Línuleg tölfræðilíkön (09.15.46)

Vikublað 4

1: Define the SLR problem using matrices and use matrix algebra to derive the OLS solutions, variances and covariances of parameter estimates.

2: Consider data collected using two instruments.

x	y1	y2
1	7.94	-3.67
2	7.58	7.42
3	2.51	8.38
4	6.48	17.93
5	10.25	14.55
6	4.17	17.70
7	19.28	14.52
8	18.17	13.39
9	18.37	22.20
10	19.27	20.44

Set up the X-matrices for two independent regression equations, two lines with the same slope, two lines with the same intercept and one line through the entire data set. This gives 4 different models.

With the largest model being $E[y_{ij}] = \alpha_i + \beta_i x_{ij},\ i = 1,2,\ j = 1,\ldots,10$, the possible null hypotheses are:

$$H_1 : \alpha_1 = \alpha_2 \quad (1)$$

$$H_2 : \beta_1 = \beta_2 \quad (2)$$

$$H_3 : \alpha_1 = \alpha_2 \text{ and } \beta_1 = \beta_2 \quad (3)$$

In addition to these null hypotheses, consider a general alternative hypothesis,

$$H_a : \alpha_1 \neq \alpha_2, \beta_1 \neq \beta_2,$$

which places no restriction on the parameters.

The first two null hypotheses can only be tested against the general alternative (H_a) but the third can be tested against all three ohter hypotheses as possible alternative hypotheses. Note that we only consider two-tailed alternatives. This gives 5 different tests.

Test all 5 combinations of hypotheses concerning slopes and intercepts by using matrix algebra in R and Excel. Verify using lm() in R.

3: In order to test hypotheses concerning whether a line is the "right" model, this hypothesis needs to be compared with some "larger" model. If the "larger" model explains significantly more variation that the straight line then it will be rejected. It is for example possible to investigate the significance of a quadratic term but also the following approach can be used: Assume that there are many measurements available at each x-value, i.e. $y_{i1}, \ldots, y_{in_i}$. Write down and develop the F-test which is based on the model $y_{ij} = \mu_i + e_{ij}$ to test the hypothesis $H_0 : \mu_i = \alpha + \beta x_i$ against the alternative $H_a :$ not H_0.

4: The following table lists the catches in numbers of cod in Icelandic waters, by age and years (in millions of fish).

A single yearclass can be tracked by considering a diagonal from the table. For example, 2.6 million fish from the 1974 yearclass were caught at age 3 in 1977, 16.3 at age 4 in 1978 and so on.

Pick a yearclass and use it as y-data with age as the x-variable. Plot y against x and select a model to fit y as a function of x.

Year	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	c13	c14
1976	23.578	39.790	21.092	24.395	5.803	5.343	1.297	0.633	0.205	0.155	0.065	0.029
1977	2.614	42.659	32.465	12.162	13.017	2.809	1.773	0.421	0.086	0.024	0.006	0.002
1978	5.999	16.287	43.931	17.626	8.729	4.119	0.978	0.348	0.119	0.048	0.015	0.027
1979	7.186	28.427	13.772	34.443	14.130	4.426	1.432	0.350	0.168	0.043	0.024	0.004
1980	4.348	28.530	32.500	15.119	27.090	7.847	2.228	0.646	0.246	0.099	0.025	0.004
1981	2.118	13.297	39.195	23.247	12.710	26.455	4.804	1.677	0.582	0.228	0.053	0.068
1982	3.285	20.812	24.462	28.351	14.012	7.666	11.517	1.912	0.327	0.094	0.043	0.011
1983	3.554	10.910	24.305	18.944	17.382	8.381	2.054	2.733	0.514	0.215	0.064	0.037
1984	6.750	31.553	19.420	15.326	8.082	7.336	2.680	0.512	0.538	0.195	0.090	0.036
1985	6.457	24.552	35.392	18.267	8.711	4.201	2.264	1.063	0.217	0.233	0.102	0.038
1986	20.642	20.330	26.644	30.839	11.413	4.441	1.771	0.805	0.392	0.103	0.076	0.040
1987	11.002	62.130	27.192	15.127	15.695	4.159	1.463	0.592	0.253	0.142	0.046	0.058
1988	6.713	39.323	55.895	18.663	6.399	5.877	1.345	0.455	0.305	0.157	0.114	0.025
1989	2.605	27.983	50.059	31.455	6.010	1.915	0.881	0.225	0.107	0.086	0.038	0.005
1990	5.785	12.313	27.179	44.534	17.037	2.573	0.609	0.322	0.118	0.050	0.015	0.020
1991	8.554	25.131	15.491	21.514	25.038	6.364	0.903	0.243	0.125	0.063	0.011	0.012
1992	12.217	21.708	26.524	11.413	10.073	8.304	2.006	0.257	0.046	0.032	0.012	0.008
1993	20.500	33.078	15.195	13.281	3.583	2.785	2.707	1.181	0.180	0.034	0.011	0.013
1994	6.160	24.142	19.666	6.968	4.393	1.257	0.599	0.508	0.283	0.049	0.018	0.006
1995	10.782	9.113	16.848	13.081	4.120	1.598	0.313	0.184	0.156	0.141	0.029	0.008
1996	5.356	14.886	7.372	12.307	9.43	2.157	0.837	0.208	0.076	0.065	0.055	0.005
1997	1.722	16.442	17.298	6.711	7.379	5.958	1.147	0.493	0.126	0.028	0.037	0.021
1998	3.548	7.707	25.394	20.167	5.893	3.856	2.951	0.5	0.196	0.055	0.033	0.013
1999	2.684	20.834	14.764	25.193	12.004	2.472	1.37	0.849	0.138	0.049	0.01	0.005

Assorted exercises from the book: 6.15*, 6.22, 6.23, 6.26, 6.27, 7.1, 7.4, 7.7, 8.8*, 8.17*, 8.21