Department of Statistics

COURSE STATS 330

Assignment 2, 2008

Instructions: Hand in your completed assignment to the Student Resource Centre by 4pm on Thursday 21 Aug.

The data set for this assignment is in the file **earthquakes.txt** which is available on the course web page.

The data in the file **earthquakes.txt** consist of measurements made on 54 major earthquakes which occurred in the period 1900-1975.

The variables in the data set relevant to this assignment are

RootSeisMmt: the "root" of the seismic moment of the earthquake
ExpSeisMmt: the "exponent" of the seismic moment of the earthquake
Displacement: The average displacement of the fault line (meters)
Length: the length of the fault
Magnitude: the magnitude of the earthquake, on the Richter scale

Note: The seismic moment is a measure of the shear effect of an earthquake. In terms of the variables above, it is calculated from the "root" and "exponent" as follows: if the root has a value of 7.6, and the exponent a value of 27, then the seismic moment is 7.6×10^{27}, measured in dynes/cm.

Earthquake theory states that the magnitude is related to the seismic moment, the displacement and the length by an equation of the form

$$M = a + b \log(M_0 D/L) \qquad (1)$$

where M is the magnitude, M_0 is the seismic moment, D is the displacement, L is the length and a and b are constants. This is a special case of the equation

$$M = \beta_0 + \beta_1 \log(M_0) + \beta_2 \log(D) + \beta_3 \log(L). \qquad (2)$$

Requirements

1. Load the data into R, and make a data frame **earthquakes.df** to contain the data. Check for any typographical errors. Create a new variable SeisMmt from the variables RootSeisMmt and ExpSeisMmt as described above and add it to the data frame. [5 marks]

2. Fit a regression using Magnitude as the response and the variables log(SeisMmt), log(Displacement) and log(length) as the explanatory variables. Comment on the fit. [20 marks, 10 for the fitting and diagnostics and 10 for the discussion]

3. If the fit is unsatisfactory, modify the model to improve the fit. [10 marks]

4. Do you think the data are well described by the model (1)? That is, is it plausible that in model (2), we have $\beta_1 = \beta_2$ and $\beta_1 = -\beta_3$? [5 marks]

Extra Question for STATS 762 only

Test the hypothesis that $\beta_1 = \beta_2 = -\beta_3$ by performing a simultaneous test of the hypotheses that $\beta_1 = \beta_2$ and $\beta_1 = -\beta_3$, using the theory below.

Theory for the extra question (for more detail see the STATS 310 notes on p 120 and following pages)

To test the hypothesis that Aβ= 0 where A is a matrix and β is the vector of regression coefficients, we use the test statistic

$$F = (A\hat{\beta})^T (A\,\mathrm{cov}(\hat{\beta})A^T)^{-1} A\hat{\beta} / q$$

which has an F distribution with q and df degrees of freedom (q=number of rows in A, df = residual df from the regression) when the hypothesis is true.
The estimated covariance matrix is $\hat{\sigma}^2 (X^T X)^{-1}$ and can be calculated by extracting the estimate $\hat{\sigma}^2$ and the matrix $(X^T X)^{-1}$ from the summary function as follows:
If the regression fit is in the "regression object" `reg.lm` returned by the lm function then the estimate of σ is returned by `summary(reg.lm)$sigma` and the matrix $(X^T X)^{-1}$ by `summary(reg.lm)$cov.unscaled`. Alternatively, you can get X using `model.matrix(reg.lm)`.

Data follows overleaf

Earthquake data

	Magnitude	RootSeisMmt	ExpSeisMmt	Displacement	Length	Width	LogEnergy
1	7.9	7.6	27	2.10	130	70	26.09
2	7.5	4.6	26	3.00	35	13	25.60
3	7.0	2.0	26	3.00	20	11	25.48
4	7.0	6.8	25	1.00	20	10	24.53
5	8.0	3.0	28	3.30	185	100	26.73
6	6.3	2.8	25	0.20	30	15	23.27
7	7.1	4.8	26	2.00	70	11	25.14
8	7.0	4.2	27	1.10	170	50	25.43
9	7.4	3.6	26	2.50	33	13	25.44
10	8.0	1.5	28	3.10	120	80	26.59
11	7.1	8.7	25	2.20	12	11	25.20
12	8.1	1.5	28	3.10	120	80	26.59
13	7.3	3.3	26	2.50	30	13	25.44
14	8.1	1.7	28	1.90	180	100	26.25
15	7.7	2.0	27	4.60	60	18	26.19
16	8.3	1.5	29	5.00	700	60	27.03
17	7.1	1.3	26	2.00	36	6	24.86
18	8.3	2.7	30	24.00	800	200	28.91
19	7.0	9.0	25	2.50	12	10	25.27
20	6.9	3.3	25	0.60	20	8	24.00
21	6.8	1.9	26	0.66	34	13	24.57
22	6.7	1.2	26	1.05	32	11	24.60
23	8.2	7.5	28	3.00	250	150	26.95
24	6.5	3.8	25	0.48	27	9	23.83
25	6.5	3.0	26	0.87	28.5	17	24.96
26	8.5	7.5	29	7.00	500	300	28.02
27	6.9	1.8	26	1.25	33	9	24.83
28	6.9	4.3	26	1.20	50	20	25.01
29	7.4	6.2	27	3.50	80	31	26.43
30	7.5	3.2	27	3.30	80	30	26.12
31	6.9	5.8	26	1.50	40	8	25.34
32	7.9	1.4	29	2.50	500	150	26.85
33	7.5	3.4	27	1.20	50	80	25.91
34	6.3	1.6	26	0.65	32	12	24.51
35	6.4	2.9	25	0.90	28.5	4.5	23.96
36	7.2	2.3	26	1.63	35	12	25.02
37	5.4	2.0	24	0.05	18	6	21.78
38	5.9	8.3	24	0.30	10	10	23.40

39	7.5	2.0	28	2.60	140	80	26.57
40	7.1	8.3	26	1.70	80	20	25.25
41	6.7	6.3	25	0.58	33	11	24.04
42	7.9	2.8	28	4.10	150	100	26.88
43	6.1	1.9	25	0.92	10	6	24.24
44	6.6	2.1	26	0.90	19	19	25.00
45	7.3	1.0	27	2.10	80	20	25.42
46	7.9	6.0	27	2.50	80	50	26.27
47	7.8	2.2	28	2.90	180	85	26.55
48	6.6	4.3	25	0.72	20	10	24.19
49	7.6	1.0	28	1.60	130	70	26.09
50	6.2	2.2	25	0.65	14	8	24.01
51	6.6	1.1	26	1.40	19	14	24.91
52	6.9	9.5	26	1.50	70	25	25.31
53	7.4	6.7	27	1.60	100	60	26.03
54	6.9	1.1	26	1.00	25	15	24.64