KARLSTADS UNIVERSITET

12307683

TENTAMEN / EXAMINATION

Fylls i av **student** / *To be completed by the* ***student***

Skriv anonymiseringskoden på samtliga svarsblad / *Write your anonymity code on each sheet*	Anonymiseringskod / *Anonymity code*: S T G A 0 1 - 0 1 4 5 - T A K
Provbenämning / *Exam name*: Skriftlig tentamen	Oanmäld
Kurskod / *Course code*: S T G A 0 1	Modul / *Module*: 1 0 0 2
Tentamensdatum / *Examination date*: 2 0 2 5 - 0 1 - 1 6	
Jag har tagit del av regler som gäller vid tentamen / *I have read the current rules for examinations* ☒ Ja / *Yes*	Antal inlämnade blad med anonymiseringskod / *Number of sheets with anonymity code*: 0 8 ✓

Fylls i av **skrivvakt** / *To be completed by the* ***invigilator***

Kontroll av legitimation / *Identification checked*	☒ Ja / Yes	Härmed intygas att kontroller utförts / *This is to certify that the checks have been carried out*
Kontroll av inlämnade blad / *Answer sheets checked*	☒ Ja / Yes MH	Tydlig sign. / *Signature*
Inlämningstid / *Time of submission*	11:56	BMA

Fylls i av **lärare** / *To be completed by the* ***examiner***

Bedömning av uppgifter / Questions attempted

1	2	3	4	5	6	7	8	9	10	~
9,5	10	9,5	9,5	10	10					

11	12	13	14	15	16	17	18	19	20	~

21	22	23	24	25	26	27	28	29	30	~

Totalt antal poäng / *Total points*: 58,5	Examin. lärare / Kursansvarig signatur / *Signature of the examiner*: A.W
Betyg / *Grade*: VG	Namnförtydligande / *Clarification of the signature*

12307683

Försättsbladet ska alltid lämnas in även om ingen uppgift behandlats / *Examination should always be submitted even if no questions are answered*

Utskriven 2025-01-03 kl. 09:42:27

Häftområde

Skriv ej i detta område
Leave this area blank

KARLSTADS UNIVERSITET

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)

STGA01-0145-TAK

Löpande sidnr
Consecutive no: 1

Uppgift nr / *Question no:* 1

Poäng / *Points awarded:* →

Lärarens anteckning
Examiner's remarks:

Uppgift 1

För att besvara frågorna i uppgift 1 har jag skapat en tabell för att göra beräkningarna smidigare

x	y	x·y	x^2	y^2
40	110	4400	1600	12100
35	145	5075	1225	21025
35	140	4900	1225	19600
32	160	5120	1024	25600
30	160	4800	900	25600
28	162	4536	784	26244
25	150	3750	625	22500
30	155	4650	900	24025
Σ 255	1182	37231	8283	176694

n = 8

a) korrelationskoefficienten $r = \dfrac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$

från tabellen får vi: $r = \dfrac{8 \cdot 37231 - 255 \cdot 1182}{\sqrt{[8 \cdot 8283 - 255^2][8 \cdot 176694 - 1182^2]}}$

→ $r = \dfrac{-3562}{\sqrt{1239 \cdot 16428}} = \dfrac{-3562}{\sqrt{20354292}} \approx -0.7895$

Eftersom korrelationskoefficienten är mindre än −0.7 finns det ett starkt linjärt samband mellan variablerna 'pris i kr' och 'kvantitet i liter'. Detta innebär att kvantiteten påverkas negativt av priset, det vill säga när priset ökar minskar kvantiteten.

3

b) $\hat{y} = a + bx$

$b = \dfrac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \dfrac{8 \cdot 37231 - 255 \cdot 1182}{8 \cdot 8283 - 255^2} = \dfrac{-3562}{1239} \approx -2.8749$

$a = \bar{y} - b\bar{x}$

→ $\bar{y} = \dfrac{1182}{8} = 147.75$

→ $\bar{x} = \dfrac{255}{8} = 31.875$

$a = 147.75 - (-2.8749) \cdot 31.875 \approx 239.39$

2,5

regressionslinjen får följande ekvation: $\hat{y} = 239.39 - 2.8749x$

(fortsättning nästa sida)

Universitetstryckeriet

Häftområde

Skriv ej i detta område
Leave this area blank

KARLSTADS UNIVERSITET

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)

STGA01-0145-TAK

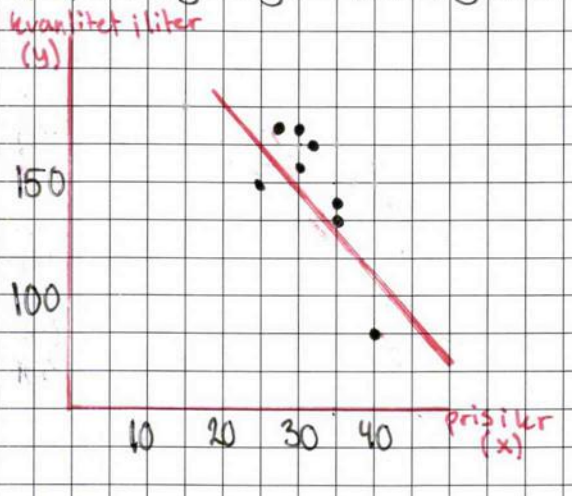
Löpande sidnr
Consecutive no: 2

Uppgift nr / *Question no:* 1

Poäng / *Points awarded:* 9,5

Lärarens anteckning
Examiner´s remarks:

b) spridningsdiagram med regressionslinjen inritad:



c) för att uppskatta hur mycket Harry skulle få sälja om han satte priset 38 kr sätter jag in 38 i ekvationen från b) som värdet på x. Jag får då följande resultat:

$\hat{y} = 239.39 - 2.8749 \cdot 38 \approx 130$ liter

1

d) residualerna $= y - \hat{y}$

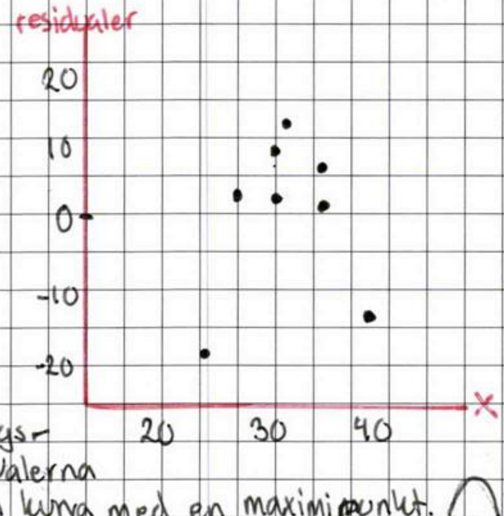
för att få fram värdena på $\hat{y}$ sätter jag in värdena på x i ekvationen för regressionslinjen och får då följande värden:

- $x = 40 \rightarrow \hat{y} \approx 124$
- $x = 35 \rightarrow \hat{y} \approx 139$
- $x = 32 \rightarrow \hat{y} \approx 147$
- $x = 30 \rightarrow \hat{y} \approx 153$
- $x = 28 \rightarrow \hat{y} \approx 159$
- $x = 25 \rightarrow \hat{y} \approx 168$

1

residualerna blir då:

- $x = 40: 110 - 124 = -14$
- $x = 35: 145 - 139 = 6$
- $x = 35: 140 - 139 = 1$
- $x = 32: 160 - 147 = 13$
- $x = 30: 160 - 153 = 7$
- $x = 28: 162 - 159 = 3$
- $x = 25: 150 - 168 = -18$
- $x = 30$ $155 - 153 = 2$



1

Som man kan se i spridningsdiagrammet uppvisar residualerna ett mönster som liknar en kurva med en maximipunkt. $\cap$
En annan typ av linje hade därför varit bättre för att beskriva sambandet mellan pris i kr och kvantitet i liter

1

Universitetstryckeriet

Behandla endast en uppgift per sida / *Answer only one question per page*

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)

STGA01-0145-TAK

Löpande sidnr / Consecutive no: 3

Uppgift nr / Question no: 2

Poäng / Points awarded: 10

Lärarens anteckning / Examiner's remarks

Uppgift 2
För att besvara frågorna i uppgift 2 har jag gjort en korstabell utifrån informationen som jag har fått. Jag antar 1000 lärare.

	B	B^c	
A	160	480	640
A^c	50	310	360
Σ	210	790	1000

a1) $P(A) = 0.64$ — 2

a2) $P(B|A) = 0.25$

a3) $P(B) = 0.21$

b) $P(A^c \text{ och } B) = \frac{50}{1000} = 0.05$ — 2

c) $P(B|A^c) = \frac{50}{360} \approx 0.1389$ — 2

d) $P(A^c|B) = \frac{50}{210} \approx 0.2381$ — 2

e) för att avgöra om A och B är oberoende kan man kolla om $P(B|A)$ och $P(B|A^c)$ är olika. I detta fall är $P(B|A) = 0.25$ och $P(B|A^c) \approx 0.1389$, alltså är de olika vilket betyder att A och B inte är oberoende. — 2

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)

STGA01-0145-TAK

Löpande sidnr / *Consecutive no:* 4

Uppgift nr / *Question no:* 3

Poäng / *Points awarded:* 9,5

Lärarens anteckning / *Examiner's remarks:*

Häftområde

Skriv ej i detta område / Leave this area blank

Uppgift 3

a) sannolikheten att en anställd är positiv $= \frac{900}{2000} = 0.45 = 45\%$.

Jag kommer anta att det slumpmässiga urvalet av de 10 personerna är oberoende av tidigare försök och att sannolikheten att den anställda är positiv är densamma varje gång, vilket den kan anses vara eftersom $\frac{n}{N} < 0.1 : \frac{10}{2000} = 0.005$

Jag får då $X \sim Bin(n=10, p=0.45)$ X = ... ange

50% av 10 = 5, så den sökta sannolikheten blir $P(X \leq 5)$, vilket är detsamma som $P(0)+P(1)+P(2)+P(3)+P(4)+P(5)$:

$P(0) = \binom{10}{0} \cdot 0.45^0 \cdot 0.55^{10} \approx 1 \cdot 1 \cdot 0.0025 = 0.0025$
$P(1) = \binom{10}{1} \cdot 0.45^1 \cdot 0.55^9 \approx 10 \cdot 0.45 \cdot 0.0046 = 0.0207$
$P(2) = \binom{10}{2} \cdot 0.45^2 \cdot 0.55^8 \approx 45 \cdot 0.2025 \cdot 0.0084 = 0.0763$
$P(3) = \binom{10}{3} \cdot 0.45^3 \cdot 0.55^7 \approx 120 \cdot 0.0911 \cdot 0.0152 = 0.1665$
$P(4) = \binom{10}{4} \cdot 0.45^4 \cdot 0.55^6 \approx 210 \cdot 0.0410 \cdot 0.0277 = 0.2384$
$P(5) = \binom{10}{5} \cdot 0.45^5 \cdot 0.55^5 \approx 252 \cdot 0.0185 \cdot 0.0503 = 0.2340$

$P(X \leq 5) = 0.0025 + 0.0207 + 0.0763 + 0.1665 + 0.2384 + 0.2340 = 0.7384$

sannolikheten att högst 50% är positiva när man väljer 10 slumpmässiga personer är ca 73.8%.

4,5

b) Jag kommer nu räkna på andelen positiva istället för antal. Det krävs dock att stickprovsandelen är approximativt normalfördelad vilket den är om villkoren $n \cdot p \geq 15$ och $n \cdot (1-p) \geq 15$ är uppfyllda:

då $n=100$ får vi: $100 \cdot 0.45 = 45$ och $100 \cdot 0.55 = 55$. Villkoren är uppfyllda.

Utifrån informationen i uppgiften får jag $\hat{p} \sim N\left(0.45, \sqrt{\frac{0.45 \cdot 0.55}{100}}\right)$ $\hat{p}$

Den sökta sannolikheten är $P(\hat{p} \leq 0.50)$, vilket blir:

$P\left(z \leq \frac{0.50-0.45}{\sqrt{\frac{0.45 \cdot 0.55}{100}}}\right) \approx P\left(z \leq \frac{0.05}{0.0497}\right) \approx P(z \leq 1.01) \approx 0.8438$

Sannolikheten att högst 50% av 100 slumpmässigt utvalda personer är positiva är ca 84.4%.

5

Universitetstryckeriet

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)

STGA01-0145-TAK

Löpande sidnr / Consecutive no: 5

Uppgift nr / Question no: 4

Poäng / Points awarded: 9,5

Lärarens anteckning / Examiner's remarks:

Uppgift 4

a) Person nummer:

1	2	$\bar{X}$	sannolikhet
→37	37	37	0.09 (0.30·0.30)
→37	43	40	0.15 (0.30·0.50)
→37	49	43	0.06 (0.30·0.20)
→43	37	40	0.15 (0.30·0.50)
→43	43	43	0.25 (0.50·0.50)
43	49	46	0.10 (0.50·0.20)
→49	37	43	0.06 (0.30·0.20)
49	43	46	0.10 (0.50·0.20)
49	49	49	0.04 (0.20·0.20)

Tabellen visar de möjliga utfallen när man väljer 2 slumpmässiga personer och sannolikheterna för dessa.

Pilarna pekar på de utfall där medellönen blir mindre än 44tkr. för att beräkna sannolikheten att medellönen blir högst 44 tkr adderar jag sannolikheterna för dessa utfall och får då:

$P(\bar{X} \leq 44) = 0.09 + 0.15 + 0.06 + 0.15 + 0.25 + 0.06 = 0.76$

Sannolikheten att medellönen hos 2 slumpvis utvalda personer är högst 44 tkr är 76%.

4

b) Eftersom n=40 kan jag tillämpa CGS som lovar att om stickprovsstorleken är tillräckligt stor (minst 30) blir samplingfördelningen för stickprovsmedelvärdet approximativt normalfördelad.

Jag får då $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

→ $\mu = \sum x P(x) = 37 \cdot 0.30 + 43 \cdot 0.50 + 49 \cdot 0.20 = 11.1 + 21.5 + 9.8 = 42.4$

→ $\sigma^2 = \sum (x-\mu)^2 P(x) = (37-42.4)^2 \cdot 0.30 + (43-42.4)^2 \cdot 0.50 + (49-42.4)^2 \cdot 0.20$
$= 8.748 + 0.18 + 8.712 = 17.784$ ✓

→ $\sigma = \sqrt{\sigma^2} = \sqrt{17.784} \approx 4.2171$

$\bar{X} \sim N\left(42.4, \frac{4.2171}{\sqrt{40}}\right)$

Den sökta sannolikheten är $P(\bar{X} \leq 44) = P\left(Z \leq \frac{44-42.4}{(4.2171/\sqrt{40})}\right) \approx P\left(Z \leq \frac{1.6}{0.6668}\right)$

$\approx P(Z \leq 2.40) \approx 0.9918$

Sannolikheten att medellönen hos 40 slumpvis utvalda personer är högst 44 tkr är ca 99.2%.

5.5

Universitetstryckeriet

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)

STGA01-0145-TAK

Löpande sidnr / *Consecutive no:* 6

Uppgift nr / *Question no:* 5

Poäng / *Points awarded:* 5

Lärarens anteckning / *Examiner's remarks:*

Uppgift 5

a1) $n=6$

$\bar{x} = (178+177+174+168+181+172)/6 = 175$

$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{(178-175)^2+(177-175)^2+(174-175)^2+(168-175)^2+(181-175)^2+(172-175)^2}{6-1}}$

$= \sqrt{\frac{9+4+1+49+36+9}{5}} = \sqrt{\frac{108}{5}} = \sqrt{21.6} \approx 4.6476$

$\rightarrow se = \frac{4.6476}{\sqrt{6}}$

antaganden:
- längd är en kvantitativ variabel
- slumpmässigt urval av de 6 männen
- stickprovsmedellängden är normalfördelad eftersom populationen är normalfördelad.

Då $n=6$ får jag 5 frihetsgrader vilket innebär att $t=2.57$
Jag kan nu beräkna konfidensintervallet:
- nedre gräns: $175 - 2.57 \cdot \frac{4.6476}{\sqrt{6}} \approx 170.12$
- övre gräns: $175 + 2.57 \cdot \frac{4.6476}{\sqrt{6}} \approx 179.88$

Detta ger konfidensintervallet [170.12, 179.88]

Med 95% säkerhet finns den verkliga medellängden bland alla män på Maldiverna mellan 170.12 cm och 179.88 cm. 3,5

a2) Eftersom 181 inte är med i konfidensintervallet kan man med 95% säkerhet våga påstå att medellängden för vuxna är lägre på Maldiverna än i Sverige. 1

b1) antaganden:
- kvantitativ variabel i 2 grupper (Maldiverna och Thailand)
- oberoende slumpmässiga urval i de 2 grupperna
- sampling fördelningen för stickprovsmedelvärdena är approximativt normalfördelade eftersom populationerna är normalfördelade.

Datan för Thailand ger:
$n=6$
$\bar{x} = (176+162+165+156+158+173)/6 = 165$
$s = \sqrt{\frac{121+9+0+81+49+64}{6-1}} = \sqrt{\frac{324}{5}} = \sqrt{64.8} \approx 8.0498$

Jag kommer kalla Maldivernas läges- och spridningsmått för $\bar{x}_M$ och s_M
Thailands mått kommer jag kalla $\bar{x}_T$ och s_T

antal frihetsgrader blir $df = n_M + n_T - 2 = 6+6-2 = 10$, så $t = 2.23$

(fortsättning nästa sida)

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)

STGA01-0145-TAK

Löpande sidnr
Consecutive no: 7

Uppgift nr / *Question no:* 5

Poäng / *Points awarded:* 10

Lärarens anteckning
Examiner´s remarks:

b1) forts.

$$se = \sqrt{\frac{s_M^2}{n_M} + \frac{s_T^2}{n_T}} = \sqrt{\frac{4.6476^2}{6} + \frac{8.0498^2}{6}} \approx \sqrt{14.4} \approx 3.7947$$

Konfidensintervallet för $\mu_M - \mu_T$ blir:
- nedre gräns: $(175-165) - 2.23 \cdot 3.7947 \approx 1.54$
- övre gräns: $(175-165) + 2.23 \cdot 3.7947 \approx 18.46$

Med 95% säkerhet finns skillnaden mellan populationsmedelvärdena på Maldiverna och i Thailand i intervallet [1.54, 18.46]   4,5

b2) Eftersom talet noll inte finns med i intervallet kan man med 95% säkerhet våga påstå att det är skillnad mellan medellängderna i de två länderna.   1

Man kan dessutom konstatera (med 95% säkerhet) att männen i Thailand är kortare än männen på Maldiverna eftersom det endast är positiva tal i intervallet.

Universitetstryckeriet

Uppgift 6.

a) 1. antaganden: 0,5
- kategorisk variabel (kollar på programmet/kollar inte)
- slumpmässigt urval av de 200 personerna 0,5
- tillräckligt stort stickprov för att normalapproximation ska gälla:
  → $200 \cdot 0.20 = 40$ och $200 \cdot 0.80 = 160$. Båda är större än 15 så villkoren är uppfyllda. 0,5

2. hypoteser:
- $H_0: p = 0.20$ 1
- $H_a: p < 0.20$

3. teststatistika:
- $se_0 = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{0.20 \cdot 0.80}{200}} \approx 0.02828$
- $\hat{p} = \frac{29}{200} = 0.145$ 0,5
- $z = \frac{\hat{p} - p_0}{se_0} = \frac{0.145 - 0.20}{0.02828} \approx -1.94$ 1

4. P-värde: 0,5
- $P(z < -1.94) \approx 0.0262$ 0,5

5. Slutsats: 0,5
- eftersom P-värdet $< \alpha$ (0.05) kan vi förkasta H_0 på 5% signifikansnivå. Med andra ord vågar vi påstå att andelen som kollar på Harrys TV-program bland befolkningen är lägre än 20% på 5% signifikansnivå 0,5

b) Typ II-fel innebär att man accepterar H_0 fast H_0 är falsk. 0,5
Gränsen för att förkasta H_0 är där $P(Z < z) = 0.05$. Detta sker enligt normalfördelningstabellen när $z = -1.645$.

Gränsen ligger därför -1.645 standardavvikelser från p_0. Vi får då:
→ $0.20 - 1.645 \cdot 0.02828 \approx 0.1535$ 0,5

0,5
Detta innebär att om $\hat{p} < 0.1535$ ska man förkasta H_0 men om $\hat{p} \geq 0.1535$ ska man "acceptera" H_0. I detta fall är $\hat{p} = 0.12$

Vi får då $\hat{P} \sim N\left(0.12, \sqrt{\frac{0.12 \cdot 0.88}{200}}\right)$ 1

0,5
Den sökta sannolikheten blir: $P(\hat{p} \geq 0.1535) = P\left(Z \geq \frac{0.1535 - 0.12}{\sqrt{\frac{0.12 \cdot 0.88}{200}}}\right)$

0,5
$\approx P(Z \geq 1.46) \approx 1 - 0.9279 = 0.0721$ 0,5

Sannolikheten att göra ett typ II-fel om den verkliga andelen är 12% är ca 7.2%.