



Antal blad /
Number of sheets

18 ✓

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
---	---	---	---	---	---	---	---	---	---

Stga01 Statistik

Kurskod + Kurs / Course Code + Course:

Delkurs / Part course:

Anonymitetskod / Anonymous code =
Kurskod + kodnr / course code + code number

S	T	G	A	Ø	1	-	1	Ø	8
---	---	---	---	---	---	---	---	---	---

Tentamensdatum /
Examination date:

22 Jan 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
8,5	10	9	6,5	2,5	8,5									
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30

Poäng / Marks gained: 45

Betyg / Grade: VG

Arle
Examin. lärare / Kursansvarig signatur / Signature of the examiner

Max poäng / Total marks gained: _____

Namnförtydligande / Clarification of the signature

För Gk poäng / Marks gained to be passed: _____



Stga01-108

1

Häftområde

Skriv ej i detta område
 Leave this area blank

1A | Förklarande variabeln x (Hushållsstor.)
 Responsvariabeln y (Vattenförbr.)
 Vattenförbrukningen är
 responsvariabeln.

Uppgift nr /
 Question no: 1

Poäng / Points
 awarded:

Lärens
 anteckning
 Examiner's remarks:

1B

Korrelationskoefficienten: $n=8$

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

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
\sum 33	911	4570	189	122095

$$r = \frac{8 \cdot 4570 - 33 \cdot 911}{\dots}$$

Kvantitativa variabler!



1.b

$$r = \frac{8 \cdot 4570 - 33 \cdot 911}{\sqrt{[8 \cdot 189 - 33^2][8 \cdot 182095 - 911^2]}}$$

$$r = \frac{36560 - 30063}{\sqrt{[1512 - 1089][976760 - 829921]}}$$

$$r = \frac{6497}{\sqrt{423 \cdot 146839}}$$

$$r = \frac{6497}{\sqrt{62112897}} = \frac{6497}{7881,1735} \approx 0,8243696$$

Svar: Korrelationskoefficienten (r) är i detta fall $\approx 0,824$.

Delta betyder att det finns ett starkt linjärt samband mellan variablerna då $r > 0,7$.

Bra linjärsamband finns då r är antingen $r > 0,7$ eller mindre än $-0,7$.

Uppgift nr /
Question no: 1

Poäng / Points
awarded:



Lärens
anteckning
Examiner's remarks:



1c $\hat{y} = 50,518 + 15,359x$

o Tolkning 50,518: står för talet där linjen skär y-axeln. I detta fall innebär det minimum vattenförbrukning. Dvs det finns inget hushåll som använder 0 vatten utan minimum är 50,518 (m³). Så i ett hushåll där 0 personer bor förbrukas ändå vatten. Trogen i form av vattenburna element eller liknande.

o Tolkning 15,359: står för den genomsnittliga ökningen på y-axeln då man går ett steg till höger på x-axeln. I detta fall innebär det att 1 till person i hushållet ökar i genomsnitt vattenförbrukningen med 15,359 (m³). 2 personer ökning ger + 30,718 (m³) osv.

Uppgift nr /
Question no: 1

Poäng / Points
awarded:

Lärarens
anteckning
Examiner's remarks:



Stga 01-108

1d Residualer: $\hat{y} - y$

OBS!

Avrundat till heltal

x Hushåll:	y	$\hat{y} (*)$	$\hat{y} - y$
2	59	81,236	22
5	148	127,313	-21
1	52	65,877	14
7	202	158,031	-44
2	75	81,236	6
9	40	188,749	38
3	104	96,595	-7
4	131	111,954	-19

Uppgift nr /
 Question no: 1

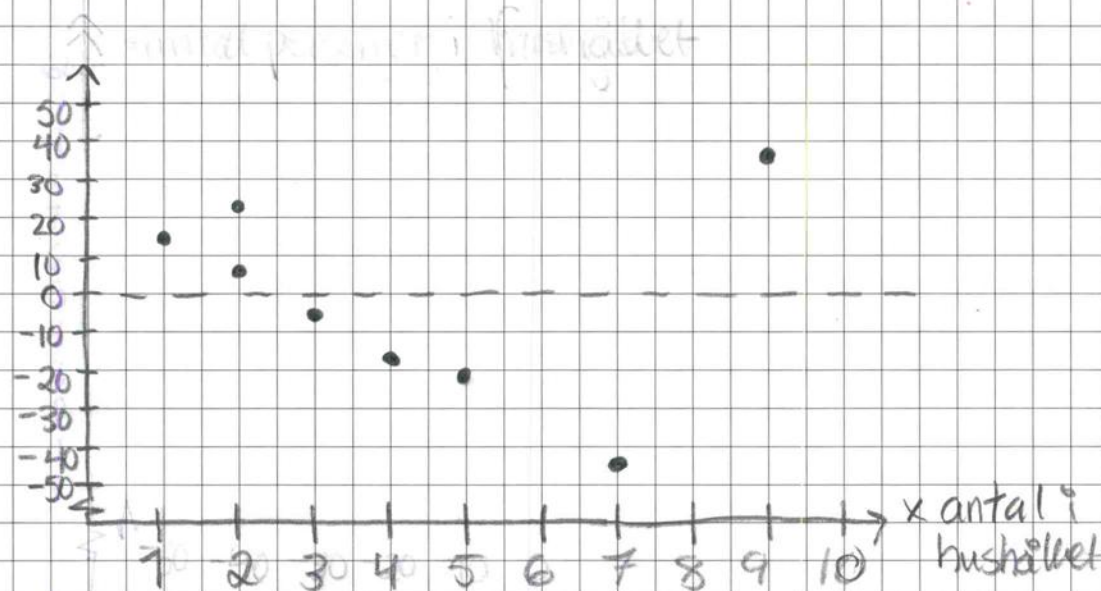
Poäng / Points
 awarded:

Lärarens anteckning
 Examiner's remarks:

(*) Här sätter jag in de olika x värdena i linjens ekvation och får fram \hat{y} för respektive hushåll.

OBS! Jag avrundar till heltal i $\hat{y} - y$ för enkelhetens skull.

Residualer





1e

Att rita ut residualer i en graf ger en bild av om det är tillförlitligt att rita en linje och dra slutsatser från den. Finns det ett mönster i residualerna så kanske inte linjen går att dra slutsatser från. Man vill ha en ojämn spridning på residualerna.

I vårt fall verkar det finnas något slags mönster då det nästan går att rita en linje vid 1-7. Detta är inte bra och är därför något man bör ha i åtanke om man vill dra slutsatser från linjen som nämnts tidigare. Det verkar som att ju fler man är i hushållet ju mer övertrasslar man den beräknade vattenmängden.

(med ett undantag)

Uppgift nr /
Question no: 1

Poäng / Points
awarded:

8,5

Lärarens
anteckning
Examiner's remarks:



Stga01-108

6

Häftområde

Skriv ej i detta område
 Leave this area blank

2.a

H: Händelsen minst 1 förälder med högskoleutb.

K: Händelsen studenten känner krav från minst 1 förälder.

Uppgift nr /
 Question no: 2

Poäng / Points
 awarded: 10

Lärarens
 anteckning
 Examiner's remarks:

o $P(H) = 0,48$: 48% är sannolikheten att en elev har minst 1 förälder med högskoleutbildning.

o $P(K^c \text{ och } H) = 0,40$: 40% är sannolikheten att en elev har en förälder med högskoleutbildning och inte känner krav.

o $P(K|H^c) = 0,10$: 10% är sannolikheten som känner krav givet att eleven inte har någon förälder med högskoleutbildning.

Jag valde att göra en tabell över det hela.

	H	H ^c	Σ
K	80	52	132
K ^c	400	468	868
Σ	480	520	1000 studenter

R



Stga01-108

7

Häftområde

Skriv ej i detta område
Leave this area blank

Uppgift nr /
Question no: 2

Poäng / Points
awarded:

Lärarens
anteckning
Examiner's remarks:

2 [b]: $P(K) = \frac{132}{1000} = 0,132$ dvs 13,2%

[c]: $P(K|H) = \frac{80}{480} \approx 0,1667$ dvs 16,67%

[d]: $P(H|K) = \frac{80}{132} \approx 0,606$ dvs 60,6%

[e]: $P(K \text{ och } H) = \frac{80}{1000} = 0,08$ dvs 8%

[f]: $P(K \text{ eller } H) = \frac{80 + 52 + 400}{1000} = 0,532$
dvs 53,2%

[g]: För att ta reda på det krävs en "oberoende kontroll".

$$P(H \text{ och } K) = P(H) \cdot P(K)$$

$$8\% = 0,48 \cdot 13,2 \text{ (5,94\%)}$$

Dessa 2 är olika och detta betyder att de inte är oberoende. Alltså det finns ett samband mellan krav och föräldrars utbildning. Hade det inte funnits ett samband hade båda sidorna av = tecknet givits samma svar. Så Svar JA!

X



Häftområde

Skriv ej i detta område
Leave this area blank

3 A: 3 m bräddor 200st/bunt
10% är defekta

A: $n=15$ om $P(X>1)$ $p=0,1$

Kategorisk variabel (defekt/ej defekt)

X = Slumpvis utvald brädda som är defekt.

Villkoret $\frac{n}{N} < 0,1$ antar vi är uppfyllt då $n=15$ och N = stort antal. Då kan vi använda bin formeln! $X \sim \text{Bin}(n=15, p=0,1)$

Bör nämnas är att även villkoren OSU och flera likadana upprepningar är uppfyllda.

Men alltså $X \sim \text{Bin}(15, 0,1)$ vi söker $P(X \leq 1)$ alltså att högst 1 brädda är defekt så Kalle får sälja en bunt.

$$P(0) = \frac{15!}{0!(15-0)!} \cdot 0,1^0 \cdot 0,9^{15} \approx 0,2058908$$

$$P(1) = \frac{15!}{1!(15-1)!} \cdot 0,1^1 \cdot 0,9^{14} \approx 0,3431514$$

$$0,2058908 + 0,3431514 = 0,5490422$$

Svar: Sannolikheten att Kalle får sälja en bunt till Augusta är ca 54,9%

R



3B 710 bräddor köper 800
 Sannolikheten högst 90 defekta.

är det, approximativt med
 Bin går ~~ei~~ *ett*. Normalfördelning villkor:
 $np > 15$ (Ja 80), $n(1-p) > 15$ (Ja 720) $\frac{n}{N} < 0,1$ får
 vi anta är uppfyllt. $X = 1$ defekt 0,01 brädda.

$$\mu = np = 800 \cdot 0,01 = 80$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{72} = 8,4852813 \approx 8,485$$

$$X \sim N(80, 8,485) \quad P(X \leq 90)$$

$$Z: \frac{90-80}{8,485} = P(Z \leq 1,18) \text{ enl tabell:}$$

0,8810.

Svar: Sannolikheten att bob får
 högst 90 defekta bräddor är *R*
 88,1%.



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)

Stga01-108

Löpande sidnr
 Consecutive no:

10

Uppgift nr /
 Question no: 4

Poäng / Points
 awarded:

Lärarens
 anteckning
 Examiner's remarks:

Häftområde

Skriv ej i detta område
 Leave this area blank

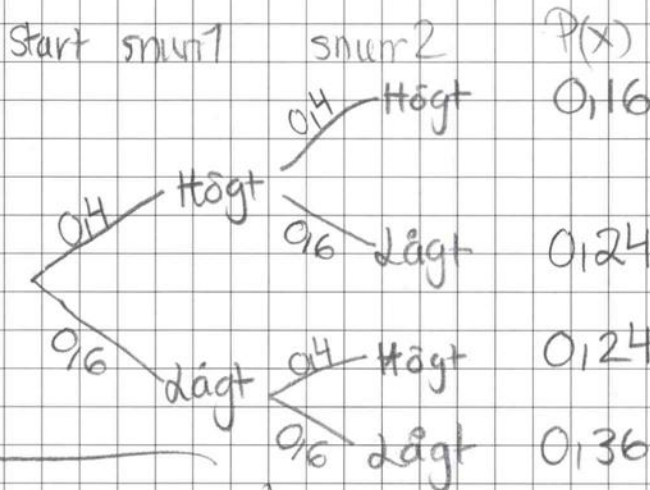
4a:



Högt {4, 5} 0,16
 Lågt {1, 2, 3} 0,16

X	1	2	3	4	5
P(x)	0,2	0,2	0,2	0,2	0,2

X = Antalet gånger du får ett högt tal.



X	P(x)
0	0,36
1	0,48
2	0,16

Svar A: Så här ser sannolikhetsfördelningen för antalet gånger med högt värde ut. ✓ 3

4b)

X	Andel
0	0/3 (0)
1	1/3 (0,333...)
2	2/3 (0,6666...)

Svar B så här ser sannolikhetsfördelningen för andel ut! ✓ 0



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)

Stga01-108

Löpande sidnr
 Consecutive no:

11

Häftområde

Skriv ej i detta område
 Leave this area blank

4.c



Högt värde {4,5}
 lågt värde {1,2,3}

Uppgift nr /
 Question no: 4

Poäng / Points
 awarded:



Lärarens
 anteckning
 Examiner's remarks:

Utfall:		Poängmedelv	Sannolikhet
Spel 1:	2:		
1	1	1	0,04 (0,02 · 0,02)
1	2	1,5	0,04
1	3	2	...
1	4	2,5	...
1	5	3	...
2	1	1,5	Samma sannolikhet
2	2	2	hela vägen
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	





Stga01 - 108

12

Häftområde

Skriv ej i detta område
 Leave this area blank

4c1

Sammanfattad tabell.

Medelvärde	Sannolikhet för det.
1	0,04 (1gång)
1,5	0,08 (2ggr)
2	0,12 (3ggr)
2,5	0,16 (4ggr)
3	0,2 (5ggr)
3,5	0,16 (4ggr)
4	0,12 (3ggr)
4,5	0,08 (2ggr)
5	0,04 (1gång)

Uppgift nr /
 Question no: 4

Poäng / Points
 awarded: 6,5

Lärarens
 anteckning
 Examiner's remarks:

Svar C så här ser sannolikhets
 fördelningen för medelvärdet
 ut!

Totalsannolikhet att få högt värde
 $0,12 + 0,08 + 0,16 = 0,36$
 alltså 36% sannolikhet.

3,5



5	DAG:	1	2	3	4	5	6	7	8
X_1	A:	28	31	34	30	27	29	35	26
X_2	B:	32	31	29	35	24	30	32	28

✓ Situationen är "penning"
A: 2 populationer kvantitativ variabel

$$(\bar{x}_1 - \bar{x}_2) \pm t \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

df = $n_1 + n_2 - 2$ om s_1 och s_2 och n_1 och n_2 är nära varandra.

$$\bar{x}_1 = (\bar{x} = \frac{\sum x}{n}) = \frac{24}{8} = 3$$

$$\bar{x}_2 = \dots = \frac{24,1}{8} = 3,0125$$

$$n_1 = 8$$

$$n_2 = 8$$

$$s_1 = \left(\sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} \right) = \frac{0,04 + 0,01 + 0,16 + 0 + 0,09 + 0,01 + 0,25 + 0,16}{7} \approx 0,11$$

OBS Jag arrunderar här till 3 decimaler ↓

$$s_2 = \dots = \frac{0,035 + 0,008 + 0,013 + 0,238 + 0,375 + 0 + 0,035 + 0,045}{7} \approx 0,11$$

Jag inser nu att i och med att det är 1 prov delat på 2 alltså samma vatten kan man ju anta att de har samma spridning.



5.a) Då är det formeln

$$(\bar{x}_1 - \bar{x}_2) \pm t \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad \text{som gäller.}$$

Delta borde jag ju såklart förställt från början... Jaja nu räknar vi resten!

$$df = \text{Då } s_1 \text{ och } s_2 \text{ och } n_1 \text{ och } n_2 = \text{samma } 8 + 8 - 2 = 14$$

T-tabellen säger att vid 95% och df 14 är $t = 2,14$

$$\text{övre: } (3 - 3,0125) + 2,14 \sqrt{\frac{0,11^2}{8} + \frac{0,11^2}{8}} = 0,0945$$

$$\text{nedre: } (3 - 3,0125) - 2,14 \sqrt{\frac{0,11^2}{8} + \frac{0,11^2}{8}} = -0,1195$$

Svar: Med 95% säkerhet finns skillnaden inom KI $[-0,1195, 0,0945]$

B: Antaganden:

Kvantitativ variabel

OSU obundet slumpmässigt urval
lika spridning i populationen.



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)

Stga01-108

Löpande sidnr
Consecutive no:

15

Uppgift nr /
Question no:

5

Poäng / Points
awarded:

Lärarens
anteckning
Examiner's remarks:

5 C:

Eftersom KI skulle kunna
vara null[?] förs jag inte
dra slutsatsen att det är
någon skillnad med 95% säkerhet.

Häftområde

Skriv ej i detta område
Leave this area blank



6. Karin påstår 25% socialdemokrat-
stöd. Du påstår högre. $n=800$

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

1. Antaganden: • Kategorisk variabel
• OSU

• Villkoren ($np > 15$ ($0,25 \cdot 800 = 200$)
är uppfyllda ($n(1-p) > 15$ ($800 \cdot 0,75 = 600$))
 $\left(\frac{n}{N} = \text{Ja eftersom storpop.}\right)$

2 Hypoteser: $H_0: p = 0,25$
 $H_a: p > 0,25$

3 Teststatistika:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0,28125 - 0,25}{0,0153} = P(Z \geq 2,04)$$

$$\sqrt{\frac{0,25 \cdot 0,75}{800}} \approx 0,0153$$

4. P-värde

$$\text{Tabellen ger } P(Z \geq 2,04) = 1 - 0,9793 = 0,0207$$

5. Slutsats: P-värdet är $< \alpha(0,05)$ så
 H_0 förkastas. På 5% signifikansnivå
tors vi säga att andelen socialar är över 25%!



6. B: Verkligheten 29% stöd.

Typ II-fel är när man accepterar H_0 trots att den är falsk.

Detta händer då $P(Z < z) = 0,05$
alltså vid $-1,645$ så vad har vi
för \hat{p} vid $z = -1,645$:

$$\frac{\hat{p} - 0,25}{0,0153} = -1,645$$

$$65,359477x - 16,339869 = 1,645$$

$$65,359477x = 17,984869$$

Mohamed?
 $\hat{p} \approx 0,275$ ^{0,5} Alltså förkastar

vi H_0 om $P(\hat{p} \geq 0,275)$ och accepterar

\downarrow H_0 om $P(\hat{p} \leq 0,275)$ ^{0,5}

$$z: \frac{0,275 - 0,29 \text{ (verkligarna)}}{0,0153} = P(Z \leq -0,98)$$

enl tabell $\approx 0,1635$ ^{0,5}

Svar: Sannolikheten att göra ett
typ II-fel är $16,35\%$ ^{0,5 (falskt)}

Uppgift nr /
Question no: 6

Poäng / Points
awarded:

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)

Stga 01-108

Löpande sidnr
Consecutive no:

18

Uppgift nr /
Question no: 6

Poäng / Points
awarded:

Lärarens
anteckning
Examiner's remarks:

6c: Eftersom sannolikheten att
man riktar acceptera H_0 vid $p=0,29$
i förra frågan var 16,35%
så är sannolikheten att man
förkastar H_0 vid $1 - 0,1635 = 0,8365$
83,65%

d: Approximativ sannolikhet ✓

Häftområde

Skriv ej i detta område
Leave this area blank