Chapter 4

Logistic Regression Modeling

In this chapter, we used logistic regressions to fit a model. In the first section, we fit a model to estimate the probability of the fertility based on year, region, age group and period (season). In the second section, we construct model curves for age-specific fertility in each region, period and season based on coefficients in model.

4.1 Logistic regression model to estimate the fertility

As indicated in preliminary analysis result, the fertility rate outcome is of continuous data type. We now consider modeling the birth outcome as binary (0 is non-birth and 1 is birth). The determinants used in this analysis were age-group.region (1-56), period.region (1-16) and season (1-4).

Table 4.1 shows the result of fitting a logistic regression model to the status of birth in a month by region, age group and period is statistically significant. The denominators (the numbers of women giving birth) are estimated using the method described in Chapter 3. The model contains coefficients for each combination of age group and region (56 parameters given by adding the age-group: region coefficients to the constant), as well as eight parameters describing the change from 2002-2003 to 2004-2005 for each region, and a further three parameters describing a quarterly seasonal effect. However, the fit is inadequate with residual deviance 902.3 and 381 degrees of freedom. This is due to the very large size of the data set.

Determinant	Coefficient	St Error	z-value	p-value
Constant	-5.3715	0.0165	-325.64	0.0000
Age-group. Region:	Baseline: 15-19. west<20			0.0000
15-19. east<20	-0.1945	0.0198	-9.85	0.0000
15-19. west20-49	0.2870	0.0335	8.58	0.0000
15-19. east20-49	0.2371	0.0281	8.43	0.0000
15-19. west50-79	-0.0186	0.0283	-0.66	0.5114
15-19. east50-79	-0.0259	0.0230	-1.13	0.2605
15-19. west80+	0.2941	0.0390	7.55	0.0000
15-19. east80+	0.0313	0.0231	1.35	0.1756
20-24. west<20	0.5246	0.0188	27.88	0.0000
20-24. east<20	0.3276	0.0184	17.79	0.0000
20-24. west20-49	0.6892	0.0301	22.87	0.0000
20-24. east20-49	0.5184	0.0263	19.72	0.0000
20-24. west50-79	0.4906	0.0255	19.26	0.0000
20-24. east50-79	0.6266	0.0206	30.40	0.0000
20-24. west80+	0.6075	0.0359	16.93	0.0000
20-24. east80+	0.7282	0.0207	35.25	0.0000
25-29. west<20	0.6683	0.0186	36.02	0.0000
25-29. east<20	0.5286	0.0181	29.21	0.0000
25-29. west20-49	0.6387	0.0312	20.47	0.0000
25-29. east20-49	0.4369	0.0271	16.15	0.0000
25-29. west50-79	0.6377	0.0250	25.54	0.0000
25-29. east50-79	0.6637	0.0207	32.13	0.0000
25-29. west80+	0.6282	0.0365	17.22	0.0000
25-29. east80+	0.8184	0.0207	39.53	0.0000
30-34. west<20	0.2125	0.0201	10.55	0.0000
30-34. east<20	0.2150	0.0187	11.50	0.0000
30-34. west20-49	0.2236	0.0357	6.26	0.0000
30-34. east20-49	0.0719	0.0304	2.37	0.0179
30-34. west50-79	0.2335	0.0275	8.50	0.0000
30-34. east50-79	0.4651	0.0215	21.65	0.0000
30-34. west80+	0.4167	0.0398	10.46	0.0000
30-34. east80+	0.7391	0.0215	34.75	0.0000

Table 4.1: Results of logistic model to outcome monthly birth rate

Determinant	Coefficient	St Error	z-value	p-value
35-39. east<20	-0.6542	0.0257	-25.43	0.0000
35-39. east<20	-0.5902	0.0212	-27.78	0.0000
35-39. west20-49	-0.4097	0.0458	-8.94	0.0000
35-39. east20-49	-0.4765	0.0376	-12.66	0.0000
35-39. west50-79	-0.3608	0.0335	-10.76	0.5054
35-39. east50-79	-0.0160	0.0239	-0.67	0.0206
35-39. west80+	-0.1177	0.0509	-2.31	0.0000
35-39. east80+	0.4202	0.0231	18.19	0.0000
40-44. west<20	-1.9960	0.0454	-44.01	0.0000
40-44. east<20	-2.1515	0.0346	-62.24	0.0000
40-44. west20-49	-1.5471	0.0788	-19.62	0.0000
40-44. east20-49	-1.4933	0.0601	-24.84	0.0000
40-44. west50-79	-1.4900	0.0551	-27.06	0.0000
40-44. east50-79	-1.0513	0.0342	-30.74	0.0000
40-44. west80+	-0.9763	0.0788	-12.39	0.0000
40-44. east80+	-0.5259	0.0319	-16.49	0.0000
45-49. west<20	-3.9852	0.1289	-30.92	0.0000
45-49. east<20	-4.3331	0.0999	-43.37	0.0000
45-49. west20-49	-2.2887	0.1229	-18.62	0.0000
45-49. east20-49	-2.4921	0.1043	-23.88	0.0000
45-49. west50-79	-2.5537	0.0985	-25.94	0.0000
45-49. east50-79	-2.2625	0.0625	-36.21	0.0000
45-49. west80+	-2.1703	0.1523	-14.25	0.0000
45-49. east80+	-1.5744	0.0530	-29.68	0.0000
Period. Region:	Baseline: 2002-3. all regions			0.0000
2004-5. west<20	0.0763	0.0120	6.37	0.0000
2004-5. east<20	0.0642	0.0080	7.98	0.0000
2004-5. west20-49	-0.0915	0.0240	-3.82	0.0001
2004-5. east20-49	0.0119	0.0192	0.62	0.5355
2004-5. west50-79	0.0562	0.0175	3.20	0.0014
2004-5. east50-79	-0.0064	0.0116	-0.55	0.5827
2004-5. west80+	-0.1134	0.0288	-3.93	0.0001
2004-5. east80+	-0.1303	0.0116	-11.28	0.0000

Table 4.1: Results of logistic model to outcome monthly birth rate (ctd.)

Determinant	Coefficient	St Error	z-value	p-value
Season:	Baseline: April-June			0.0000
January-March	-0.0531	0.0067	-7.94	0.0000
July-September	0.0582	0.0065	8.96	0.0000
October-December	0.0120	0.0066	1.83	0.0665

r-sq: 0.0376 df: 38061569 ResDev: 2267542.98 nRecords: 5376 sampleSize: Grouped data (448 cells): df: 381 ResDev: 902.28

Table 4.1: Results of logistic model to outcome monthly birth rate (ctd.)

4.2 Fertility trends in Southern Thailand

The coefficients obtained from the model are plotted in Figure 4.1. It can be seen from this plot that the fertility rates of the mother in both east and west regions where the percent Muslim was more than 80% were higher than the mother in both east and west region where the percent Muslim was less than 20%. Among regions with more than 80% Muslims, the fertility of mother in the east region was higher than in the west region.

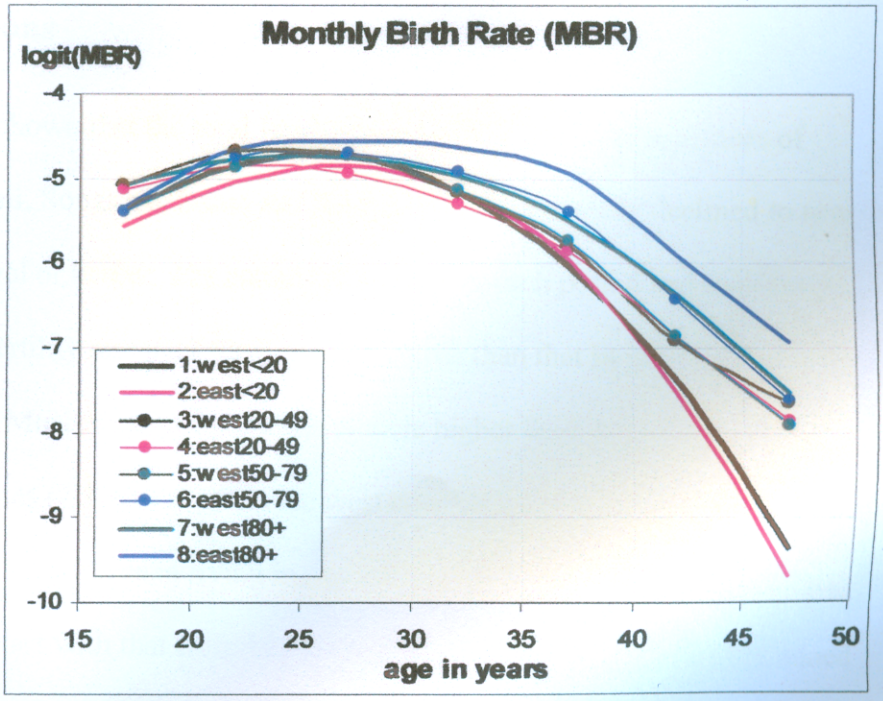

Figure 4.1: Logits of monthly birth rates based on logistic regression model