

Tentamen

Linköpings universitet, Institutionen för datavetenskap, STIMA

Kurskod och namn:	732G71 Statistik B
Datum och tid:	2018-04-09, 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)

En bankdirektör vill anpassa en enkel linjär regressionsmodell för hur bankens årliga vinstmarginal i procent beror på årliga nettointäkten per insatt krona från bankens kunder. Därför valde man ut 7 på varandra följande år, vilket gav följande tabell:

År	Nettointäkt per insatt krona	Vinstmarginal (i %)
1	3,96	73
2	3,38	68
3	3,48	70
4	3,86	78
5	4,22	92
6	3,82	74
7	3,63	77

- Skatta regressionslinjens skärning och lutning i den enkla linjära regressionsmodellen. Tolka skattningarna i ord. Motivera om tolkningen av skärningen är relevant. (2p)
- Testa på 1 % signifikansnivå om nettointäkten per insatt krona signifikant bidrar till att förklara vinstmarginalen. SSE = 115.2. (2p)
- Beräkna ett 95 % konfidensintervall för den genomsnittliga vinstmarginalen per år då nettointäkten per krona är lika med 3,54. Tolka konfidensintervallet i ord. SSE = 115.2. (2p)
- Ett av modellantagandena för regressionsmodellen bygger på normalfördelningen. Förklara vad detta modellantagande innebär och hur residualerna bör se ut för att detta antagande ska vara uppfyllt. (1p)

Uppgift 2 (6p)

Bankdirektören vill utvidga modellen i uppgift 1 med information om hur årtalet och antalet bankkontor också påverkar vinstmarginalen per år. Därför valde man i en större studie ut 11 på varandra följande år och noterade värden på respektive variabel nedan för varje år:

y = Vinstmarginal

x1 = Nettointäkt per insatt krona

x2 = Antalet bankkontor i tusentals

x3 = Årtal: antar värdet 1 för år 1, 2 för år 2, osv. till 11 för år 11

Skattning av en multipel linjär regressionsmodell med samtliga förklaringsvariabler x1, x2 och x3 gav följande reducerade Minitab-utskrift:

Regression Analysis: y versus x1; x2; x3

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	490,63	163,54		
Error	7	94,10	13,44		
Total	10	584,73			

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	71,9	45,5	1,58	0,158	
x1	24,55	4,87	5,04	0,001	1,11
x2	-12,95	6,07	-2,13	0,070	2,64
x3	-0,611	0,582	-1,05	0,329	2,77

- a) Testa på 5 % signifikansnivå om den multipla linjära regressionsmodellen är signifikant eller inte. Dra fullständig slutsats av ditt test i ord. (2p)

Skattning av en enkel linjär regressionsmodell med förklaringsvariabel x1 gav följande reducerade Minitab-utskrift:

Regression Analysis: y versus x1

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	417,3	417,28	22,43	0,001
Error	9	167,4	18,60		
Total	10	584,7			

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-20,7	20,8	-0,99	0,346	
x1	25,70	5,43	4,74	0,001	1,00

- b) Testa med ett test på 5 % signifikansnivå om minst en av förklaringsvariablerna x2, x3, bidrar signifikant till att förklara bankens vinstmarginal i modellen, givet att förklaringsvariabeln x1 redan finns med i modellen. Dra fullständig slutsats av ditt test i ord. (2p)
- c) Beräkna en prognos för vinstmarginalen år 12 utifrån den multipla linjära regressionsmodellen med $x_1 = 4$ och $x_2 = 6,5$. Tolka prognosen i ord. (1p)
- d) Beräkna förklaringsgraden för den multipla linjära regressionsmodellen och tolka förklaringsgraden i ord. (1p)

Uppgift 3 (4p)

Ett livsmedelsföretag säljer en yoghurt som kallas för Jo-gut. Yoghurten hade en prisutveckling mellan 2015 och 2017 enligt följande:

År	Pris/liter
2015	11,8
2016	12,3
2017	12,5

- a) Beräkna ett enkelt prisindex för åren 2015, 2016 och 2017 med basår 2017 som visar hur priserna per liter har utvecklats för företagets yoghurt Jo-gut. (1p)

- b) Livsmedelsföretaget säljer också en fil som kallas för Fi-dolus. Filen hade en prisutveckling mellan 2015 och 2017 enligt följande:

År	Pris/liter
2015	14,2
2016	14,8
2017	14,4

Försäljningsvärdet år 2015 var för Jo-gut 3,7 miljoner kronor och för Fi-dolus 6,1 miljoner kronor. Beräkna ett sammansatt prisindex för Jo-gut och Fi-dolus. Använd basår 2015. Vilket viktsystem, Paasches eller Laspeyres, behöver användas för beräkningen? Motivera. (3p)

Uppgift 4 (3p)

Klassisk komponentuppdelning har anpassats på ett prisindex för hushållsmaskiner nedan. Prisindexet sträcker sig från kvartal 1 år 2015 till kvartal 4 år 2017.

Time Series Decomposition for Hushållsmaskiner

Data Hushållsmaskiner
Length 12
NMissing 0

Fitted Trend Equation

$$Y_t = 98,458 + 0,071 \times t$$

Seasonal Indices

Period	Index
1	0,782
2	0,965
3	1,022
4	1,231

- a) Vilken metod (additiv eller multiplikativ) har använts för den klassiska komponentuppdelningen? Motivera. (1p)
- b) Tolka samtliga säsongindex i ord. (1p)
- c) Gör en prognos för kvartal 1 år 2018. (1p)

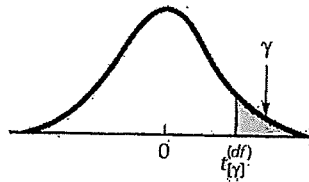


TABLE A2 A t-Table: Values of $t_{(df)}^{(\gamma)}$

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.

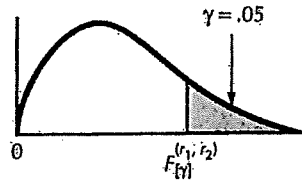


TABLE A3 An F-Table: Values of F_{γ}

Denominator Degrees of Freedom, r_2	Numerator Degrees of Freedom, r_1								
	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Source: From "Tables of Percentage Points of the Inverted Beta (F)-Distribution," by Maxine Merrington and Catherine M. Thompson, *Biometrika* 33 (1943), 73-88. Reproduced by permission of the *Biometrika* Trustees.