

TENTAMEN I STATISTIKENS GRUNDER 2  
2019-04-25

**Skrivtid:** 15.00-20.00

**Godkända hjälpmedel:** Miniräknare, språklexikon.

Tentamen består av fem uppgifter. För full poäng på en uppgift krävs tydliga, utförliga och väl motiverade lösningar.

Vänligen använd inte röd penna på tentamen. Lycka till!

**Uppgift 1.** (20 poäng) Vid en fabrik tillverkas rep av nylon. Enligt en tidigare produktionsmetod har den förväntade hållfastheten på repen varit 7500 kg. Man vill undersöka om ett nytt produktions sätt ger ökad hållfasthet och kontrollerar därför hållfastheten för 80 repstumpar tillverkade på det nya sättet. Man antar oberoende mellan repstumparna. Medelvärde och standaravvikelsen av hållfastheten för de 80 repstumparna är:

Variable: Hållfastheten

N	Mean	Std Dev
80	7620.00	570.00

- Utför ett lämpligt test för att testa om den nya tillverknings sättet ger ökan hållfasthet på  $\alpha = 10\%$ . Tolka resultat.
- Beräkna ett tvåsidigt 95% konfidensintervall för hållfastheten. Tolka resultat.
- Vad är sannolikhet att medelvärdet hamnar i intervall  $(-\infty, 7550)$ ?
- Den stokastiska variabeln  $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$  leder till ett konfidensintervall för  $\mu$  i form  $I_\mu = (\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}})$ . Ange formel för ett konfidensintervall för  $\mu_1 - \mu_2$  som baseras på  $\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim N(0,1)$ .

**Uppgift 2.** (20 poäng) I Tabell 1 hittar du antal lägenheter per hustyp i landets två befolkningsmässigt största kommuner utanför Stockholm i 2017.

Kommun	"Hustyp"		
	Småhus	Flerbostadshus	Övriga hustyp
Göteborg	53 317	205 413	20609
Malmö	27 140	125 011	9495

Tabell 1: Källa: SCB

- Analysera om det finns skillnader mellan kommunerna vad gäller olika hustyp. Använda lämpligt hypotestest på signifikansnivå 2.5%.
- Analysera om stokastisk variabel "Hustyp" i Göteborg följer:  
 $P(\text{"Hustyp"} = \text{Småhus}) = 19\%$ ,  
 $P(\text{"Hustyp"} = \text{Flerbostadshus}) = 73.7\%$ ,  
 $P(\text{"Hustyp"} = \text{Övriga hustyp}) = 7.3\%$ .  
 Svara med lämpligt hypotestest på signifikansnivå 5%.

**Uppgift 3.** (20 poäng) Ett företag har ett stort kundregister av 2000 e-post adresser. Man vill bedöma andelen icke-aktiva e-post adresser i detta register. Man gör ett slumpmässigt urval med 50 e-post adresser för att pröva hypotes:

$H_0$ : 8% av alla e-post adresser är icke-aktiva,

$H_A$ : mer än 8% av alla e-post adresser är icke-aktiva.

- Vi fick 5 stycken icke-aktiva e-post adresser i urvalet av storlek 50. Kan man förkasta  $H_0$  på signifikansnivå 1%?
- Beräkna  $p$ -värde för a).
- Hur många e-post adresser måste minst vara icke-aktiva för att vi ska förkasta nollhypotesen på  $\alpha = 5\%$ ?

**Uppgift 4.** (25 poäng) Elnätspriser för lägenheterna i Linköping och i Norrköping ges i tabell (pris per kWh, öre):

År	2013	2014	2015	2016	2017	2018
Medelelnätspris i Linköping	64.1	66.7	68.0	70.8	73.8	76.3
Medelelnätspris i Norrköping	65.1	66.5	68.5	71.2	73.6	76.3

- Finns det någon signifikant skillnad mellan elnätspriserna i Linköping och Norrköping? Svara med lämpligt test på signifikansnivå 5%. Du får anta att elnätspriserna är normalfördelade.
- Index för elnätspris i Linköping i 2010 (med basår 2013) är 89. Vad var elnätspris i Linköping i 2010 (ge pris per kWh i öre)?
- Markus är ägare till en lägenhet i Linköping och en i Norrköping. Den årliga elförbrukningen för lägenheterna finns i tabellen:

År	2013	2014
Elanvändning i kWh (lgh i Linköping)	2000	2100
Elanvändning i kWh (lgh i Norrköping)	3300	3500

Beräkna Laspeyres index för utvecklingen av Markus elkostnader från 2013 till 2014. Tolka resultat.

- Fortsätt 3 påståenden som finns nedan.  
 Löpande priser är ...  
 KPI är ....  
 Urvall är ...

**Uppgift 5.** (15 poäng) I en företag fick beslutstagare följande beslut matris:

	Naturtillstånd			
	$E_1$	$E_2$	$E_3$	$E_4$
	$p_1 = 0.4$	$p_2 = 0.3$	$p_3 = 0.2$	$p_4 = 0.1$
$S_1$	-20	200	400	80
$S_2$	-100	100	600	200
$S_3$	300	50	150	-300

där  $S_1, S_2, S_3$  betecknar möjliga strategier.

- Om vi baserar våra beslut på givet matris är beslut tagit under säkerhet, under risk eller under osäkerhet? Välj en alternativ och motivera kort svaret.
- Vilken strategi ska man välja för att maximera den förväntade nyttan? Motivera svaret.
- Beräkna varians för nyttan vid strategi  $S_2$ .
- Vilken strategi ska man välja enligt Laplace-kriteriet? Motivera svaret.

Statistiska institutionen



Stockholms  
universitet

## Rättningsblad

**Datum:** 25/4-2019

**Sal:** Ugglevikssalen

**Tenta:** Statistikens grunder 2

**Kurs:** Statistikens grunder

**ANONYMKOD:**

0011-NRJ

Jag godkänner att min tenta får läggas ut anonymt på hemsidan som studentsvar.

**OBS! SKRIV ÄVEN PÅ BAKSIDAN AV SKRIVBLADEN**

Markera besvarade uppgifter med kryss

1	2	3	4	5	6	7	8	9	Antal inl. blad
X	X	X	X	X	X	X	X	X	4 2c
Lär.ant. 19	20	18	16	12.5					

<b>POÄNG</b> 85.5	<b>BETYG</b> B	<b>Lärarens sign.</b> Y. Prekator
----------------------	-------------------	--------------------------------------



Uppgift 7

a)  $H_0: \mu = 7500$     $H_A: \mu > 7500$

$n = 80$     $\bar{X} = 7620$     $s = 570$     $\alpha = 0,1$

$\sigma^2$  okänd    $n > 30$  Vi kan enligt CGS approximera  $N(0,1)$ -fördelning väljer  $z_{0,05} = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$

$z_{0,05} = \frac{7620 - 7500}{570/\sqrt{80}} = 1,88$     $z_{0,1} = 1,2816$

Vi förkastar  $H_0$  om  $z_{obs} > 1,2816$

Vi förkastar  $H_0$  då  $\alpha = 0,1$

dvs Vi kan påvisa att den nya metoden

ger en hållfasthet som är större än 7500kg. (den har ökat med den nya metoden)

b)  $\bar{X} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$

$7620 \pm 1,6449 \cdot \frac{570}{\sqrt{80}} = [7556,3, 7684]$

-0,5 fel  $\alpha$   
-0,5 beträkning s fel  
+63,73

Det betyder att hållfastheten med 95% säkerhet ligger mellan 7556,3 kg & 7684kg eller med andra

ord att 95% av de tillverkade Nylonrepen kommer att ha en hållfasthet mellan 7556,3 & 7684kg.

1/9

1/3



$$1c) P(\bar{X} \leq 7550) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq \frac{7550 - 7500}{570/\sqrt{80}}\right)$$

$$= P(Z \leq 0,79) = 0,78524$$

$$0,7845 \approx 0,79$$

bra /3

Svar: Det är ca 79% Sth.

$$d) \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$\bar{X} - \bar{Y} \pm z_{\alpha/2} \cdot \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}$$

bra /4

### Uppgift 2

a)

Kommun	$n_i$	Småhus	Flerbostadshus	Övriga	Total
Göteborg		53 317	205 413	20 669	279 339
Malmö		27 140	125 011	9 425	161 646
		80 457	330 24	30 104	440 985

$H_0$ : Det finns inga skillnader mellan kommunerna vad gäller husstyper

förväntade frekvenser om det råder oberoende dvs produkten av marginalstheterna.

Kommun	$E(n_i)$	Småhus	Flerbostadshus	Övriga	Total
Göteborg		50 965	209 305	19 069	279 339
Malmö		29 492	121 119	11 235	161 646
		80 457	330 424	30 104	440 985

$H_1$ : Det finns skillnader mellan kommunerna vad gäller husstyper

$$\left(\frac{80457}{440985}\right) \cdot \left(\frac{279339}{440985}\right) \cdot 440985 \approx 50965$$

$$\left(\frac{161646}{440985}\right) \cdot \left(\frac{80457}{440985}\right) \cdot 440985 \approx 29492$$

$$\chi^2 = \frac{(n_i - E(n_i))^2}{E(n_i)} \Leftrightarrow \frac{(53317 - 50965)^2}{50965} + \frac{(27140 - 29492)^2}{29492}$$

OSV  $\chi^2_{obs} \approx 8,33$

$$\chi^2(2) = 7,378 \quad V = (2-1) \cdot (3-1) = 2$$

förvänta  $H_0$  om  $\chi^2_{obs} > 7,378$   
 Vi förkastar  $H_0$  på signifikansnivå 2,5%.

Uppgift 2

$H_0$ : Variabeln "hustyp i Göteborg" följer fördelningen 0,19, 0,737 = 0,073 R

b) Vi gör om procent till antal

Kommun	Småhus	Flerbostadshus	Övriga	Total
Göteborg	53074	205873	20392	279339

$\downarrow E(n)$

Totalt antal hus i Göteborg = 279339

$279339 \cdot 0,19 \approx 53074$  osv

$\chi^2_{obs} = \frac{\sum (n_i - E(n_i))^2}{E(n_i)}$  R

$H_1$ : Variabeln "hustyp i Göteborg" följer ej fördelningen 0,19, 0,737 = 0,073 R

$\chi^2_{obs} = \frac{(53317 - 53074)^2}{53074} \dots$  R

	Småhus	Flerbostadshus	Övriga
Göteborg	53317	205413	20609

$\uparrow n$

$\chi^2_{obs} \approx 4,5$  R  
 $\downarrow$   
 $4,44958 \approx 4,5$

$V = 1 - 2 = 2$  R

förkasta  $H_0$  om

$\chi^2_{0,95}(2) = 5,991$  R

$\chi^2_{obs} > 5,991$  R

Vi kan ej förkasta  $H_0$  på signifikansnivå 0,05 dvs vi kan ej påvisa att variabeln "hustyp i Göteborg" ej följer fördelningen 0,19, 0,737 = 0,073

blm

//10p



# UPPGIFT 3

a)  $\frac{N-n}{N-1} \Rightarrow \frac{2000-50}{2000-1} = 0,975$  eftersom  $0,975 > 0,95$  behöver vi ej göra ändlighetskorrekturion R

$H_0: \pi = 0,08$      $H_A: \pi > 0,08$     ensidigt test =  $z_\alpha$  ej  $z_{\alpha/2}$

$P = \frac{5}{50} = 0,1$  eftersom  $n > 30$  så gäller CGS & vi kan approximera  $Nw(0,1)$ -fördelning  
kontrollera  $np \geq 5$   
 $n(1-p) \geq 5$

$$Z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0,1 - 0,08}{\sqrt{\frac{0,0736}{50}}} = Z_{obs} \approx 0,52$$

$z_{0,01} = 2,3263$  förkasta  $H_0$  om  $Z_{obs} > 2,3263$

Vi kan ej förkasta  $H_0$  på sign. nivå  $0,01$  /10

b) p-värde:  $P(Z > 0,52) = 1 - P(Z \leq 0,52)$  03 53

$$P(Z < -0,52) = 1 - P(Z < 0,52) = 0,30153 \quad 0,69847$$

$$2 \cdot 0,30153 = 0,60306$$

Svar: ca 60% eller 0,6 /4

$$\frac{P - 0,08}{\sqrt{\frac{0,0736}{50}}} \geq 1,6449 \quad z_{0,05} = 1,6449$$

$$\frac{P - 0,08}{\sqrt{0,001472}} \geq 1,6449 \quad (0,1431 \dots)$$

$$P \cdot 50 = 7,1554$$

minst 8

Svar ca 7 e-post-adresser måste vara icke aktiv för att vi ska förkasta  $H_0$  på  $\alpha = 0,05$  /4



Uppgift 4

du har beroende stickprov -3

1)  $H_0: \mu_1 = \mu_2$   $H_A: \mu_1 \neq \mu_2$  dubbelsidigt test

$\sigma_L^2$  &  $\sigma_N^2$  är okända Vi vet att enstapfserna är

$n_1, n_2 \leq 30$  Vi måste alltså att  $\sigma = \sigma_N$   $N(0,1)$  fördelade

Vi använder  $T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2 (\frac{1}{n_1} + \frac{1}{n_2})}}$

$\bar{X} = 69,95$   $\sum (\bar{X} - \bar{X})^2 = 3^2 \cdot 20$   
 $\bar{X}_N = 70,2$   $n-1$

$\sum_N^2 = 18,472$   $\frac{(64,1 - 69,95)^2 + (66,7 - 69,95)^2 + \dots + 0,5V}{6-1}$

$\bar{X}_N = 70,2$   $\bar{X}_L = 69,95$   
 $\sum_N^2 = 18,472$   $\sum_L^2 = 20,891$

$S_p^2 = 19,8165$   
 $\sqrt{(5 \cdot 18,472 + 20,891 \cdot 5)} / (12-2) =$

$T = \frac{70,2 - 69,95}{\sqrt{19,8165 (\frac{1}{6} + \frac{1}{6})}} = t_{obs} = 0,0973$

$t_{(6+6-2)} = 2,228$   
 $0,025$

Förkasta  $H_0$  om  $t_{obs} > 2,228$  eller ...

Vi kan ej förkasta  $H_0$  på sign.nivå 0,05

117p

# UPPS 4

b)  $\frac{X}{64,1} \cdot 100 = 89 \quad X = 57,049$

R Svar: Elnätpriset är ca 57,1 år 2010 /4

c)  $L_+^0 = \frac{\sum P_{+i} \cdot q_{0i}}{\sum P_{0i} \cdot q_{0i}} \cdot 100 \quad P_+^0 = \frac{\sum P_{+i} \cdot q_{+i}}{\sum P_{0i} \cdot q_{+i}}$

	2013	2014	2013	2014
Elanvändning Lgh Lill	64,1	66,7	Kvantitet	2000
— " — Norr	65,1	66,5	Kvantitet	3300
				2100
				3500

$L = \frac{66,7 \cdot 2000}{65,1 \cdot 3300} \cdot 100 = 62,0956 \approx 62,1$

$P = \frac{66,7 \cdot 2100}{65,1 \cdot 3300} \cdot 100 = 65,20039 \approx 65,2$  /1

Med Lasperres index är elnätpriset ca 62,1 öre per kWh & med Paasche-index 65,2 öre/kWh

d) KPI är konsumentprisindex & används för att beräkna inflationen i Sverige t.d.

— Löpande priser är priser som förändras mycket under varje år

R Urval är ett stickprov taget ur en population. /4



Uppgift 5

a) Vilken risk eftersom vi vet sannolikhetserna för varje strategi? /2,5

	E1	E2	E3	E4	$E(U_i)$
$P=0,4$		0,3	0,2	0,1	$(-20 \cdot 0,4) + (200 \cdot 0,3) + (400 \cdot 0,2) + (80 \cdot 0,1)$
$S_1$	-20	200	400	80	140
$S_2$	-100	100	600	200	130
$S_3$	300	50	150	-300	135

Vi väljer  $S_1$  med störst förväntad nytta /5

	E1	E2	E3	E4	$E(U_i)$
0,25	0,25	0,25	0,25	0,25	115
$S_1$	-20	200	400	80	165
$S_2$	-100	100	600	200	200
$S_3$	300	50	150	-300	50

Laplace alla  $S$  lika sannolika  
vilken strategi väljer man? /4

$-100 \cdot 0,4 + 100 \cdot 0,3 + 600 \cdot 0,2 + 200 \cdot 0,1 = 130$  R

$S = \frac{(-20 + 200)^2 \cdot 0,5V}{4} = 86666,66..$   $\bar{S}_2 = 200$  /1



