Eksamen på Økonomistudiet sommeren 2016
Sandsynlighedsteori og Statistik
2. årsprøve
14. juni, 2016
(3-timers prøve med hjælpemidler)

Dette eksamenssæt består af 5 sider (forsiden inklusiv).

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.

Den simultane fordeling af X og Y er givet ved

	$X=0$	$X=1$	$X=2$
$Y=0$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$
$Y=1$	$\frac{1}{4}$	0	$\frac{1}{8}$
$Y=2$	0	$\frac{1}{8}$	0

Opgave 1

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

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

3. Udregn den betingede middelværdi og den betingede varians af Y givet $X = 1$. dvs. udregn $E(Y|X = 1)$ og $V(Y|X = 1)$.

4. Er X og Y uafhængige? Begrund svaret.

Opgave 2

I den sidste PISA undersøgelse foretaget i 2012 deltog 7.500 elever fra Danmark og 4.700 elever fra Sverige. De danske elevers score i matematik kan beskrives med en normalfordeling, der har middelværdi 500 og en spredning (standard afvigelse) på 80. Dvs. $N(500, 80^2)$. De svenske elevers score i matematik kan beskrives med en normalfordeling, der har middelværdi 480 og en spredning på 80. Dvs. $N(480, 80^2)$.Det antages, at alle elevers score er uafhængige af hinanden.

1. Beregn sandsynligheden for at en tilfældig svensk elev får en score, der er mindst 500.

   I alt er der i de to lande 12.200 elever. Der er nu udvalgt en elev, hvis score er mindst 500.

2. Hvad er sandsynligheden for, at eleven kommer fra Sverige?

   I Danmark udvælges de elever, der har en score på mindst 600. Antallet af elever der har en score på mindst 600 kaldes Z.

3. Hvilken fordeling kan beskrive Z? Beregn $E(Z)$.

   Eleverne er testet både i matematik og læsning. Læsescoren i Danmark kan beskrives med en normalfordeling med middelværdi 490 og en spredning på 80. Dvs. $N(490, 80^2)$.

   Korrelationskoefficienten mellem læse- og matematikscoren er på 0,9. dvs

$$\rho = \frac{cov(X,Y)}{\sqrt{V(X) * V(Y)}} = 0,9.$$

   Her angiver X læsescoren og Y matematikscoren.

Undervisningsministeriet ønsker at anvende en samlet score som kaldes S. Den samlede score beregnes udfra formlen:

$S = \frac{1}{3}X + \frac{2}{3}Y$. Man ønsker altså at lægge mere vægt på læsning end matematik i dette samlede mål

1. Angiv fordelingen for S.

Opgave 3

På to store indfaldsveje (A og B) til København har Vejdirektoratet i henholdsvis 35 og 33 uger registreret antallet af trafikuheld på en hverdag.

Registreringerne er angivet i nedenstående tabel.

antal trafikuheld	A	B
0	10	13
1	13	10
2	8	7
3	3	3
4	1	0
5+	0	0
i alt	35	33

Der opstilles følgende model:

$X_1, ...X_{35} \sim Poisson(\lambda)$ og $Y_1, ...Y_{33} \sim Poisson(\mu)$ alle stokastiske variable er uafhængige af hinanden.

Ud fra tabellen fås

$\sum_{i=1}^{35} X_i = 42$ og $\overline{X} = \frac{1}{35} \sum_{i=1}^{35} X_i = 1,2$ samt

$s^2 = \frac{1}{35-1} \sum_{i=1}^{35} (X_i - 1,2)^2 = 1,11$

$\sum_{i=1}^{33} Y_i = 33$ og $\overline{Y} = \frac{1}{33} \sum_{i=1}^{33} Y_i = 1,0$ samt

$\mathrm{s}^2 = \frac{1}{33-1} \sum_{i=1}^{33} (Y_i - 1,2)^2 = 1,00$

1. Argumenter for at det er en rimelig model, der opstilles. Angiv fordelingerne for $\sum_{i=1}^{35} X_i$ og $\sum_{i=1}^{33} Y_i$

2. Opskriv likelihood funktionen $L(\lambda, \mu)$ samt log-likelihood funktionen $log[L(\lambda, \mu)]$ for det samlede datamateriale $X_1, ...X_{35}$ og $Y_1, ...Y_{33}$

   dvs. for alle 68 observationer. Angiv scorefunktionerne og Hesse-matricen. Udregn MLE (maksimumlikelihood-estimaterne) for λ og μ.

3. Angiv et 95% (approximativt) konfidensinterval for λ.

   Der ønskes nu undersøgt om $\lambda = \mu$. Kald den fælles intensitet for γ.

4. Vis at MLE for γ bliver $\frac{42+33}{35+33} = 1,103$. Test $H_0 : \lambda = \mu$ mod $H_A : \lambda \neq \mu$ ved brug af et likelihood ratio test.

   Antag nu at alle 68 observationer er Poissonfordelt med parameteren γ.

5. Test $H_0 : \gamma = 1$ mod $H_A : \gamma \neq 1$.