $\widehat{\mu}_1 \pm 1,96\sqrt{\frac{1}{10}}$

   Antag nu at der gælder $\mu_1 = \mu_2 = \mu$.

7. Angiv estimatet for μ

   $l(\mu) =$

   $30\ln(\frac{1}{\sqrt{2\pi}}) - \frac{1}{2}\sum_{i=1}^{10}(x_i - \mu)^2 - \frac{1}{2}\sum_{j=1}^{20}(y_i - \mu)^2$

   $l'(\mu) = \sum_{i=1}^{10}(x_i - \mu) + \sum_{j=1}^{20}(y_i - \mu) =$

   $\sum_{i=1}^{10} x_i + \sum_{j=1}^{20} y_i - 30\mu$

   $\widehat{\mu} = \frac{\sum_{i=1}^{10} x_i + \sum_{j=1}^{20} y_i}{30} = \frac{24,3+56,1}{30} = 2,7$

8. Test $\mathrm{H}_0 : \mu_1 = \mu_2$ mod $\mathrm{H}_A : \mu_1 \neq \mu_2$ med et LR test.

   $l(\widehat{\mu}) =$

1. $30\ln(\frac{1}{\sqrt{2\pi}}) - \frac{1}{2}\sum_{i=1}^{10}(x_i - \widehat{\mu})^2 - \frac{1}{2}\sum_{j=1}^{20}(y_i - \widehat{\mu})^2 =$

   $30\ln(\frac{1}{\sqrt{2\pi}}) - \frac{1}{2} * 16,1$

   Dermed bliver teststørrelsen:

   $= -2[l(\widehat{\mu}) - l(\widehat{\mu}_1, \widehat{\mu}_2)] = 16,1 - 3,2 - 11,9 = 1,0$

   som er χ^2 med df=2-1=1

   Da teststørrelsen er mindre end $\chi^2_{.95}(1) = 3,84$ kan vi ikke afvise H_0.(Så vi opretholder H_0)

   Note: det klassiske U-test bliver

   $U = \frac{\overline{X}-\overline{Y}}{\sqrt{\frac{1}{10}+\frac{1}{20}}} = \frac{2,4-2,8}{\sqrt{\frac{1}{10}+\frac{1}{20}}} = -1,03$ og $U^2 = 1,01$