Chapter 4

Statistical Modeling

This chapter covers statistical models used to explain under-five death rates by PHA, gender, and year. The models comprise the log-linear regression models for all-cause death rates and perinatal originating conditions death rates. The models were fitted to the data based on sum contrasts. Confidence intervals for each factor level adjusted for other factors were obtained from the models.

4.1 All- cause death rates

Model fitting

A linear model was fitted to the log transformed death rates outcome with gender, PHA, and year as determinants. The r-squared of log-linear model was 80%. The residuals plot shows few residual values at the lower and upper ends departing from the straight line. Thus the normality assumption of residuals was acceptable. Therefore log-linear regression was considered to be used.

Model diagnostic

Figure 4.1 shows plots of residuals versus the normal quantile. The log-linear regression model is appropriate for modeling all-cause deaths with less residual values departing from a diagonal line. This results in normally distributed residual.
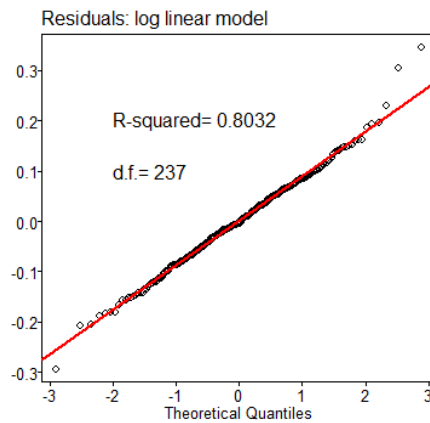
Figure 4.1: Residuals plot for log-linear regression model of all-cause deaths

Figure 4.2 shows plot of observed death rates per 100,000 population against fitted values obtained from log-linear model.
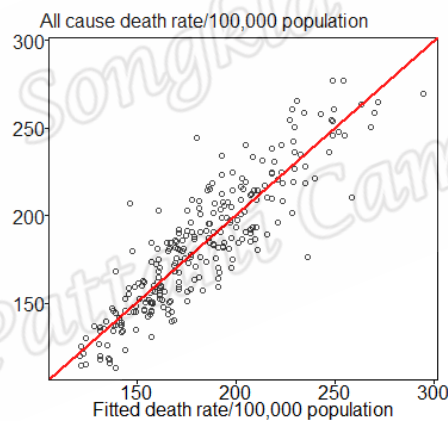

Figure 4.2: Plot of observed death rates against fitted values

from log-linear regression model of all-cause

Table 4.1 shows the coefficients and their standard errors from log-linear model based on sum contrasts. The results show that PHA, gender, and year are statistically significantly associated with under-five death rates.

Table 4.1: Coefficients, standard errors, and p-values based on log-linear regression model fitted to all-cause death rates

Factors	Coef	SE	P-value
Constant	5.190	0.006	< 0.001
PHA			
PHA1	-0.008	0.020	0.707
PHA2	0.137	0.020	<0.001
PHA3	0.102	0.020	<0.001
PHA4	0.186	0.020	<0.001
PHA5	-0.241	0.020	<0.001
PHA6	-0.162	0.020	<0.001
PHA7	-0.117	0.020	<0.001
PHA8	0.012	0.020	0.546
PHA9	0.015	0.020	0.450
PHA10	0.018	0.020	0.369
PHA11	-0.083	0.020	<0.001
PHA12	0.268	0.020	<0.001
PHA13	-0.129	0.020	<0.001
Gender			
Boy	0.103	0.006	< 0.001
Girl	-0.103	0.006	< 0.001
Year			
2000	0.030	0.017	0.085
2001	0.125	0.017	<0.001
2002	0.036	0.017	0.039
2003	-0.031	0.017	0.077
2004	-0.020	0.017	0.248
2005	0.013	0.017	0.463
2006	-0.022	0.017	0.220
2007	-0.043	0.017	0.014
2008	-0.043	0.017	0.014
2009	-0.045	0.017	0.011

Confidence intervals

The 95% confidence interval graphs of under-five death rates per 100,000 population for each factor level adjusted for other factors from log-linear regression model using sum contrasts are shown in Figure 4.3. The blue horizontal line represent the average of all-cause death rates which is 177.9 per 100,000 population.

All-cause death rates in PHA2, PHA3, PHA4, and PHA12 were higher than the average death rates, whereas death rates in PHA5, PHA6, PHA7, PHA11, and PHA13 were lower than the average. PHA12 had the highest death rates. The death rates for boys were higher than those for girls. In 2001, the death rates peaked and higher than the average. The death rates in 2007, 2008, and 2009 were lower than the average.
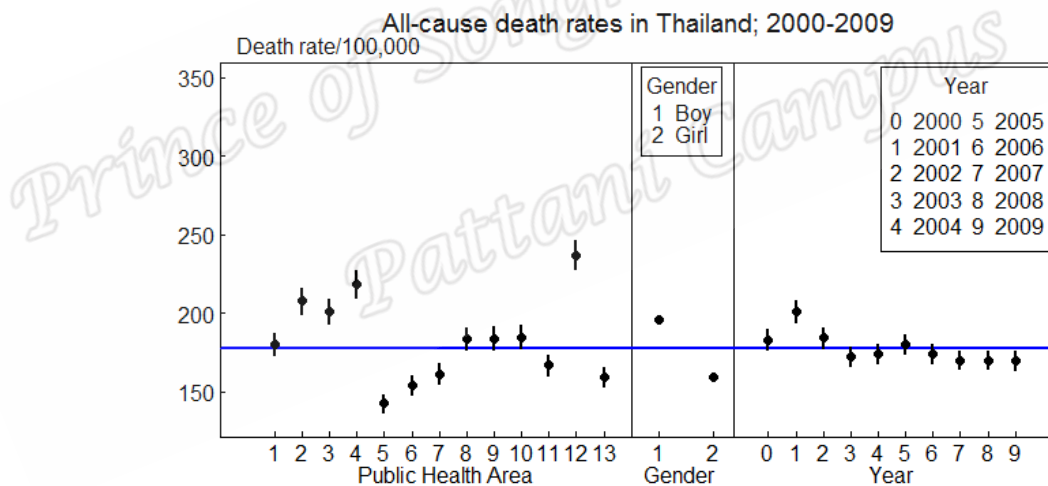

Figure 4.3: Confidence intervals for PHA gender and year of all-cause death rates per 100,000 population from log-linear regression model

Figures 4.4 shows the bar chart of all-cause death rates per 100,000 population by PHA based on confidence intervals from log-linear regression model. The graph clearly shows that death rates in the Central (PHA 2, PHA3, and PHA4) and in the lower South regions (PHA12) were higher than the average.
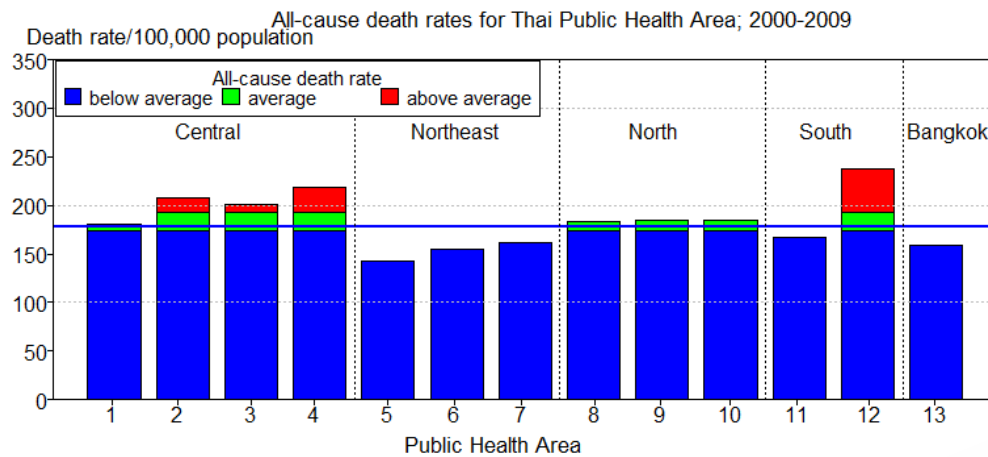
Figure 4.4: Bar chart of all-cause death rates per 100,000 population

Figure 4.5 shows the map of all-cause death rates per 100,000 population by PHA, based on confidence intervals from log-linear regression model. The map clearly shows that the PHA2 (AngThong, Lop Buri, Sing Buri, ChaiNat, Saraburi, and Nakhon Nayok), PHA3 (Samut Prakan, Chon Buri, Rayong, Chanthaburi, Trat, Chachoengsao, Prachin Buri, and SaKaeo), PHA4 (Ratchaburi, Kanchanaburi, Suphan Buri, Nakhon Pathom, Samut Sakhon, Samut Songkhram, Phetchaburi, and Prachuap Khiri Khan) and PHA12 (Songkhla, Satun, Trang, Phattalung, Pattani, Yala, and Narathiwat) had death rates higher than the average.
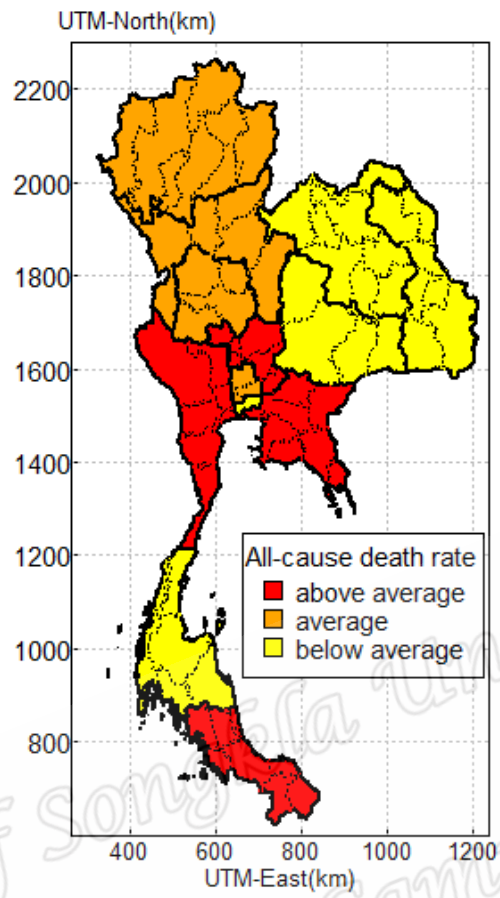
Figure 4.5: The map of all-cause death rates for each PHA

from log-linear regression model

4.2 Perinatal originating conditions death rates

Model fitting

A linear model was first fitted to the log transformed of perinatal originating conditions death rates outcome with PHA, gender, and year as determinants. The r-squared of log-linear model was 89%. The residuals plot shows few residual values at the upper end depart from the straight line. Thus the normality assumption of residuals was acceptable.

Model diagnostic

Figure 4.6 shows plots of residuals versus the normal quantile. The log-linear regression model is appropriate for modeling perinatal originating conditions death rates with less residual values departing from a diagonal line. This results in normally distributed residual.
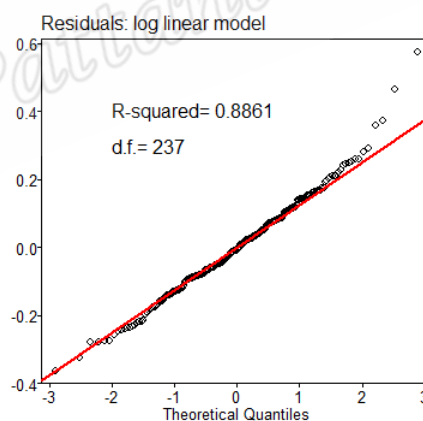

Figure 4.6: Residuals plot for log-linear regression model of perinatal originating conditions

Figure 4.7 shows plot observed death rates per 100,000 population against fitted values obtained from log-linear model.
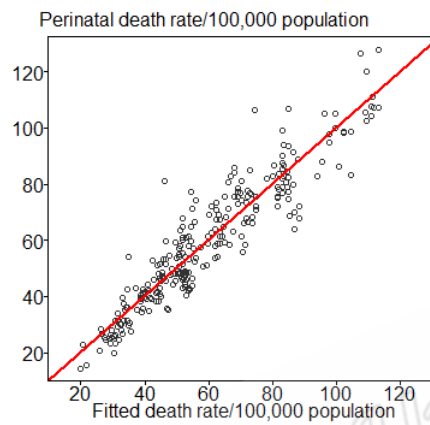

Figure 4.7: Plot of observed death rate against fitted values from log-linear model of perinatal originating conditions

Table 4.2 shows the coefficients and their standard errors from log-linear model based on sum contrasts. The results show that PHA, gender, and year are statistically significantly associated with perinatal originating conditions death rates.

Table 4.2: Coefficients, standard errors, and p-values based on log-linear regression model fitted to perinatal originating conditions death rates

Factors	Coef	SE	P-value
Constant	4.008	0.009	< 0.001
PHA			
PHA1	-0.001	0.031	0.976
PHA2	0.260	0.031	<0.001
PHA3	0.199	0.031	<0.001
PHA4	0.143	0.031	<0.001
PHA5	-0.156	0.031	<0.001
PHA6	-0.035	0.031	0.255
PHA7	-0.040	0.031	0.195
PHA8	-0.033	0.031	0.290
PHA9	-0.014	0.031	0.648
PHA10	-0.150	0.031	<0.001
PHA11	-0.476	0.031	<0.001
PHA12	0.272	0.031	<0.001
PHA13	0.031	0.031	0.320
Gender			
Boy	0.242	0.009	<0.001
Girl	-0.242	0.009	<0.001
Year			
2000	-0.377	0.027	<0.001
2001	0.026	0.027	0.332
2002	-0.265	0.027	<0.001
2003	-0.313	0.027	<0.001
2004	-0.083	0.027	0.002
2005	0.210	0.027	<0.001
2006	0.182	0.027	<0.001
2007	0.202	0.027	<0.001
2008	0.215	0.027	<0.001
2009	0.203	0.027	<0.001

Confidence intervals

The 95% confidence interval graphs of perinatal originating conditions death rates per 100,000 population for each factor level adjusted for other factors from log-linear regression model using sum contrasts are shown in Figures 4.8. The blue horizontal line in each graph represents the average of perinatal originating conditions death rates, which is 58.4 per 100,000 population.

Perinatal originating conditions death rates in PHA2, PHA3, PHA4, and PHA12 were higher than the average, whereas death rates in PHA5, PHA10, and PHA11 were lower than the average. PHA12 had the highest death rates. The death rates for boys were higher than those for girls. The death rates in 2005, 2006, 2007, 2008, and 2009 were higher than the average.
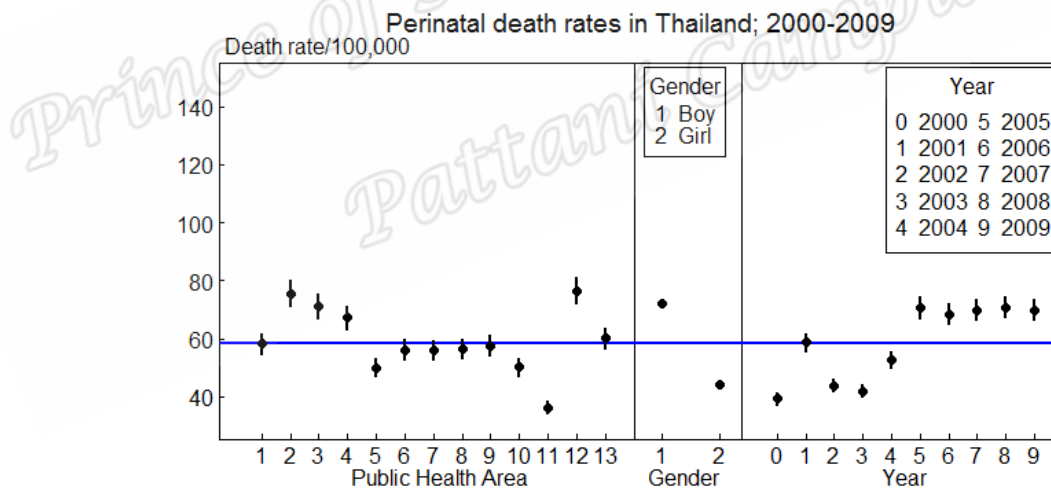
Figure 4.8: Confidence intervals for PHA gender and year of perinatal originating conditions death rates per 100,000 population from log-linear regression model

Figures 4.9 shows the bar chart for perinatal originating conditions death rates per 100,000 population by PHA based on confidence intervals from log-linear regression.

The graph shows that death rates in the Central (PHA 2, PHA3, and PHA4) and in the lower South regions (PHA12) were higher than the average.
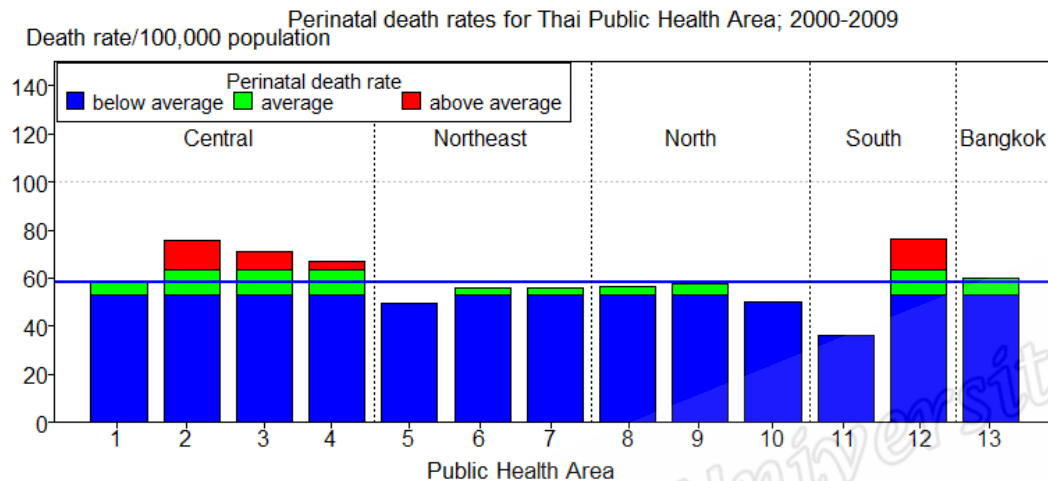

Figure 4.9: Bar chart of perinatal originating conditions death rates per 100,000 population

Figure 4.10 shows the map of perinatal originating conditions death rates per 100,000 population by province, based on confidence intervals from log-linear regression model. The map clearly shows that the PHA2 (AngThong, Lop Buri, Sing Buri, ChaiNat, Saraburi, and Nakhon Nayok), PHA3 (Samut Prakan, Chon Buri, Rayong, Chanthaburi, Trat, Chachoengsao, Prachin Buri, and SaKaeo), PHA4 (Ratchaburi, Kanchanaburi, Suphan Buri, Nakhon Pathom, Samut Sakhon, Samut Songkhram, Phetchaburi, and Prachuap Khiri Khan) and PHA12 (Songkhla, Satun, Trang, Phattalung, Pattani, Yala, and Narathiwat) had death rates higher than the average.
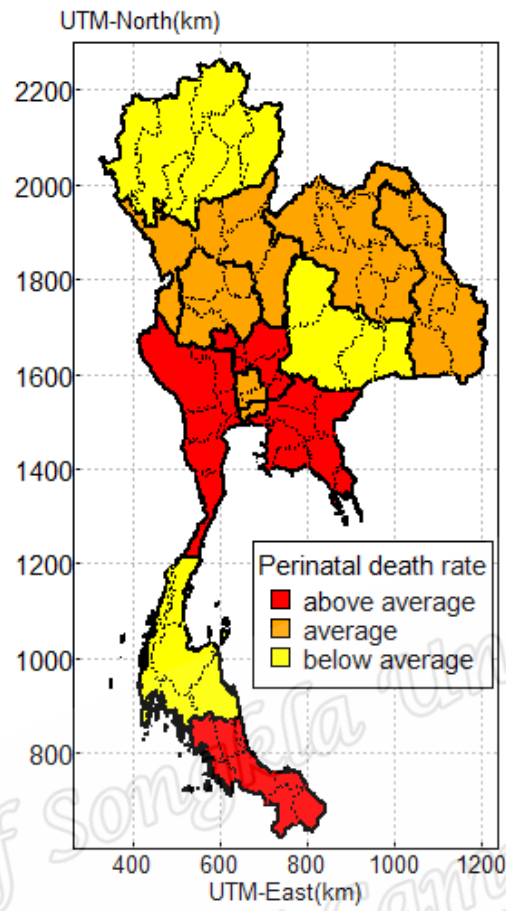
Figure 4.10: The map of perinatal originating conditions death rate for each PHA

from log-linear regression model