



Antal blad /
Number of sheets

15 ✓

TENTAMEN / EXAMINATION

Anvisningar: Skriv din anonymitetskod på varje blad.
Endast en uppgift får lösas på varje blad.
Var vänlig skriv tydligt!

Instructions: Write your anonymous code on each sheet.
Answer only one question on each sheet.
Please write clearly!

Vänligen texta anonymitetskoden i textboxen enligt exempel nedan!
Please write the Anonymous Code clearly in the textbox like example below!

Bokstäver/Letters:

A-B-C-D-E-F-G-H-I-J-K-L-M-N-O
P-Q-R-S-T-U-V-W-X-Y-Z-Å-Ä-Ö

Siffror/Numbers:

Ø-1-2-3-4-5-6-7-8-9

Exempel:

A	B	C	1	7	Ø	-	Ø	1	7
---	---	---	---	---	---	---	---	---	---

STGA Statistik

Kurskod + Kurs / Course Code + Course:

Delkurs / Part course:

Anonymitetskod / Anonymous code =
Kurskod + kodnr / course code + code number
STGAØ1-122 ✓

Tentamensdatum /
Examination date:
22/1-2016

Behandlade uppgifter / Solved problems

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	X	X	X	X	X									
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

Ifylles av lärare / To be completed by the examiner

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
7	10	9	9.5	9.5	10									
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

Poäng / Marks gained: 55

Betyg / Grade: VG

Exam. lärare / Kursansvarig signatur / Signature of the examiner
A.W

Max poäng / Total marks gained: _____

Namnförtydligande / Clarification of the signature

För Gk poäng / Marks gained to be passed: _____



Fråga 1

a) Svar: Vattenförbrukningen är responsvariabeln

$$b) r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

x	y	xy	x ²	y ²
2	59	118	4	3481
5	148	740	25	21904
1	52	52	1	2704
7	202	1414	49	40804
2	75	150	4	5625
9	140	1260	81	19600
3	104	312	9	10816
4	131	524	16	17161
Summa:	33	911	189	122140

n = 8

$$r = \frac{8 \times 911 - 33 \times 33}{\sqrt{[8 \times 189 - 33^2] \times [8 \times 122140 - 911^2]}} \approx -0,46725$$

$$r^2 \approx 0,21832$$

Svar: Korrelationskoefficienten är ca -0,467 och kan tolkas så att eftersom r inte är mellan -1 till -0,7 och 0,7 så är det inte optimalt att anpassa en rätt linje. Determinationskoefficienten (r²) säger att endast 21,8% av vattenförbrukningen kan förklaras av antalet personer i hushållet.

Häftområde

Skriv ej i detta område
 Leave this area blank



Ange anonymitetskod / Write your anonymity code
 (Vid icke anonym tentamen ange kurskod + namn + personnummer)
 (For non-anonymous exams write the course code + name + civic registration number)

ST6A01-122

Löpande sidnr
 Consecutive no:

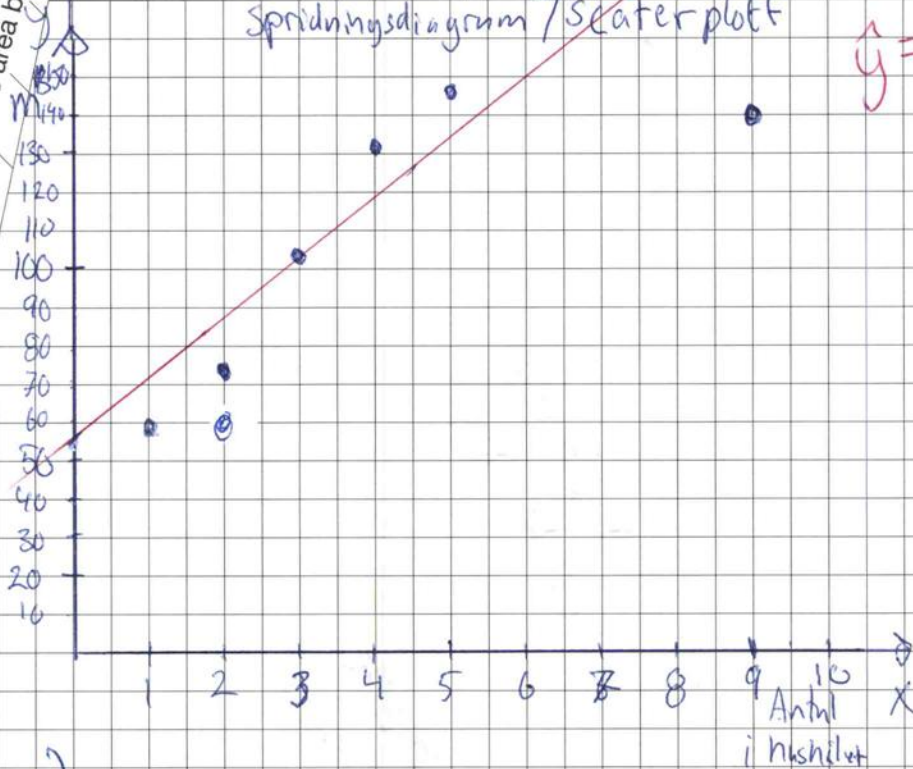
2

Häftområde

Skriv ej i detta område
 Leave this area blank

Spridningsdiagram / Scatter plot

$$\hat{y} = 50,518 + 15,359x$$



Uppgift nr /
 Question no:
 1

Poäng / Points
 awarded:

Lärarens
 anteckning
 Examiner's remarks:

c)

Tolkning av regressionslinjen:

Linjan säger att vid 1 person i hushållet ~~ligger~~ ^{skattas} förbrukningen ligger runt 70m^3 . För varje ytterliggare person i hushållet så ökar förbrukningen med ca $15,4\text{m}^3$ per år.

- $50,518$ berättar vart linjen skär y-axeln (alltså vid 0 personer i hushållet förbrukar man ca $50,5\text{m}^3$ enligt linjen)

- $15,359$ berättar hur mycket en ^{ytterliggare} person i hushållet ~~ökar~~ ökar förbrukningen.

Regressionslinjen är en skattad linje och kan ses som den bäst anpassade räta linjen.

R



Ange anonymitetskod / Write your anonymity code
 (Vid icke anonym tentamen ange kurskod + namn + personnummer)
 (For non-anonymous exams write the course code + name + civic registration number)

STGA01-122

Löpande sidnr
 Consecutive no:

3

Häftområde

Skriv ej i detta område
 Leave this area blank

d) Residualerna $\hat{y} - y$

$y - \hat{y}$

$$\hat{y} = 50,578 + 15,359x$$

Uppgift nr /
 Question no:

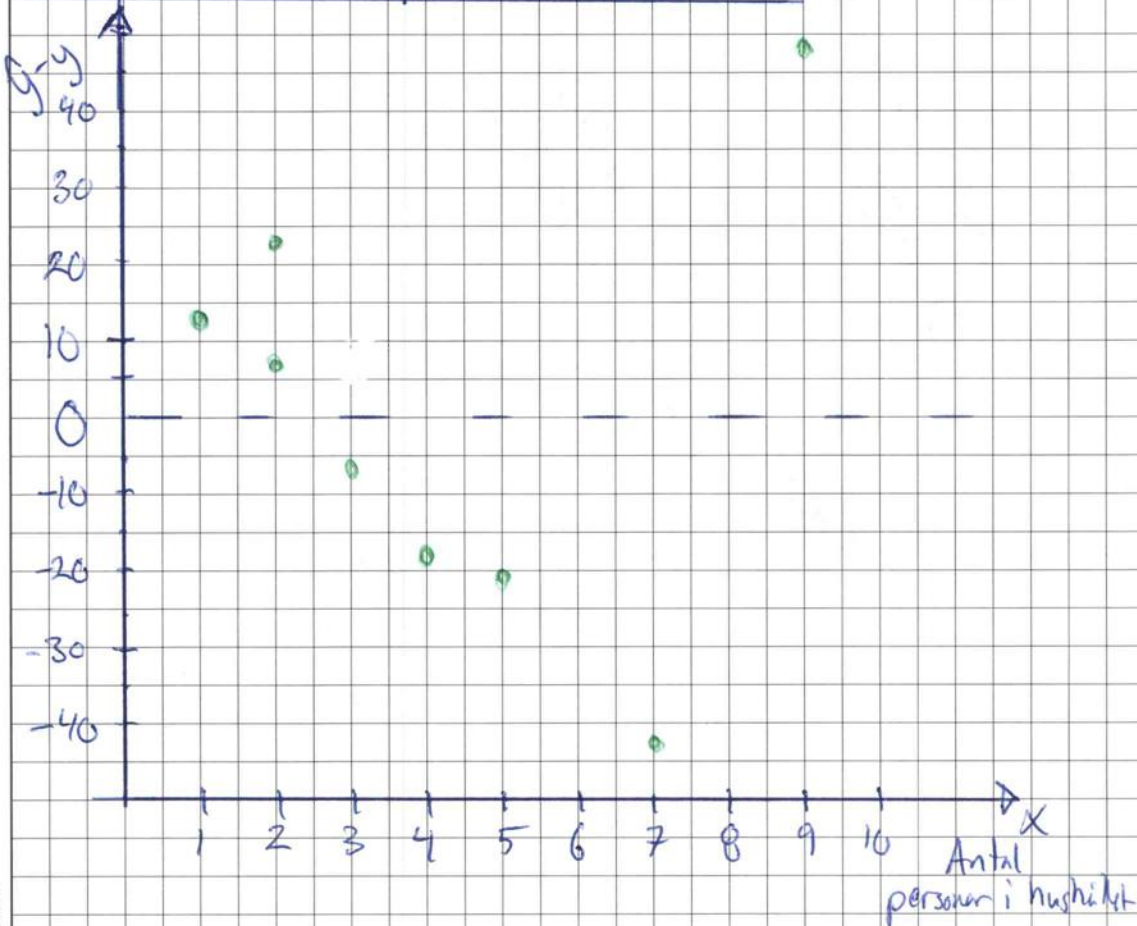
1

Poäng / Points
 awarded:

→

Lärarens
 anteckning
 Examiner's remarks:

x	y	\hat{y}	$\hat{y} - y$
2	59	81,236	22,236
5	148	127,313	-20,687
1	52	65,877	13,877
7	202	158,091	-43,969
2	75	81,236	6,236
9	140	188,749	48,749
3	104	96,595	-7,405
4	131	111,954	-19,046





Ange anonymitetskod / Write your anonymity code
(Vid icke anonym tentamen ange kurskod + namn + personnummer)
(For non-anonymous exams write the course code + name + civic registration number)

STGAG1-122

Löpande sidnr
Consecutive no:

4

Uppgift nr /
Question no:

1

Poäng / Points
awarded:

7

Lärarens
anteckning
Examiner's remarks:

e) Genom att rita in residualerna i ett spridningsdiagram kan vi se om det finns något mönster. Om en rät-linje skulle vara det optimala skulle det inte finnas något mönster utan residualerna skulle ligga spridda.

I det här fallet kan vi se att residualerna till en början ligger över 0, sedan ligger de under 0, och sedan över igen.

Vi kan där för dra slutsatsen att en linje inte är det optimala i detta fallet utan att en kurva antagligen passar bättre.

Ⓢ

Häftområde

Skriv ej i detta område
Leave this area blank



Ange anonymitetskod / Write your anonymity code
 (Vid icke anonym tentamen ange kurskod + namn + personnummer)
 (For non-anonymous exams write the course code + name + civic registration number)

STGA01-122

Löpande sidnr
 Consecutive no:

5

Häftområde

Skriv ej i detta område
 Leave this area blank

Uppgift 2

a) $P(H) = 0,48$: Sannolikheten att en slumpvis utvald student har förälder med högskoleutbildning är 48%

~~b)~~ $P(K^c \text{ och } H) = 0,10$ Sannolikheten att en slumpvis utvald student inte känner krav och har förälder med högskoleutbildning är 40%

$P(K|H^c) = 0,10$ Sannolikheten att en slumpvis utvald student känner krav givet att de inte har förälder med högskoleutb. är 10%

	K	K ^c	total
H	800	4000	4800
H ^c	520	4680	5200
total	1320	8680	10000

b) $P(K) = \frac{1320}{10000} = 0,132$ svar: 13,2%

c) $P(K|H) = \frac{800}{4800} = 0,1667$ svar: 16,7%

d) $P(H|K) = \frac{800}{1320} = 0,606$ svar: 60,6%

Uppgift nr /
 Question no:
 2

Poäng / Points
 awarded:

10

Lärarens
 anteckning
 Examiner's remarks:



Ange anonymitetskod / Write your anonymity code
(Vid icke anonym tentamen ange kurskod + namn + personnummer)
(For non-anonymous exams write the course code + name + civic registration number)

STGA01-122

Löpande sidnr
Consecutive no:

6

Uppgift nr /
Question no:
2

Poäng / Points
awarded:

Lärarens
anteckning
Examiner's remarks:

e) $P(K \text{ och } H) = \frac{800}{10000} = 0,08$ svar: 8%

f) $P(K \text{ eller } H) = \frac{4000 + 800 + 520}{10000} = 0,532$ svar: 53,2%

g) svar: Ja det verkar som att föräldrarnas utbildning påverkar hur mycket studenterna känner krav eftersom $P(K|H) \neq P(K|H^c)$
Det verkar alltså som att det finns en association mellan variablerna föräldras utbildning och hurvidt studenten kände krav. Eftersom variablerna inte verkar vara oberoende.

K

Häftområde

Skriv ej i detta område
Leave this area blank



Uppgift 3

a) $n = 15$ $N = 200$ $p = 0,1$

Antaganden:

• Eftersom $\frac{n}{N} < 0,1$ följer sannolikhetsfördelningen
approximativt en Binomial fördelning

• Kategorisk variabel, slumpmässigt urval

$X =$ Antalet bräder som är defekta

$X \sim \text{Bin}(15, 0.1)$

$P(X \leq 1) = P(0) + P(1) = 0,2059 + 0,3432 = 0,5491$

$P(0) = {}^{15}n_C0 \times 0,1^0 \times 0,9^{15} = 0,2059$
 $P(1) = {}^{15}n_C1 \times 0,1^1 \times 0,9^{14} = 0,3432$ } $P(X \leq 1)$

svare: Sannolikheten att kalle får sälja en
bunt är ca 54,9% R

b) Antaganden:

- Kategorisk variabel
- Slumpmässigt urval
- Samplingsfördelningen för stockpart
approximativt normalfördelat eftersom
 $np \geq 15$ och $n(1-p) \geq 15$
(80) (120)

$n = 800$

$p = 0,1$

$X =$ antalet defekta bräder

$X \sim N(np, \sqrt{np(1-p)})$

$\mu = np = 800 \times 0,1 = 80$

$\sigma = \sqrt{800 \times 0,1 \times (1-0,1)} \approx 8,48528$

$X \sim N(80, \sqrt{80 \times 0,9})$



Ange anonymitetskod / Write your anonymity code
(Vid icke anonym tentamen ange kurskod + namn + personnummer)
(For non-anonymous exams write the course code + name + civic registration number)

STGA01-122

Löpande sidnr
Consecutive no:

8

Uppgift nr /
Question no:

3

Poäng / Points
awarded:

Lärarens
anteckning
Examiner's remarks:

fortsättning uppgift 3b

$$P(X \leq 90) = P\left(\frac{90-80}{\sqrt{100 \cdot 0,01}} \leq Z\right) \approx P$$
$$= P\left(Z \leq \frac{90-80}{\sqrt{100 \cdot 0,01}}\right) \approx P(Z \leq 1,18) = 0,8810$$

Svar: Sannolikheten för att Bob får högst
90 defekta bräddor i sitt köp är 88,1% R

Häftområde

Skriv ej i detta område
Leave this area blank



Ange anonymitetskod / Write your anonymity code
 (Vid icke anonym tentamen ange kurskod + namn + personnummer)
 (For non-anonymous exams write the course code + name + civic registration number)

STGA01-122

Löpande sidnr
 Consecutive no:

9

Häftområde

Skriv ej i detta område
 Leave this area blank

Uppgift 4

Antaganden

a) Kategorisk variabel; upprepade försök (n)

V varje försök samma sannolikhet att lyckas

V varje försök har två svarsalternativ (lyckas / inte lyckas)

$p = \frac{2}{5}$

$n = 2$

X = antalet gånger du får högt värde

$X \sim \text{Bin}(2, \frac{2}{5})$

Antal

X	P(x)	
0	0,36	$2nCn0 \times 0,4^0 \times 0,6^2$
1	0,48	$2nCn1 \times 0,4^1 \times 0,6^1$
2	0,16	$2nCn2 \times 0,4^2 \times 0,6^0$

3

Svar Antal

b)

X	P(x)
0%	0,36
50%	0,48
100%	0,16

annan beräkning än x

Svar Andel

2,5

~~$$\sum x P(x) = 1 \times \frac{1}{5} + 2 \times \frac{1}{5} + 3 \times \frac{1}{5} + 4 \times \frac{1}{5} + 5 \times \frac{1}{5}$$~~

~~$$= 3$$~~

c) ~~Kategorisk~~ kvantitativ variabel

~~Antal~~
 fortsätter blad
 nästa





Ange anonymitetskod / Write your anonymity code
 (Vid icke anonym tentamen ange kurskod + namn + personnummer)
 (For non-anonymous exams write the course code + name + civic registration number)

Stga 01-122

Löpande sidnr
 Consecutive no:

10

Häftområde

Skriv ej i detta område
 Leave this area blank

c)

Resultat Medelvärde

1+1=2	1
1+2=3	1,5
1+3=	2
1+4	2,5*
1+5	3
2+1	1,5
2+2	2
2+2 2+3	2,5
2+4	3
2+5	3,5
3+1	2
3+2	2,5
3+3	3
3+4	3,5
3+5	4
4+1	2,5
4+2	3
4+3	3,5
4+4	4
4+5	4,5
5+1	3
5+2	3,5
5+3	4
5+4	4,5
5+5	5

Sannolikhetsfördelning

Medelvärde(x)	Sannolikhet P(x)
1	1/25 = 0,04
1,5	2/25 = 0,08
2	3/25 = 0,12
2,5*	4/25 = 0,16
3	5/25 = 0,20
3,5	4/25 = 0,16
4	3/25 = 0,12
4,5	2/25 = 0,08
5	1/25 = 0,04

Antal Permutationer $5 \times 5 = 25$

4

Uppgift nr /
 Question no:

4

Poäng / Points
 awarded:

9,5

Lärarens
 anteckning
 Examiner's remarks:



9,5

Uppgift 5a

a) Anläggande:

- En kvantitativ variabel
- De 8 vatten proven slumpvis utvalda i två grupper
- ~~Att skillnaden i stickprovsmedel~~

Att samplingfördelningen för skillnaden i stickprovsmedelvärden är approximativt normalfördelat eftersom populationen för skillnaden i medelvärde är det (normalfördelat)

0,5 Fuset parvis

0,5 Motivering ges i b-uppgiften

Vi skapar en ny variabel (X_d) som står för skillnaden mellan resultatet i de olika metoderna.

$$X_d = X_B - X_A$$

$$\bar{X}_d = \frac{\sum X_d}{n} \quad S_d = \sqrt{\frac{\sum (X_d - \bar{X}_d)^2}{n-1}}$$

Day	X_d	$(X_d - \bar{X}_d)^2$
1	0,4	0,1502
2	0	0,0002
3	-0,5	0,2627
4	0,5	0,2377
5	-0,3	0,0977
6	0,1	0,0077
7	-0,3	0,0977
8	0,2	0,0352
Summa	0,1	0,8891

$$S_d = \sqrt{\frac{0,8891}{8-1}} \approx 0,3564$$

$$95\% \text{ KI: } \bar{X}_d \pm t \cdot \frac{S_d}{n}$$

$$t = 2,361 \quad (\text{Från } t\text{-tabel där } df = 7)$$

$$0,0125 \pm 2,36 \times \frac{0,3564}{\sqrt{8}}$$

$$\begin{aligned} \text{övre gräns} &: 0,3001 \\ \text{undre gräns} &: -0,285 \end{aligned}$$

$$\bar{X}_d = \frac{0,1}{8} = 0,0125$$

Se nästa blad →



Ange anonymitetskod / Write your anonymity code
(Vid icke anonym tentamen ange kurskod + namn + personnummer)
(For non-anonymous exams write the course code + name + civic registration number)

STGA01-122

Löpande sidnr
Consecutive no:

12

Uppgift nr /
Question no:
5

Poäng / Points
awarded:

Lärens
anteckning
Examiner's remarks:

Fortsättning uppgift 5a

Tolkning: Med 95% säkerhet finns ^{Medelskillnaderna} skillnaden mellan metoderna ^{0,5} inom intervallet: $[-0,285, 0,310]$

b) För att kunna göra konfidensintervallet med så här få värden antog jag att populationen ^{0,5} för skillnaden mellan metoderna var normalfördelad

- Jag antog även: - kvantitativ variabel ^{0,5}
- slumpmässigt urval i två grupper ^{0,5} med ett beroende/association

c) Nej jag skulle inte dra slutsatsen att det är någon skillnad mellan metoderna eftersom värdet noll finns med i intervallet.



Uppgift 6

a) Hypotesbäst

1. Antagande

- Kategorisk variabel
- Slumpmässigt urval
- Eftersom $\frac{n}{N} < 0,1$ och $np \geq 15$ & $n(1-p) \geq 15$

är samplingfördelningen för stickprovsandelen
 approximativt normalfördelad $\hat{p} \sim (p, \sqrt{\frac{p(1-p)}{n}})$

$$np = 800 \times 0,25 = 200 > 15$$

$$n(1-p) = 800 \times 0,75 = 600 > 15$$

2. Hypotes

$$H_0: p = 0,25$$

$$H_a: p > 0,25$$

3. Teststatistiken

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \quad 0,5$$

$$\hat{p} = \frac{225}{800} = 0,28125$$

$$z = \frac{0,28125 - 0,25}{\sqrt{\frac{0,25 \times 0,75}{800}}} \approx 2,04 \quad |$$

4. P-värde 0,5

$$P(Z \geq 2,04) = 1 - 0,9793 = 0,0207 \quad 0,5$$

Fortsätter på
 nästa blad



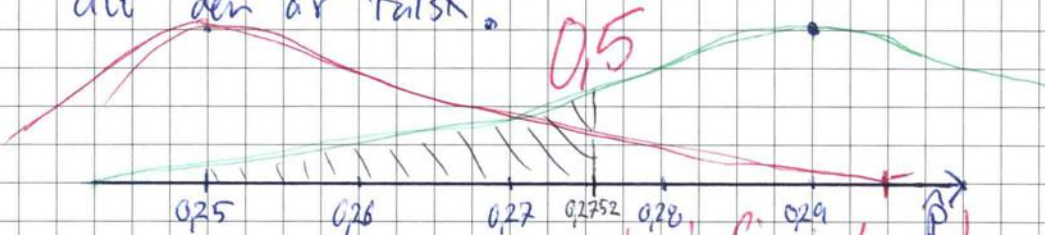
fortsättning på uppgift 5a

5. Slutsets

Eftersom p-värdet $< \alpha$ ($0,0207 < 0,05$) förkastas H_0 0,5

Svar: Med en signifikansnivå på 5% vågar vi påstå att Karins påstående inte stämmer och att andelen som röstar på socialdemokraterna är mer än ~~25%~~ 25%. 0,5

b) Typ II fel är när man inte förkastar H_0 fast att den är falsk.



Gränsen för att förkasta H_0

$$\frac{\hat{p} - 0,25}{\sqrt{\frac{0,25 \times 0,75}{800}}} = 1,645$$

$$\hat{p} \approx 0,2752 \quad \text{0,5}$$

(Motivområde från under antaganden)

$$P(\text{typ II fel}) = P(\hat{p} \leq 0,2752)$$

$$= P(Z \leq \frac{0,2752 - 0,29}{\sqrt{\frac{0,29 \times 0,71}{800}}})$$

$$= P(Z \leq -0,92) = 0,1788 \quad \text{0,5}$$

Svar: Om S i verkligheten är 29% är sannolikheten att göra ett typ II fel 17,88%

Uppgift nr /
 Question no:
 6

Poäng / Points
 awarded:

Lärarens
 anteckning
 Examiner's remarks:



fortsättning uppgift 6

$$\begin{aligned} c) \quad & P(\hat{p} \geq 0,2752) \quad \text{om } p = 0,29 \\ & = P\left(Z \geq \frac{0,2752 - 0,29}{\sqrt{\frac{0,29 \cdot 0,71}{800}}}\right) \approx P(Z \geq -0,92) = (1 - 0,1788) \\ & = 0,8212 \end{aligned}$$

~~svar~~ eller $1 - \beta = 1 - 0,1788 = 0,8212$ |

Svar: Sannolikheten att vi förkastar H_0 om det sanna värdet är 29% är ca 82,1%

d) Hypotestestets styrka brukar användas för ~~ett~~ som benämning för beräkningen i c-uppgiften och räknas genom formeln $1 - \beta$ där β står för sannolikheten för ett typ II fel. 0,5

Uppgift nr /
Question no:

6

Poäng / Points
awarded:

Lärarens
anteckning
Examiner's remarks: