

Försättsblad till skriftlig tentamen vid Linköpings universitet



Datum för tentamen	2016-04-16
Sal (1)	TER1
Tid	8-12
Kurskod	732G71
Provkod	TENT
Kursnamn/benämning Provnamn/benämning	Statistik B Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Bertil Wegmann
Telefon under skrivtiden	070-1128321
Besöker salen ca klockan	09 ³⁰
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Annelie Almquist 013-282934 annelie.almquist@liu.se
Tillåtna hjälpmedel	Ett A4-blad med egna handskrivna anteckningar (båda sidor) samt räknedosa.
Övrigt	
Antal exemplar i påsen	55

Tentamen

Linköpings Universitet, Institutionen för datavetenskap, Statistik

- Kurskod och namn: 732G71 Statistik B
- Datum och tid: 2016-04-16, 8-12
- Jourhavande lärare: Bertil Wegmann
- Tillåtna hjälpmedel: Ett A4-blad med egna handskrivna anteckningar (båda sidor) samt räknedosa.
- Betygsgränser: Tentamen omfattar totalt 20p. Godkänt från 12p, väl godkänt från 16p.
Siffrorna i uppgifterna är delvis påhittade.
-

Redovisa och motivera tydligt alla dina lösningar!

Uppgift 1 (7p)

I ett slumpmässigt urval av 7 stycken företagsbranscher samlade en marknadsundersökare in total annonseringskostnad (i miljontals kronor) och omsättning (i miljarder kronor) för respektive bransch i syfte att undersöka hur annonseringskostnader påverkar omsättningen i branscher. Detta gav

SSE = 14.38 och följande tabell:

Bransch	Annonseringskostnad	Omsättning
1	24	29
2	14	15
3	16	19
4	21	27
5	17	17
6	20	24
7	15	14

- a) Skatta en linjär regressionsmodell där omsättningen antas bero på branschens annonseringskostnad. Beräkna ett 90 %-igt konfidensintervall för lutningen i modellen. Tolka konfidensintervallet. Avgör med hjälp av konfidensintervallet om branschens annonseringskostnader signifikant bidrar till att förklara omsättningen och ange på vilken signifikansnivå man avgör det. (4p)
- b) Beräkna en prognos och ett tillhörande 95 %-igt prognosintervall för omsättningen i en bransch där annonseringskostnaden är 19 miljoner kronor. Tolka intervallet i ord. (3p)

Uppgift 2 (2p)

Ett företag i mejeribranschen säljer en ost som kallas för Bragdost. Osten hade en prisutveckling mellan 2013 och 2015 enligt följande:

År	Pris/kg
2013	118
2014	121
2015	123

Under samma tidsperiod var konsumentprisindex för varugruppen mejeriprodukter enligt följande:

År	KPI_mejeri
2013	236
2014	240
2015	243

Beräkna en indexserie för åren 2013, 2014 och 2015, med basår 2013, som visar hur priserna per kilo för företagets Bragdost har utvecklats jämfört med den allmänna prisutvecklingen av mejeriprodukter under perioden.

Uppgift 3 (6p)

Ett engelskt läkemedelsföretag vill undersöka hur försäljningen av receptbelagda läkemedel beror på olika variabler. Man har därför samlat in information från 17 stycken slumpmässigt utvalda apotek under ett år och studerat följande variabler:

Sales (y) Årsgenomsnittet av veckoförsäljningen av receptbelagda läkemedel (1000-tals pund)

FloorSp (x1) Butiksyta (kvadratfot)

PresPct (x2) Procentandel av butiksytan som används till försäljning av receptbelagda läkemedel

Parking (x3) Antalet parkeringsplatser som är avsedda för apotekskunder

Income (x4) Genomsnittlig veckoinkomst per capita i apotekets upptagningsområde (100-tals pund)

ShopCtr (x5) = 1 om apoteket ligger i ett shoppingcentrum och annars 0.

Man prövar först följande regressionsmodell för variabeln Sales (y):

$$y = \beta_0 + \beta_1 x_5 + \beta_2 x_1 + \beta_3 x_1^2 + \varepsilon$$

där en ny förklaringsvariabel $x_1^2 = (\text{FloorSp})^2 = \text{FloorSp_sq}$ har skapats. Detta gav följande parameterskattningar av den multipla linjära regressionsmodellen från Minitab (se nästa sida):

Regression Analysis: Sales versus ShopCtr; FloorSp; FloorSp_sq

Coefficients

Term	Coef
Constant	18,4
ShopCtr	-10,26
FloorSp	-0,00239
FloorSp_sq	0,000001

- a) Antag att ett apotek har butiksytan 4000 kvadratfot och att det ligger i ett shoppingcentrum. Använd den skattade regressionsmodellen ovan för att uppskatta förändringen i årsgenomsnittlig försäljning (förändringen i Sales(y)) för detta apotek om butiksytan skulle öka med 500 kvadratfot. (2p)

Man vill också pröva följande regressionsmodeller för variabeln Sales (y):

$$y = \beta_0 + \beta_1 x_4 + \beta_2 x_5 + \beta_3 x_4 x_5 + \varepsilon$$

och

$$y = \beta_0 + \beta_1 x_4 + \beta_2 x_5 + \varepsilon$$

Detta gav följande skattningsresultat av respektive multipel linjär regressionsmodell från Minitab:

MODELL 1:

Regression Analysis: Sales versus Income; ShopCtr; Income*ShopCtr

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	234,485	78,162		
Income	1	65,265	65,265		
ShopCtr	1	68,932	68,932		
Income*ShopCtr	1	9,355	9,355		
Error	13	261,632	20,126		
Lack-of-Fit	11	209,632	19,057		
Pure Error	2	52,000	26,000		
Total	16	496,118			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
4,48615	47,26%	35,09%	18,25%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	14,41	3,34			
Income	0,620	0,344			
ShopCtr	-13,12	7,09			
Income*ShopCtr	0,347	0,509			

MODELL 2:

Regression Analysis: Sales versus Income; ShopCtr

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	2	225,13	112,57	5,82	0,015
Income	1	189,44	189,44	9,79	0,007
ShopCtr	1	183,14	183,14	9,46	0,008
Error	14	270,99	19,36		
Lack-of-Fit	12	218,99	18,25	0,70	0,722
Pure Error	2	52,00	26,00		
Total	16	496,12			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
4,39957	45,38%	37,58%	26,58%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	13,03	2,61	5,00	0,000	
Income	0,778	0,249	3,13	0,007	1,75
ShopCtr	-8,71	2,83	-3,08	0,008	1,75

- Redovisa den skattade regressionsekvationen för ett apotek som ligger i ett shoppingcentrum för modellen med interaktionsterm mellan Income och ShopCtr. (1p)
- Testa på 5 % signifikansnivå om interaktionstermen är signifikant i modell 1. Dra fullständig slutsats av ditt test i ord. (2p)
- Vilken modell är bäst utifrån ditt svar i c) och utskrifterna ovan? Motivera. (1p)

Uppgift 4 (5p)

Antag att du vill skatta följande modell: $y_t = TR_t + SN_t + CL_t + IR_t$

Försäljningsvärden för en viss produkt under tre år gav:

År	2013			2014			2015		
	jan-apr	maj-aug	sep-dec	jan-apr	maj-aug	sep-dec	jan-apr	maj-aug	sep-dec
Försäljning	195	180	160	210	197	165	220	204	176

- Beräkna centrerade glidande medelvärden för försäljningsvärdena och ange vad dessa medelvärden skattar i modellen. (2p)
- Beräkna skattningar för säsongskomponenterna sn_t . (3p)

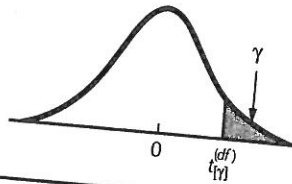


TABLE A2 A t-Table: Values of $t_{\gamma}^{(df)}$

df	$t_{[.10]}^{(df)}$	$t_{[.05]}^{(df)}$	$t_{[.025]}^{(df)}$	$t_{[.01]}^{(df)}$	$t_{[.005]}^{(df)}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
inf.	1.282	1.645	1.960	2.326	2.576

Source: From "Table of Percentage Points of the t-Distribution," by Maxine Merrington, *Biometrika* 32 (1941), 300. Reproduced by permission of the *Biometrika* Trustees.