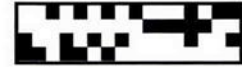




# TENTAMEN / EXAMINATION



8164617

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 0 9 3 - S Y Z
Provbetsbenämning / Exam name		Oonmald	
Statistik			
Kurskod / Course code	Provkod / Exam code	Tentamensdatum / Examination date	
S T G A 0 1	1 0 0 2	2 0 2 0 - 0 1 - 1 4	
Jag har tagit del av regler som gäller i tentamenssalen / I have read the current exam hall rules		Antal inlämnade blad / Number of sheets	
<input checked="" type="checkbox"/> Ja / Yes		1 2	

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

Kontroll av legitimation / Identification checked	<input checked="" type="checkbox"/> Ja / Yes	Härmed intygas att ovanstående kontroller utförts / This is to certify that the above mentioned checks have been carried out
Kontroll av inlämnade blad / Answer sheets checked	<input checked="" type="checkbox"/> Ja / Yes	
Inlämningstid / Time of submission	1 7 : 1 5	Tydlig sign. / Signature

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	~
10	9,5	10	9	9,5	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				Examin. lärare / Kursansvarig signatur / Signature of the examiner						
58				A.W						
Betyg / Grade				Namnförtydligande / Clarification of the signature						
VG										

8164617

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



$$1a) \bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{0 \cdot 1 + 1 \cdot 4 + 2 \cdot 4 + 3 \cdot 3 + 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 2 + 7 \cdot 1 + 8 \cdot 1}{1 + 4 + 4 + 3 + 1 + 1 + 2 + 1 + 1}$$

$$\bar{x} = \frac{0 + 4 + 8 + 9 + 4 + 5 + 12 + 7 + 8}{18} = \frac{57}{18} = 3,1667$$

Svar:  $\bar{x} = 3,1667$  medelvärdet för män är 3,1667

$$1b) \bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{0 \cdot 2 + 1 \cdot 4 + 2 \cdot 4 + 3 \cdot 5 + 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 1 + 7 \cdot 1 + 8 \cdot 1}{2 + 4 + 4 + 5 + 1 + 1 + 1 + 1 + 1}$$

$$\bar{x} = \frac{0 + 4 + 8 + 15 + 4 + 5 + 6 + 7 + 8}{20} = \frac{57}{20} = 2,85$$

Svar:  $\bar{x} = 2,85$  medelvärdet för kvinnorna är 2,85

$$1c) \bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{0 \cdot 3 + 1 \cdot 8 + 2 \cdot 8 + 3 \cdot 8 + 4 \cdot 2 + 5 \cdot 2 + 6 \cdot 3 + 7 \cdot 2 + 8 \cdot 2}{18 + 20}$$

$$\bar{x} = \frac{0 + 8 + 16 + 24 + 8 + 10 + 18 + 14 + 16}{38} = \frac{114}{38} = 3$$

Svar:  $\bar{x} = 3$  medelvärdet för alla personer är 3

1d) Män:

0, 1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 5, 6, 6, 7, 8

$\uparrow$   $\uparrow$   $\uparrow$   
 Q1      md      Q3

$$Q1 = 1$$

$$md = (2+3)/2 = 2,5$$

$$Q3 = 5$$

$$Q3 + 1,5 IQR = 5 + 6 = 11$$

$$Q1 - 1,5 IQR = 1 - 6 = -5$$

$$IQR = Q3 - Q1 \quad IQR = 5 - 1 = 4$$

→ inga outliers

$$\text{max mörrhår} = 4,5 IQR \quad 1,5 \cdot 4 = 6$$



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)

ST6A01-0093-SYZ

Löpande sidnr  
Consecutive no:

2

Uppgift nr /  
Question no: 1

Poäng / Points  
awarded:

10

Lärens  
anteckning  
Examiner's remarks:

Kvinnor:

0, 0, 1, 1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 3, 4, 5, 6, 7, 8

Q1

md

Q3

$$Q1 = (1+1)/2 = 1$$

$$md = (2+3)/2 = 2,5$$

$$Q3 = (3+4)/2 = 3,5$$

$$Q3 + 1,5 IQR = 3,5 + 3,75 = 7,25$$

$$Q1 - 1,5 IQR = 1 - 3,75 = -2,75$$

$$IQR = Q3 - Q1 \quad IQR = 3,5 - 1 = 2,5$$

→ 8 är outlier pga

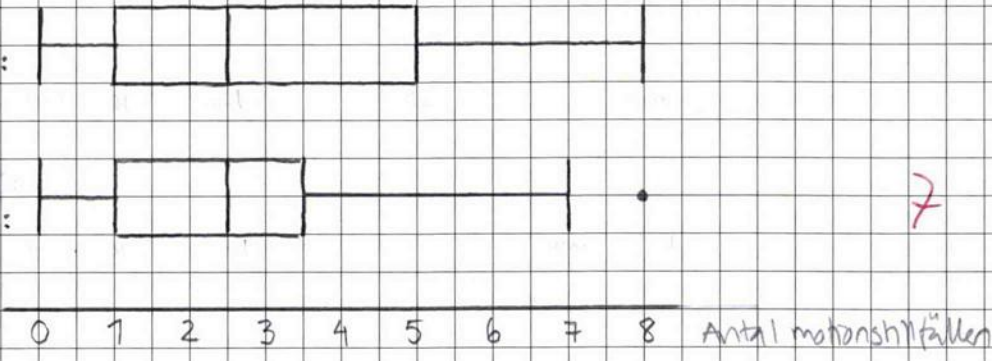
$$\text{max morrhår} = 1,5 IQR \quad 1,5 \cdot 2,5 = 3,75$$

$$8 > 7,25$$

Kön:

Män:

Kvinnor:



7



2) A: Decimalfel

2a)  $P(A^c \text{ och } B^c) = 61/100 = 0,61$  2

B: Datumfel

2b)  $P(B|A) = 4/18 = 0,222$  2

2c)  $P(A|B) = 4/25 = 0,16$  2

2d)  $P(A^c|B^c) = 61/75 = 0,813$  2

	A	A <sup>c</sup>	
B	4	21	25
B <sup>c</sup>	14	61	75
	18	82	100

2e) Om  $P(B|A) = P(B|A^c)$  är variablerna oberoende, dvs om det är lika stor sannolikhet att en fil har datumfel givet att den har decimalfel som att en fil har datumfel givet att den inte har decimalfel innebär det att de två feltyperna är oberoende.

$$P(B|A) = 0,18 \quad (\text{se 2b}) - 0,222$$

$$P(B|A^c) = 21/82 = 0,256$$

$0,18 \neq 0,256 \rightarrow$  de två feltyperna är inte oberoende

Uppgift nr /  
Question no: 2Poäng / Points  
awarded:

9,5

Lärens  
anteckning  
Examiner's remarks:



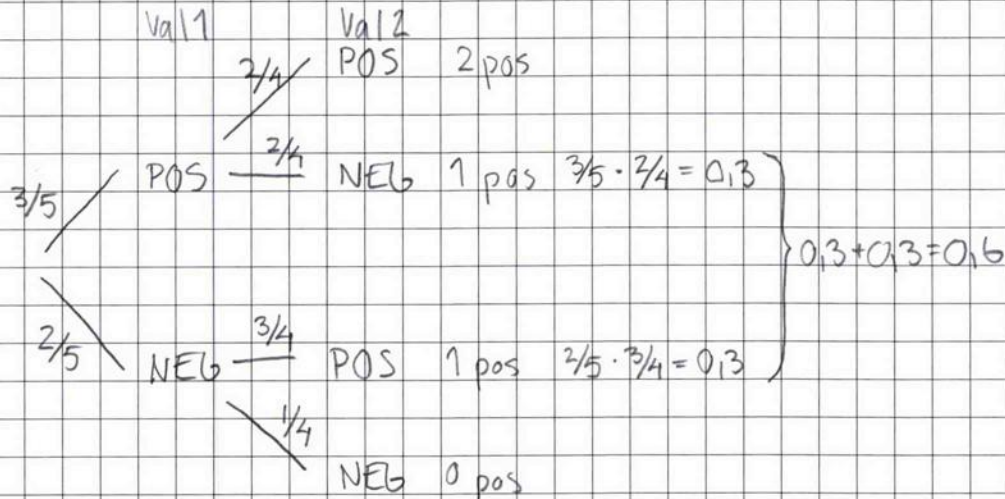

 Skriv ej i detta område  
 Leave this area blank

 3a)  $N = 5$   $n/N < 0,10$  → sannolikheten anses samma vid varje val  
 $n = 2$ 

$$2/5 = 0,20 \quad 0,20 > 0,10 \rightarrow \text{valen har}$$

3/5 positiva

inte lika sannolikhet


 Svar: Sannolikheten att välja en positiv person ur skapelsen vid  
 val av 2 är 0,60, dvs 60%

 3b)  $N = 2000$   $n/N < 0,10$  → sannolikheten anses samma vid  
 $n = 10$  varje slumpmässigt val

$$10/2000 = 0,005 \quad 0,005 < 0,10 \rightarrow \text{valen anses oberoende}$$

$$p = 900/2000 = 0,45$$

 $X$ : antal positiva personer

$$X \sim \text{Bin}(n=10, p=0,45)$$

$$P(X) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \quad 1-p = 1-0,45 = 0,55$$

$$P(0) = \binom{10}{0} 0,45^0 \cdot 0,55^{10} \approx 1 \cdot 1 \cdot 0,00253 \approx 0,00253$$

$$P(1) = \binom{10}{1} 0,45^1 \cdot 0,55^9 \approx 10 \cdot 0,45 \cdot 0,00461 \approx 0,02072$$

$$P(2) = \binom{10}{2} 0,45^2 \cdot 0,55^8 \approx 45 \cdot 0,2025 \cdot 0,00837 \approx 0,07630$$

→

 Uppgift nr /  
 Question no: 3

 Poäng / Points  
 awarded:

 Lärarens  
 anteckning  
 Examiner's remarks:

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)

ST6A01-0093-SYZ

Löpande sidnr  
Consecutive no:

5

Uppgift nr /  
Question no: 3

Poäng / Points  
awarded: →

Lärens  
anteckning  
Examiner's remarks:

$$P(3) = \binom{10}{3} 0,45^3 \cdot 0,55^7 \approx 120 \cdot 0,091125 \cdot 0,01522 \approx 0,16648$$

$$P(4) = \binom{10}{4} 0,45^4 \cdot 0,55^6 \approx 210 \cdot 0,04101 \cdot 0,02768 \approx 0,23837$$

$$P(5) = \binom{10}{5} 0,45^5 \cdot 0,55^5 \approx 252 \cdot 0,01845 \cdot 0,05033 \approx 0,23403$$

$$P(X \leq 5) = P(0) + P(1) + P(2) + P(3) + P(4) + P(5) = 0,73843$$

Svar: sannolikheten att få högst fem positiva från de  
anställda vid val av 10 slumpmässiga är 73,84%

$$(P(X \leq 5) = 0,73843)$$

3c)  $N = 2000$   $n/N = 100/2000 = 0,05$   $0,05 \leq 0,10$   
 $n = 100$  → oberoende val

$$p = 0,45$$

$$np > 15$$

$$n(1-p) > 15$$

} krav för normalapproximation

$$100 \cdot 0,45 = 45 \quad 45 > 15$$

ok!

$$100 \cdot 0,55 = 55 \quad 55 > 15$$

Antaganden:

- \* Slumpmässigt urval
- \* Oberoende försök  $n/N < 0,10$
- \* Villkor för normalapproximation uppfyllda:

$$np > 15 \quad n(1-p) > 15$$





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

$$n = 100$$

$$p = 0,45$$

X: antalet positiva

$$\mu = np \quad \mu = 100 \cdot 0,45 = 45$$

$$\sigma = \sqrt{np \cdot (1-p)} \quad \sigma = \sqrt{100 \cdot 0,45 \cdot (1-0,45)} = 4,9749$$

$$X \sim N(45, 4.9749)$$

$$Z = \frac{X - \mu}{\sigma} \quad P(X \leq 50) = P\left(Z \leq \frac{50 - 45}{4,9749}\right) = P(Z \leq 1,01)$$

$$P(Z \leq 1,01) = 0,8438$$

Svar: Sannolikheten att välja högst 50 positiva personer från de anställda vid val av 100 slumpmässiga är 84,38%

$$(P(X \leq 50) = 0,8438)$$

4,5



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)

ST6A01-0093-SYZ

Löpande sidnr  
 Consecutive no:

7

Häftområde

Skriv ej i detta område  
 Leave this area blank

4 a) utfall |  $\bar{x}$  | sannolikhet

1 + 1	1	
1 + 2	1,5	
1 + 3	2	
1 + 4	2,5	
2 + 1	1,5	
2 + 2	2	
2 + 3	2,5	
2 + 4	3	
3 + 1	2	
3 + 2	2,5	
3 + 3	3	
3 + 4	3,5	$0,25 \cdot 0,25 = 0,0625$
4 + 1	2,5	
4 + 2	3	
4 + 3	3,5	$0,25 \cdot 0,25 = 0,0625$
4 + 4	4	$0,25 \cdot 0,25 = 0,0625$

$0,0625 \cdot 3 = 0,1875$

Svar: sannolikheten att få ett medelresultat högre än 3 är 18,75%

(Alternativ uträkning: I 3 av 16 fall är medelresultatet över 3,  $3/16 = 0,1875$  18,75%)

Uppgift nr /  
 Question no: 4

Poäng / Points  
 awarded:

4

Lärarens  
 anteckning  
 Examiner's remarks:

4





4b)

$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \quad n = 30$$

$$\mu = \sum x P(x) \quad \mu = 1 \cdot 0,25 + 2 \cdot 0,25 + 3 \cdot 0,25 + 4 \cdot 0,25 = 2,5$$

$$\sigma^2 = \sum (x - \mu)^2 P(x)$$

x	P(x)	(x - μ)	(x - μ) <sup>2</sup>	(x - μ) <sup>2</sup> P(x)
1	0,25	-1,5	2,25	0,5625
2	0,25	-0,5	0,25	0,0625
3	0,25	0,5	0,25	0,0625
4	0,25	1,5	2,25	0,5625

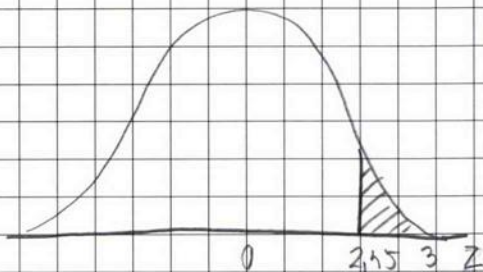
$$\Sigma = 1,25$$

$$\sigma^2 = 1,25$$

$$\sigma = \sqrt{\sigma^2} \quad \sigma = \sqrt{1,25} = 1,1180$$

$$\bar{x} \sim N\left(2,5; \frac{1,1180}{\sqrt{30}}\right) = \bar{x} \sim N(2,5; 0,2041)$$

$$P(\bar{x} > 3) = P\left(Z > \frac{3 - 2,5}{0,2041}\right) = P(Z > 2,45)$$



$$P(Z > 2,45) = 1 - 0,9929 = 0,0071$$

Svar: sannolikheten att  
medelresultatet blir  
högre än 3 när  
man skurrar skivan  
30 gånger är 0,71%

 Uppgift nr /  
 Question no: 4

 Poäng / Points  
 awarded: 9

 Lärarens  
 anteckning  
 Examiner's remarks:



5a)

	P	N	
V	124	76	200
Ö	78	82	160
	202	158	360

Antaganden

Två kategoriska variabler (plats/inställning)

Slumpmässigt urval

$$n_1 \hat{p}_1 > 10 \quad n_1 (1 - \hat{p}_1) > 10$$

$$n_2 \hat{p}_2 > 10 \quad n_2 (1 - \hat{p}_2) > 10$$

ok! se nedan:

$$V: 1 \quad n_1 = 200 \quad \hat{p}_1 = 124/200 = 0,62 \quad 200 \cdot 0,62 = 124 \quad 200 \cdot 0,38 = 76$$

$$Ö: 2 \quad n_2 = 160 \quad \hat{p}_2 = 78/160 = 0,4875 \quad 160 \cdot 0,4875 = 78 \quad 160 \cdot 0,5125 = 82$$

$$(\hat{p}_1 - \hat{p}_2) \pm z \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$z = 1,96 \text{ (av 95\% KI)}$$

$$(0,62 - 0,4875) \pm 1,96 \cdot \sqrt{\frac{0,62(1-0,62)}{200} + \frac{0,4875(1-0,4875)}{160}}$$

= [0,0299; 0,2351] Med 95% säkerhet finns skillnaden mellan populationsandelarna positiv i Västerby jämfört med Österby mellan +0,0299 och +0,2351.

(2,99%) (23,51%)

$$5b) \quad z = 2,576 \text{ (av 99\% KI)}$$

$$(0,62 - 0,4875) \pm 2,576 \cdot \sqrt{\frac{0,62(1-0,62)}{200} + \frac{0,4875(1-0,4875)}{160}}$$

$$= [-0,0023; 0,2673]$$

Med 99% säkerhet finns skillnaden mellan populationsandelarna positiv i Västerby jämfört med Österby mellan -0,0023 och +0,2673 (-0,23% o +26,73%)

Uppgift nr /  
Question no: 5Poäng / 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)

ST6A01-0093-SYZ

Löpande sidnr  
 Consecutive no:

10

Häftområde

Skriv ej i detta område  
 Leave this area blank

5c) Eftersom svaret i 5a innehöll endast positiva tal kan vi vara 95% säkra på att det finns skillnad mellan populationsandelarna åt vilket håll?

I svar 5b var ett av resultaten ett negativt värde och det andra var positivt, vilket innebär att 0 finns med i intervallet, det gör att vi inte med 99% säkerhet kan säga om det finns skillnad mellan populationsandelarna.

Svar: vi kan vara 95% säkra på att det finns skillnad, men inte 99% säkra.

Uppgift nr /  
 Question no: 5

Poäng / Points  
 awarded: 9.5

Lärarens  
 anteckning  
 Examiner's remarks:



b a)  $\sigma = 3500$   
 $n = 50$   
 $\bar{x} = 23900$

Uppgift nr /  
Question no: 6

Poäng / Points  
awarded: 10

Lärarens  
anteckning  
Examiner's remarks:

HYPOTESTEST

1. Antaganden

- Kvantitativ variabeltyp (Lön)
- Slumpmässigt urval
- $n > 30$  ok enl CGS

2. Hypoteser

$H_0: \mu = 25000$

$H_a: \mu < 25000$

3. Teststatistika

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$
 (0,5)  
( $\sigma$  känt)

$$Z = \frac{23900 - 25000}{3500 / \sqrt{50}} = -2,22$$
 (0,5)

4. P-värde

$P(\mu < 25000) = P(Z < -2,22) = 0,0136$  (0,5)

5. Slutats

P-värdet 0,0136, 1,36% är mindre än 5%, alltså förkastas  $H_0$ . (0,5)

1 ord: på 5% signifikansnivå ger data tillräckligt stöd för att vi ska kunna påstå att medellönen är lägre än 25000kr. (0,5)





b b) 1. Beräkna gränsen för att förkasta  $H_0$  om  $\mu=25000$

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \rightarrow \bar{x} = Z \cdot (\sigma/\sqrt{n}) + \mu_0$$

0,5 (indirect)

$Z = -1,645$  (från 5% signifikansnivå)

$\mu_0 = 25000$

$\sigma = 3500$

$n = 50$

$$\bar{x} = -1,645 \cdot (3500/\sqrt{50}) + 25000$$

$$\bar{x} = 24185,77$$

Gränsen för att förkasta  $H_0$  om  $\mu=25000$  går vid  $\bar{x} < 24185,77$

2. Beräkna sannolikheten att få  $\bar{x} > 24185,77$  om det korrekta  $\mu = 23700$

0,5

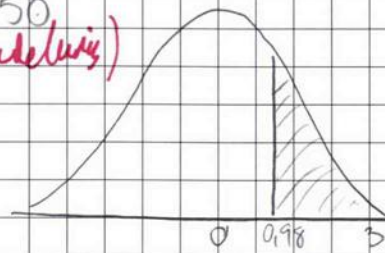
$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$\bar{x} = 24185,77 \quad \sigma = 3500$$

$$\mu = 23700 \quad n = 50$$

0,5 (Indirect korrekt Andelenis)

$$Z = \frac{24185,77 - 23700}{3500/\sqrt{50}} = 0,98$$



$$P(\bar{x} > 24185,77) = P(Z > 0,98) = 1 - 0,8365 = 0,1635$$

0,5

Svar: Sannolikheten att göra ett typ II-fel i uppgift a är 16,35% (0,1635)

b c) Sannolikheten att jag upptäcker att Harry har lygnit är  $1 - \beta$ ,  $1 - 0,1635 = 0,8365$ , dvs 83,65%

1,5