

Skriftlig eksamen i Statistik

4 timers skriftlig prøve

Dette opgavesæt består af 4 delopgaver, der indgår i bedømmelsen af den samlede opgavebesvarelse med følgende omtrentlige vægte:

<i>Opgave 1</i>	<i>35%</i>
<i>Opgave 2</i>	<i>20%</i>
<i>Opgave 3</i>	<i>25%</i>
<i>Opgave 4</i>	<i>20%</i>
<i>I alt</i>	<i>100%</i>

Til prøven er alle hjælpemidler tilladte – også elektroniske.
Der må ikke foretages opkobling til eksterne kilder.

Opgavebesvarelsen skal afleveres i henhold til skolens eksamensreglement
jf. bekendtgørelse nr. 1519 af 16. december 2013

Torsdag den 4. juni 2015
Kl. 14.00 –18.00

Baggrundsinformation

Fantomvibrations – syndromet (fremover blot fantom – SMS) er et udbredt fænomen. Det dækker over, at man griber i lommen, fordi man fornemmer, at mobilen vibrerer fra en SMS. Det viser sig så, at ingen har sendt hverken en SMS, mail eller Facebook-besked – mobilen har ligget helt i ro. En gruppe studerende har på nettet fundet en større undersøgelse af fænomenet og besluttet at undersøge udbredelsen i deres eget område.

Deres undersøgelse omfattede 45 tilfældigt udvalgte personer, der fordelte sig på aldersgrupper som vist nedenfor.

Under 25 år	25 år - 50 år	Over 50 år
16	17	12

I undersøgelsen blev respondenterne bedt om inden for den næste måned at registrere fænomenet fantom – SMS (= Antal fantom – SMS).

Opgave 1

Svarene om antal fantom – SMS for de 16 respondenter i aldersgruppen Under 25 år fremgår af tabel 1.1 nedenfor.

Tabel 1.1 Antal fantom – SMS for 16 tilfældigt udvalgte personer under 25 år.

2	5	0	5
0	6	1	4
3	0	3	6
1	3	4	7

- 1.1 Undersøg om antal fantom – SMS for personer under 25 år følger en normalfordeling.
- 1.2 Test på 5% niveauet om variansen på antal fantom – SMS for personer under 25 år kan antages at være 4.

I den undersøgelse, der blev fundet på nettet, fremgår det, at det gennemsnitlige antal fantom – SMS for personer under 25 år er 4.

- 1.3 Test på 5% niveauet om det gennemsnitlige antal fantom – SMS for personer under 25 år i de studerendes undersøgelse er under 4.
Gør rede for signifikansniveauets betydning for din konklusion.

De studerende havde en forventning om, at fænomenet fantom – SMS var mest udbredt blandt unge under 25 år. De ønskede derfor at sammenligne aldersgruppen Under 25 år med aldersgruppen 25 år – 50 år. Svarene fra de 17 respondenter i gruppen 25 år – 50 år fremgår af tabel 1.2 nedenfor.

Tabel 1.2 Antal fantom – SMS for 17 tilfældigt udvalgte personer i gruppen 25 år – 50 år.

3	2	0	0	1
3	1	0	4	0
0	3	5	4	2
3	0			

- 1.4 Test på 5% niveauet om antal fantom – SMS i gruppen Under 25 år og i gruppen 25 år – 50 år i den videre analyse kan antages at have ens varianser.
Præcisér forudsætningerne for denne test.
- 1.5 Undersøg ved hjælp af en test om det gennemsnitlige antal fantom – SMS i gruppen Under 25 år er større end det gennemsnitlige antal fantom – SMS i gruppen 25 år – 50 år.
Hvad ville konklusionen blive, hvis vi havde testet 2-sidet?

Antag i det følgende at antal fantom – SMS i en måned i gruppen Under 25 år følger en normalfordeling med en middelværdi på $\mu = 4$ og en varians på $\sigma^2 = 4$.

- 1.6 Bestem sandsynligheden for at en person i gruppen Under 25 år registrerer over 5 fantom – SMS i en måned.
- 1.7 Bestem sandsynligheden for at det gennemsnitlige antal fantom – SMS i en måned for 16 personer i gruppen Under 25 år er over 5.

Opgave 2

Eventuelle sammenhænge mellem antal fantom – SMS og alder blev yderligere undersøgt ved hjælp af en variansanalyse.

Inden selve variansanalysen blev det undersøgt, om udvælgelsen kunne anses for at være repræsentativ. Tabel 2.1 viser dels antal tilfældigt udvalgte og dels aldersgruppens andel i området.

Tabel 2.1 Antal tilfældigt udvalgte fordelt på alder sammenholdt med aldersfordelingen i området.

Gruppe	Antal udvalgte	Andel i området
Under 25 år	16	0,30
25 år – 50 år	17	0,35
Over 50 år	12	0,35

2.1 Undersøg ved hjælp af en test om stikprøven på 45 tilfældigt udvalgte er repræsentativ for området med hensyn til alder.

Resultatet af variansanalysen gengives i tabel 2.2 nedenfor.

Data til variansanalysen findes i bilag 1.

Tabel 2.2 Variansanalyse for antal fantom – SMS fordelt på alder.

Anava: Enkelt faktor

RESUME

Grupper	Antal	Sum	Gennemsnit	Varsians
Under 25 år	16	50	3,1250	5,3167
25 år – 50 år	17	31	1,8235	2,9044
Over 50 år	12	18	1,5000	1,9091

ANAVA

Variationskilde	SK	fg	MK	F	P-værdi	F krit
Mellem grupper	21,98	2	10,99	3,1352	0,0538	3,2199
Inden for grupper	147,22	42	3,51			
I alt	169,20	44				

2.2 Undersøg ved hjælp af tabel 2.2, om det gennemsnitlige antal fantom – SMS kan antages at være ens for de 3 aldersgrupper.
Kommentér dit resultat.

2.3 Vurdér om forudsætningerne for testen er opfyldt.

Opgave 3

For at finde mulige forklaringer på forskellen i antal fantom – SMS pr. måned for aldersgruppen Under 25 år, blev der lavet en regressionsanalyse med antal fantom – SMS pr. måned (nedenfor blot Antal) som den forklarede variabel.

I undersøgelsen indgik følgende mulige forklarende variable:

SMS	- viser hvor mange SMS respondenter i gennemsnit sender pr. dag.
Mobiltid	- viser hvor mange minutter respondenter i gennemsnit bruger mobiltelefonen pr. dag.
Alder	- viser respondenterens alder.
Køn	- viser respondenterens køn, idet Kvinde = 1 og Mand = 0.

De indsamlede data findes i bilag 2.

Inden regressionerne blev udarbejdet, blev korrelationerne mellem de anvendte variable undersøgt.

Bilag 3 viser en korrelationsmatrix for de anvendte variable.

3.1 Vurdér – ud fra bilag 3 – sammenhængene mellem de anvendte variable.

I bilag 4 vises følgende regressioner:

Model 1: Antal fantom – SMS forklaret ved SMS, Mobiltid, Alder og Køn.

Model 2: Antal fantom – SMS forklaret ved SMS, Mobiltid og Alder.

Model 3: Antal fantom – SMS forklaret ved SMS og Mobiltid.

Model 4: Antal fantom – SMS forklaret ved SMS.

3.2 Opstil ved hjælp af bilag 4 den bedste model til forklaring af Antal fantom – SMS.

De studerende besluttede at anvende den model, hvor Antal fantom – SMS alene forklares ved SMS (Model 4).

3.3 Opstil denne model, og forklar modellens koefficienter. Vurdér endvidere modellens forudsætninger ud fra plottene i Model 4.

3.4 Bestem, ud fra Model 4, et 95% forudsigelsesinterval for Antal fantom – SMS for en person, der i gennemsnit sender 30 SMS pr. dag.

.

Opgave 4

Den undersøgelse gruppen af studerende har fundet på nettet omfatter 992 respondenter fra hele landet. Tabel 4.1 nedenfor viser sammenhængen fra denne undersøgelse mellem aldersgruppe og hvorvidt respondenterne har oplevet fænomenet fantom – SMS eller ej.

Tabel 4.1 Sammenhæng mellem aldersgruppe og oplevelsen af fænomenet fantom – SMS.

Alder	Fantom - SMS		
	Ja	Nej	I alt
Under 25 år	212	92	304
25 år - 50 år	158	181	339
Over 50 år	148	201	349
I alt	518	474	992

- 4.1 Test på 5% niveauet om der er en sammenhæng mellem alder og oplevelsen af fantom – SMS. Kommentér en evt. sammenhæng.
- 4.2 Undersøg ved hjælp af et 95% konfidensinterval, om det kan antages, at mindst 75% af alle unge Under 25 år har oplevet fantom – SMS (svaret Ja).
Gør rede for, om konklusionen ville ændres, hvis udgangspunktet var et 99% konfidensinterval.
- 4.3 Test på 5% niveauet om andelen i gruppen 25 år – 50 år, der har oplevet fantom – SMS, kan antages at være større end andelen i gruppen Over 50 år, der har oplevet fantom – SMS.

Bilag 1 **Antal fantom – SMS for aldersgrupperne Under 25 år, 25 år – 50 år og Over 50 år.**

Under 25 år	25 år – 50 år	Over 50 år
2	3	0
5	2	2
0	0	0
5	0	4
0	1	2
6	3	0
1	1	1
4	0	2
3	4	3
0	0	1
3	0	0
6	3	3
1	5	
3	4	
4	2	
7	3	
	0	

Bilag 2 Data til regressioner

Antal	SMS	Mobiltid	Alder	Køn
2	15	110	17	0
5	40	80	22	1
0	10	90	18	1
5	42	120	18	1
0	15	60	16	0
6	45	160	20	1
1	15	120	16	0
4	30	140	17	1
3	20	100	16	0
0	5	120	16	0
3	25	200	17	0
6	40	120	23	1
1	20	100	17	0
3	25	170	21	1
4	35	70	23	0
7	50	100	24	1

Antal	- viser antal fantom – SMS pr. måned.
SMS	- viser hvor mange SMS respondenter i gennemsnit sender pr. dag.
Mobiltid	- viser hvor mange minutter respondenter i gennemsnit bruger mobiltelefonen pr. dag.
Alder	- viser respondenterens alder.
Køn	- viser respondenterens køn, idet Kvinde = 1 og Mand = 0.

Bilag 3 Korrelationsmatrix

	<i>Antal</i>	<i>SMS</i>	<i>Mobiltid</i>	<i>Alder</i>	<i>Køn</i>
Antal	1				
SMS	0,9629	1			
Mobiltid	0,2010	0,1014	1		
Alder	0,7466	0,7607	-0,1008	1	
Køn	0,6159	0,6236	0,1743	0,5603	1

Bilag 4 Regressioner**Model 1: Antal fantom – SMS forklaret ved SMS, Mobiltid, Alder og Køn.**

RESUMEOUTPUT

<i>Regressionsstatistik</i>	
Multipel R	0,9699
R-kvadreret	0,9407
Justeret R-kvadreret	0,9191
Standardfejl	0,6556
Observationer	16

ANOVA					
	<i>fg</i>	<i>SK</i>	<i>MK</i>	<i>F</i>	<i>Signifikans F</i>
Regression	4	75,02	18,76	43,6305	0,0000
Residual	11	4,73	0,43		
I alt	15	79,75			

	<i>Koefficienter</i>	<i>Standardfejl</i>	<i>t-stat</i>	<i>P-værdi</i>	<i>Nedre 95%</i>	<i>Øvre 95%</i>
Skæring	-3,0912	1,6982	-1,8203	0,0960	-6,8289	0,6464
SMS	0,1503	0,0209	7,1835	0,0000	0,1042	0,1963
Mobiltid	0,0075	0,0049	1,5441	0,1508	-0,0032	0,0183
Alder	0,0693	0,0965	0,7178	0,4878	-0,1432	0,2818
Køn	-0,0400	0,4343	-0,0920	0,9283	-0,9958	0,9159

Model 2: Antal fantom – SMS forklaret ved SMS, Mobiltid og Alder.

RESUMEOUTPUT

<i>Regressionsstatistik</i>	
Multipel R	0,9699
R-kvadreret	0,9407
Justeret R-kvadreret	0,9258
Standardfejl	0,6280
Observationer	16

ANOVA

	<i>fg</i>	<i>SK</i>	<i>MK</i>	<i>F</i>	<i>Signifikans F</i>
Regression	3	75,02	25,01	63,4107	0,0000
Residual	12	4,73	0,39		
I alt	15	79,75			

	<i>Koefficienter</i>	<i>Standardfejl</i>	<i>t-stat</i>	<i>P-værdi</i>	<i>Nedre 95%</i>	<i>Øvre 95%</i>
Skæring	-3,0478	1,5624	-1,9507	0,0748	-6,4521	0,3565
SMS	0,1497	0,0190	7,8685	0,0000	0,1082	0,1911
Mobiltid	0,0074	0,0046	1,6257	0,1300	-0,0025	0,0174
Alder	0,0673	0,0902	0,7465	0,4697	-0,1292	0,2639

Model 3: Antal fantom – SMS forklaret ved SMS og Mobiltid.

RESUMEOUTPUT

<i>Regressionsstatistik</i>	
Multipel R	0,9685
R-kvadreret	0,9379
Justeret R-kvadreret	0,9284
Standardfejl	0,6172
Observationer	16

ANOVA					
	<i>fg</i>	<i>SK</i>	<i>MK</i>	<i>F</i>	<i>Signifikans F</i>
Regression	2	74,80	37,40	98,1805	0,0000
Residual	13	4,95	0,38		
I alt	15	79,75			

	<i>Koefficienter</i>	<i>Standardfejl</i>	<i>t-stat</i>	<i>P-værdi</i>	<i>Nedre 95%</i>	<i>Øvre 95%</i>
Skæring	-1,9699	0,5870	-3,3562	0,0052	-3,2380	-0,7019
SMS	0,1607	0,0117	13,7077	0,0000	0,1354	0,1860
Mobiltid	0,0065	0,0043	1,5030	0,1567	-0,0028	0,0158

Model 4: Antal fantom – SMS forklaret ved SMS.

RESUMEOUTPUT

<i>Regressionsstatistik</i>	
Multipel R	0,9629
R-kvadreret	0,9271
Justeret R-kvadreret	0,9219
Standardfejl	0,6443
Observationer	16

ANOVA					
	<i>fg</i>	<i>SK</i>	<i>MK</i>	<i>F</i>	<i>Signifikans F</i>
Regression	1	73,94	73,94	178,0860	0,0000
Residual	14	5,81	0,42		
I alt	15	79,75			

	<i>Koefficienter</i>	<i>Standardfejl</i>	<i>t-stat</i>	<i>P-værdi</i>	<i>Nedre 95%</i>	<i>Øvre 95%</i>
Skæring	-1,2625	0,3661	-3,4483	0,0039	-2,0477	-0,4773
SMS	0,1625	0,0122	13,3449	0,0000	0,1364	0,1886

