

Försättsblad till skriftlig tentamen vid Linköpings universitet



Datum för tentamen	2016-03-22
Sal (1)	TER4
Tid	8-13
Kurskod	732G21
Provkod	TEN1
Kursnamn/benämning Provnamn/benämning	Sambandsmodeller Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Lotta Hallberg
Telefon under skrivtiden	013-281657
Besöker salen ca klockan	10
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Annelie almquist
Tillåtna hjälpmedel	Räknedosa. Läroboken: Applied linear statistical models av Kutner, Nachtsheim m fl som inte får innehålla anteckningar men får ha markeringar och flärpar. Flärpar får ha en liten anteckning. Dessutom ett handskrivet, dubbelsidigt A4 ark med egna anteckningar.
Övrigt	
Antal exemplar i påsen	

TENTAMEN I SAMBANDSMODELLER, 2016-03-22

- Skrivtid:** kl: 8-13
- Hjälpmedel:** Räknedosa. Läroboken: *Applied linear statistical models* av Kutner, Nachtsheim m fl som inte får innehålla anteckningar men får ha markeringar och flärpar. Flärpar får ha en liten anteckning. Dessutom ett handskrivet, dubbelsidigt A4 ark med egna anteckningar.
- Jourhavande lärare:** Lotta Hallberg
- Betygsgränser:** För godkänt krävs minst 12 av 20 poäng och för väl godkänt krävs minst 16 av 20 poäng.

Redovisa och motivera kort alla dina lösningar

Tolka (om möjligt) alla dina resultat!

1

En mäklare har observationer på huspriser den senaste tiden. Priset är i USD. För att ta reda på vad som är viktigt för huspriset så har man även samlat in observationer på kvadratmeteryta m_2 , tomtyta $tomt$ samt antal rum rum .

- a) Du ska med hjälp av framåtvalsmetoden välja den bästa modellen för dessa förklarande variabler (Modell 1-7). Välj även bästa modell med bakåtelimeringsmetoden. Förklara hur du går tillväga. 3p
- b) Mäklaren har valt modellen med endast kvadratmeteryta, m_2 . Han har också observationer på antal badrum i husen. Denna variabel är kodad som två dummy-variabler. Variabeln *mindre* är 1 om antalet badrum är 1 eller 1.5 och noll annars. Variabeln *medel* är 1 om antalet badrum är 2 eller 2.5 och noll annars. Båda är noll om antalet badrum ligger mellan 3 och 4. Se modell 8. Pröva med ett test om antalet badrum (kodad på detta sätt) bidrar med förklaring av priset. Signifikansnivå 5%. 2p
- c) Tolka regressionkoefficienten framför variabeln *mindre* i modell 8. 1p

REG-proceduren
Modell: MODEL1
Beroendevariabel: pris

Antal lästa observationer 150
Antal använda observationer 150

Rot MSE 29946 R-kvadrat 0.4856
Beroende medel 153775 Just. R-kvadr. 0.4821
Koeff.var. 19.47364

Parameterskattningar

Variabel	DF	Parameter-skattning	Standard-fel	t-värde	Pr > t
Skärning	1	63745	7999.97371	7.97	<.0001
m_2	1	49.37473	4.17746	11.82	<.0001

Modell: MODEL2
Beroendevariabel: pris

Antal lästa observationer 150
Antal använda observationer 150

Rot MSE 36507 R-kvadrat 0.2354
Beroende medel 153775 Just. R-kvadr. 0.2303

Koeff.var. 23.74079

Parameterskattningar

Variabel	DF	Parameter- skattning	Standard- fel	t-värde	Pr > t
Skärning	1	144025	3312.31777	43.48	<.0001
tomt	1	9054.72274	1341.32739	6.75	<.0001

Modell: MODEL3

Beroendevariabel: pris

Antal lästa observationer 150
Antal använda observationer 150

Rot MSE 34115 R-kvadrat 0.3324
Beroende medel 153775 Just. R-kvadr. 0.3278
Koeff.var. 22.18477

Parameterskattningar

Variabel	DF	Parameter- skattning	Standard- fel	t-värde	Pr > t
Skärning	1	37969	13776	2.76	0.0066
rum	1	15966	1860.07275	8.58	<.0001

Modell: MODEL4

Beroendevariabel: pris

Antal lästa observationer 150
Antal använda observationer 150

Rot MSE 24552 R-kvadrat 0.6565
Beroende medel 153775 Just. R-kvadr. 0.6519
Koeff.var. 15.96628

Parameterskattningar

Variabel	DF	Parameter- skattning	Standard- fel	t-värde	Pr > t
Skärning	1	61100	6566.39960	9.30	<.0001
m2	1	46.24260	3.44458	13.42	<.0001
tomt	1	7760.02369	907.21706	8.55	<.0001

Modell: MODEL5

Beroendevariabel: pris

Antal lästa observationer 150
Antal använda observationer 150

Rot MSE 30047 R-kvadrat 0.4856
Beroende medel 153775 Just. R-kvadr. 0.4786
Koeff.var. 19.53961

Parameterskattningar

Variabel	DF	Parameter- skattning	Standard- fel	t-värde	Pr > t
Skärning	1	64221	12766	5.03	<.0001
m2	1	49.67304	7.50702	6.62	<.0001
rum	1	-140.54127	2934.12098	-0.05	0.9619

Modell: MODEL6

Beroendevariabel: pris

Antal lästa observationer 150
Antal använda observationer 150

Rot MSE 28383 R-kvadrat 0.5410
Beroende medel 153775 Just. R-kvadr. 0.5347
Koeff.var. 18.45774

Parameterskattningar

Variabel	DF	Parameter- skattning	Standard- fel	t-värde	Pr > t
Skärning	1	33408	11475	2.91	0.0042
tomt	1	8534.31750	1044.16760	8.17	<.0001
rum	1	15328	1549.55022	9.89	<.0001

Modell: MODEL7
Beroendevariabel: pris

Antal lästa observationer 150
Antal använda observationer 150

Rot MSE 24613 R-kvadrat 0.6572
Beroende medel 153775 Just. R-kvadr. 0.6501
Koeff.var. 16.00583

Parameterskattningar

Variabel	DF	Parameter- skattning	Standard- fel	t-värde	Pr > t
Skärning	1	56820	10493	5.42	<.0001
m2	1	43.55130	6.19092	7.03	<.0001
tomt	1	7792.53527	911.58054	8.55	<.0001
rum	1	1261.77232	2409.07169	0.52	0.6012

Modell: MODEL8
Beroendevariabel: pris

Antal lästa observationer 150
Antal använda observationer 150

Variansanalys

Källa	DF	Summa av kvadrater	Medel- kvadrat	F-värde	Sh. > F
Modell	3	1.306386E11	43546204009	49.92	<.0001
Fel	146	1.2735E11	872260043		
Korrigerad total	149	2.579886E11			

Rot MSE 29534 R-kvadrat 0.5064
Beroende medel 153775 Just. R-kvadr. 0.4962
Koeff.var. 19.20601

Parameterskattningar

Variabel	DF	Parameter- skattning	Standard- fel	t-värde	Pr > t	Typ I SS
Skärning	1	77029	16751	4.60	<.0001	3.547014E12
m2	1	43.01660	5.56636	7.73	<.0001	1.252713E11
mindre	1	-9688.92100	11113	-0.87	0.3847	5106693725
medel	1	5090.02578	9311.85140	0.55	0.5855	260623678

2

Vid ett löpande band produceras en viss typ av mobiltelefon. Vid en kvalitetskontroll förfar man på följande vis. Vid en viss tidpunkt så undersöker man telefoner och räknar antalet telefoner tills man får en med någon defekt. Den defekta inkluderas i antalet. Detta görs 8 gånger. Resultat:

137 11 37 548 165 29 33 72

Låt X vara detta antal som är beskrivet ovan. X är då en geometriskt fördelad slumpvariabel. Sannolikhetfunktionen är given av

$$f(x) = (1 - \pi)^{x-1} \pi, \quad x = 1, 2, \dots$$

Maximum-Likelihood-skatta π . Sätt även in observationerna i skattningen.

3p

3

I en fabrik läggs ett ytskikt ut på ett visst material. Tjockleken på ytskiktet kan variera beroende på vilken operatör som lägger på ytskiktet och även på hur maskinen är inställd. För att undersöka hur mycket detta varierar så vill man genomföra ett experiment.

Det finns många operatörer på företaget så tre väljs slumpmässigt ut för experimentet.

Som responsvariabel används: Tjocklek=Thickness

De två faktorerna är: Operatörer=Operator och Inställning=Setting

Nedan visas några utskriftar från Minitab.

ANOVA: Thickness versus Setting; Operator

Factor	Type	Levels	Values
Setting	fixed	3	35; 44; 52
Operator	random	3	1; 2; 3

Analysis of Variance for Thickness		
Source	DF	SS
Setting	2	15676,4
Operator	2	1120,9
Setting*Operator	4	428,4
Error	27	342,5
Total	35	17568,2

Tabulated statistics: Setting; Operator

Rows: Setting	Columns: Operator			
	1	2	3	All
35	39,50 4	40,75 4	41,50 4	40,58 12
44	64,00 4	73,75 4	81,50 4	73,08 12
52	80,50 4	90,50 4	102,00 4	91,00 12
All	61,33 12	68,33 12	75,00 12	68,22 36

Cell Contents: Thickness : Mean
Count

- Testa om maskinens inställning påverkar tjockleken på ytan. Ställ upp korrekta hypoteser. 5% signifikansnivå. 1.5p
- Testa om det är nödvändigt att ta hänsyn till att det är olika operatörer som lägger på ytmaterialet. Ställ upp korrekta hypoteser. 5% signifikansnivå. 1.5p
- Skatta varianskomponenten för operatörerna. Ge även ett approximativt konfidensintervall med hjälp av Satterthwaits metod. 95% konfidensgrad. 2p
- Skatta de tre fixa effekterna för Inställningar. Beräkna även tre Bonferroni konfidensintervall. 95% simultan konfidensgrad. 2p

4

Man är intresserad av attityden i en viss fråga från fyra olika politiska partier. Man valde därför 5 politiker slumpmässigt från vardera partiet. De fick svara på en skala från 1-100, där 100 är positivt inställd. Följande resultat erhöles:

Parti nr:	1	2	3	4
\bar{y}_i	85	80	95	50
s_i	6	7	4	10

- a) Sätt upp en envägs-variationsanalysmodell och pröva om det är skillnad mellan partierna. Visa att $MSE=50.25$. 2p
- b) Använd Tukey's metod för att pröva vilka partier som skiljer sig åt. Simultan (familje) signifikansnivå är 5%. 2p

