



# TENTAMEN / EXAMINATION



12307683

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

Skriv anonymiseringskoden på samtliga svarsblad / <i>Write your anonymity code on each sheet</i>		Anonymiseringskod / <i>Anonymity code</i> N E G C 4 7 - 0 0 0 6 - R D O	
Provbeneämning / <i>Exam name</i> Tentamen Undersökningsmetodik			Onämld <input type="text"/> <input type="text"/> <input type="text"/>
Kurskod / <i>Course code</i> N E G C 4 7	Modul / <i>Module</i> 2 0 1 0	Tentamensdatum / <i>Examination date</i> 2 0 2 1 - 0 1 - 1 4	
Jag har tagit del av regler som gäller vid tentamen / <i>I have read the current rules for examinations</i>		<input type="checkbox"/> Ja / <i>Yes</i>	Antal inlämnade blad med anonymiseringskod / <i>Number of sheets with anonymity code</i> <input type="text"/> <input type="text"/>

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

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

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

Bedömning av uppgifter / <i>Questions attempted</i>										
1	2	3	4	5	6	7	8	9	10	~
10	3,5	8,5	7	9	4					
11	12	13	14	15	16	17	18	19	20	~
21	22	23	24	25	26	27	28	29	30	~
Totalt antal poäng / <i>Total points</i> -- 42					Examin. lärare / <i>Kursansvarig signatur</i> / <i>Signature of the examiner</i>					
Betyg / <i>Grade</i> 9					Namnförtydligande / <i>Clarification of the signature</i>					

12307683



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

## Fråga 1

1 a) Man gör en kvalitetsdeklaration för att kunna bedöma värdet av informationen i undersökningen. Man gör användarna uppmärksamma på fel som kan göra informationen mindre användbar i ger bättre beslutsunderlag och gör det lättare för användaren att kunna precisera kvalitetskrav inför nya undersökningar.

Man gör alltså användaren uppmärksam på i vilka situationer materialet i undersökningen är tillräckligt som beslutsunderlag och när det inte är tillräckligt. Kvalitetsdeklarationen gör att man minskar risken att man som användare "litar blint" på informationen för att den ser bra ut.

1 PSU-rapporten ingår komponenterna:

- Innehåll
- Tillförlitlighet
- Aktualitet
- Tillgänglighet
- Jämförbarhet

4p

1 b) De som hade varit röstberättigade om det hade varit riksdagsval andra söndagen i september aktuellt undersökningsår och som är bokförda i Sverige.

1p

1 c) Registret över totalbefolkning (RTB)

Registret innehåller aktuell information eftersom det uppdateras dagligen

→

Fråga 1

- 1 c) Ramen innehåller värden på bakgrunds- och hjälpvariabler (demografiska variabler som kön, ålder, boendeort etc.) 2 p
- 1 d) Bortfall och mätning anses ha störst betydelse för resultatet.  
( )  
svar helt eller delvis saknas för vissa personer      frågor missförstås eller är svarbesvarade p.g.a. minnesproblemakti 1 p
- 1 e) Antingen genom datorstödda telefonintervjuer eller genom webbenkät. 1 p
- 1 f) Eftersom undersökningen genomförts med huvudsakligen samma metod och samma huvudsakliga definitioner sedan 1972 (utom nov 1981 t.o.m. nov. 1983) är statistikerna jämförbara över tid. 1 p

## Fråga 2

- 2 a) Obundet slumpmässigt urval (OSU) är teoretiskt enkelt men kan ge dålig precision, stratifierat urval kan ge bättre precision än vad OSU gör och är bra vid sneda urval. Om man använder systematiskt urval kan man få en hög precision om man använder en lämpligt sorterad ram. Grupperat urval är enkelt och kräver ingen ram men ger ofta sämre precision än vad OSU gör.

När det kommer till hur stickprovet fördelas mellan strata (allokering) får man större precisionförbättring ju mer stratummedelvärdena skiljer sig åt och avviker från det totala medelvärdet vid proportionellt stratifierat urval i jämförelse med OSU med återläggning.

Neymanallokering kallas också optimal allokering och ger högst precision men den är svår att tillämpa praktiskt eftersom man då måste ha kännedom om standardavvikelserna innan allokeringen görs. 3,5 P

- 2 b) Genom att öka stickprovsstorleken ökar precisionen. 0 P

Fråga 3

Hushåll	A	B	C	
1	52	51	28	B1
2	52	40	48	140
3	63	45	57	165
4	52	47	44	143
$\bar{y}_a$	219	183	177	579

3a) Randomiserat blockförsök <sup>1 2 3</sup>  
 Modell:  $Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$   $\begin{cases} i = A, B, C & \text{diskmedel} \\ j = 1, 2, 3, 4 & \text{hushåll} \end{cases}$   
 Antaganden:  $\epsilon_{ij}$  oberoende, NF, homoskedastiska

$H_0: \tau_A = \tau_B = \tau_C = 0$

$H_1: \text{Något } \tau_i \neq 0$

$\alpha = 0,05 \quad F_0 = \frac{MS_{Tr}}{MSE}$

$SS_T = \sum \sum y_{ij}^2 - \frac{1}{N} y_{..}^2 = 28789 - \frac{1}{12} \cdot 579^2 = 852,25$   
 $\uparrow$   
 $52^2 + 52^2 + 63^2 + 52^2 + 51^2 + \dots + 44^2$

$SS_{Tr} = \frac{\sum y_i^2}{b} - \frac{1}{N} y_{..}^2 = \frac{219^2 + 183^2 + 177^2}{4} - \frac{1}{12} \cdot 579^2 = 258$

$SS_{e1} = \frac{\sum y_{-j}^2}{a} - \frac{1}{N} y_{..}^2 = \frac{131^2 + 140^2 + 165^2 + 143^2}{3} - \frac{1}{12} \cdot 579^2 = 208,25$

$SS_E = SS_T - SS_{Tr} - SS_{e1} = 852,25 - 258 - 208,25 = 386$

Variationskälla	SS	df	MS	$f_0$
Tr	258	2	$258/2 = 129$	$129/64,3 \approx 2,01$
B1	208,25	3	$208,25/3 \approx 69,42$	$69,42/64,3 \approx 1,08$
E	386	6	$386/6 \approx 64,3$	
T	852,25	11		

$H_0$  förkastas om  $f_0 > f_{0,05, 2, 6} = 5,14$

$2,01 < 5,14$  så  $H_0$  kan ej förkastas.

På 5% signifikansnivå kan vi påstå att det inte finns någon skillnad i medelvärde mellan diskmedlen.

## Fråga 3

- 36) I diagram 9 tycker jag att residualerna följer linjen väl, möjligtvis ligger de något högt i början och något lågt i slutet, men jag skulle ändå påstå att det ser ut som att de är normalfördelade. I tabell 2 testas normalfördelningen, signifikansnivån är högre än 0,05, vilket innebär att nollhypotesen inte kan förkastas → residualerna är på 5% sig nivå normalfördelade.

Antagandet om normalfördelning är alltså uppfyllt.

I diagram 5 kan vi se att Diskmedel 1 verkar ha lägst varians (spridning), liksom i boxplotten (3), det tyder på att variansen inte är lika. I tabell 10 testas detta och sig.  $0,364 > 0,05$  vilket innebär att  $H_0$  inte kan förkastas. Variansen beror inte på värdena på de oberoende variablerna → variansen är olik

✓  $H_0$ : Lika varians kan ej förkastas.

Oberoende?

1,5P

## Fråga 4

4 a)  $H_0: \mu_i = \mu_j \quad \forall i, j, i \neq j$

$H_1: \mu_i \neq \mu_j$

$\alpha = 0,05$

Teststatistiken  $|\bar{y}_i - \bar{y}_j|$

$H_0$  förkastas om  $|\bar{y}_i - \bar{y}_j| > \text{LSD}$

$$\text{LSD} = t_{0,025,16} \sqrt{\text{MSE} \left( \frac{1}{n_i} + \frac{1}{n_j} \right)} = 2,145 \cdot \sqrt{64,3 \left( \frac{1}{4} + \frac{1}{4} \right)} = 13,89$$

$\alpha/2$

i	j	$ \bar{y}_i - \bar{y}_j $	Slutsats
A	B	9	$H_0$ förkastas ej ( $9 < 13,89$ )
A	C	10,5	$H_0$ förkastas ej ( $10,5 < 13,89$ )
B	C	1,5	$H_0$ förkastas ej ( $1,5 < 13,89$ )

$\bar{y}_A = 217/4 = 54,25$

$\bar{y}_B = 183/4 = 45,75$

$\bar{y}_C = 177/4 = 44,25$

På 5% signifikansnivå finns ingen skillnad i medelvärde mellan någon av diskmedlen.

4 b)

/R

Fr

Op

Fråga 5

A	C
52	28
57	52
59	54
69	67
59	52
70	74
53	48
66	57
$\Sigma$ 485	432

5a)  $\bar{x} = \frac{1}{n} \sum y_i$

$\bar{x} = \frac{1}{8} \cdot 485 = 60,625$

$n = 8$   
 $N = 120$

$\hat{V}(\bar{y}) = \left(1 - \frac{n}{N}\right) \frac{s^2}{n}$ ,  $s^2 = \frac{1}{n-1} \left(\sum y_i^2 - \frac{1}{n} (\sum y_i)^2\right)$

$s^2 = \frac{1}{8-1} \left(29741 - \frac{1}{8} \cdot 485^2\right) = \frac{1}{7} (29741 - 29403,125)$

$s^2 = \frac{1}{7} \cdot (29741 - 29403,125) \approx 48,268$

$\hat{V}(\bar{y}) = \left(1 - \frac{8}{120}\right) \cdot \frac{48,268}{8} = 5,63$

95% KL:  $60,625 \pm 1,96 \cdot \sqrt{5,63}$

t-värde

$55,97 \leq \mu \leq 65,28$

Med 95% sannolikhet <sup>säkerhet</sup> är medelvärdet för populationen mellan 55,97 och 65,28.

Antagande: Normalfördelning

4,5P

5b)  $\Sigma_{tot} = 6283$

$N_i = 120$

$\bar{y}_C = 6283/120 \approx 52,36$

$\bar{x}_C = 432/8 = 54$

$\bar{y}_{kvot} = \frac{48,268}{54} = 52,36 = 46,802$

$\hat{V}(\bar{y}_{kvot}) = \left(1 - \frac{n}{N}\right) \frac{s_{kvot}^2}{n}$

$\checkmark \bar{y}_{kvot} = \frac{\bar{y}}{\bar{x}} \cdot \mu_x$   
 $= \frac{60,625}{54} \cdot 52,36$

$P_{kvot} = \frac{485}{432} \approx 1,12$

$s^2 = \frac{\sum (A_i - P_{kvot} \cdot C_i)^2}{n-1}$

$s^2 = \frac{(52 - 1,12 \cdot 28)^2 + (57 - 1,12 \cdot 52)^2 + \dots + (66 - 1,12 \cdot 57)^2}{8-1} = 91,13$

8-1



NEB(47-0006-RDD

Fråga 5

5 b forts)  $s_{kv}^2 = 91,13$

$$D(\bar{y}_{kv}) = \left(1 - \frac{n}{N}\right) \frac{s^2}{n} = \left(1 - \frac{8}{120}\right) \frac{91,13}{8} = 10,63$$

104

4.5 p



