Chapter 4

Statistical Modeling

In this chapter we present statistical modeling and maps of multiplicative models for youth non-participation rates in Pattani and Songkhla Provinces of Thailand.

4.1 Additive logistic regression models

The additive model using logistic regression was fitted to youth participation in Songkhla and Pattani Provinces. The determinants were age, gender, religion and location. The results of fitting logistic regression were as follows.

Table 4.1 shows results from fitting additive logistic regression model to the data for Pattani province.

In Figure 4.1 the top left plot shows observed counts versus expected counts, closely approximating in a straight line. The top middle plot shows observed percentages versus fitted percentages and the top right plot shows residuals versus normal quantiles. The residuals plot shows that there are no outliers. The lower panels of Figure 4.1 shows percentages of non-participation for gender-religion and super-tambon, each adjusted for the other factor.

For Pattani data when logistic regression was fitted to the non-participation rates, the additive model (2.1) in table 4.1 gave a residual deviance of 278.2 with 99 degrees of freedom.

Table 4.1: Regression coefficients and standard errors from additive model in Pattani

Factors	Coef.	SE	Factors	Coef.	SE
intercept	-1.13	0.07	PanareCity (19)	0.32	0.09
			ThaKam+Bn (20)	0.22	0.15
Sabarang+Anuru (1)	-0.18	0.06	Don (21)	-0.29	0.22
Bana (2)	-0.17	0.08	Khuan+Thanum (22)	-0.36	0.13
Jabangtiko+Talubo (3)	-0.28	0.10	NorthPanare (23)	0.42	0.07
Rusamilae+ (4)	-0.06	0.05	Mayo+TYD (24)	-0.10	0.04
KhokPho (5)	-0.10	0.11	Taluban (25)	-0.70	0.11
Makrut+Bangkro (6)	-0.24	0.15	WestSaiburi+Karubi (26)	-0.42	0.07
Pabon+Changhaitok (7)	0.18	0.14	EastSaiburi (27)	-0.01	0.05
Saikhao (8)	-0.49	0.21	MaiKaen (28)	-0.88	0.13
Napradu (9)	0.24	0.11	SouthYaring (29)	0.11	0.05
Pahlo (10)	-0.04	0.14	NorthYaring (30)	0.04	0.04
ThungPhala (11)	0.60	0.17	NorthYarang (31)	-0.01	0.04
ThaRua (12)	0.24	0.12	SouthYarang (32)	0.15	0.05
NaKet (13)	-0.09	0.12	Maelan+MungTia (33)	0.20	0.11
KhuanNori (14)	-0.37	0.17	Parai (34)	0.46	0.12
Tuyong+BangTawa (15)	0.82	0.07	Muslim male	0.63	0.02
CentralNongChick (16)	-0.05	0.08	Muslim female	0.35	0.02
SouthNongChick (17)	0.35	0.07	Other male	-0.42	0.06
WestNongChick (18)	0.49	0.05	Other female	-0.56	0.06
Null deviance: 1513.56 df 135			Residual deviance: 278.19 df 99		

In Figure 4.1 the residual plot indicates some departure from linearity. The Muslim male group had the highest non-participation rate (28.4 percent), followed by Muslim females (23.3 percent), other males (12.2 percent) and other females (10.8 percent). Non-participation rate were high overall in ThungPhala (11), Tuyong+BangTawa (15), SouthNOngChick (17), WesNongChick (18), PanareCity (19), NorthPanare (23), SouthYarang (32) and Parai (34).

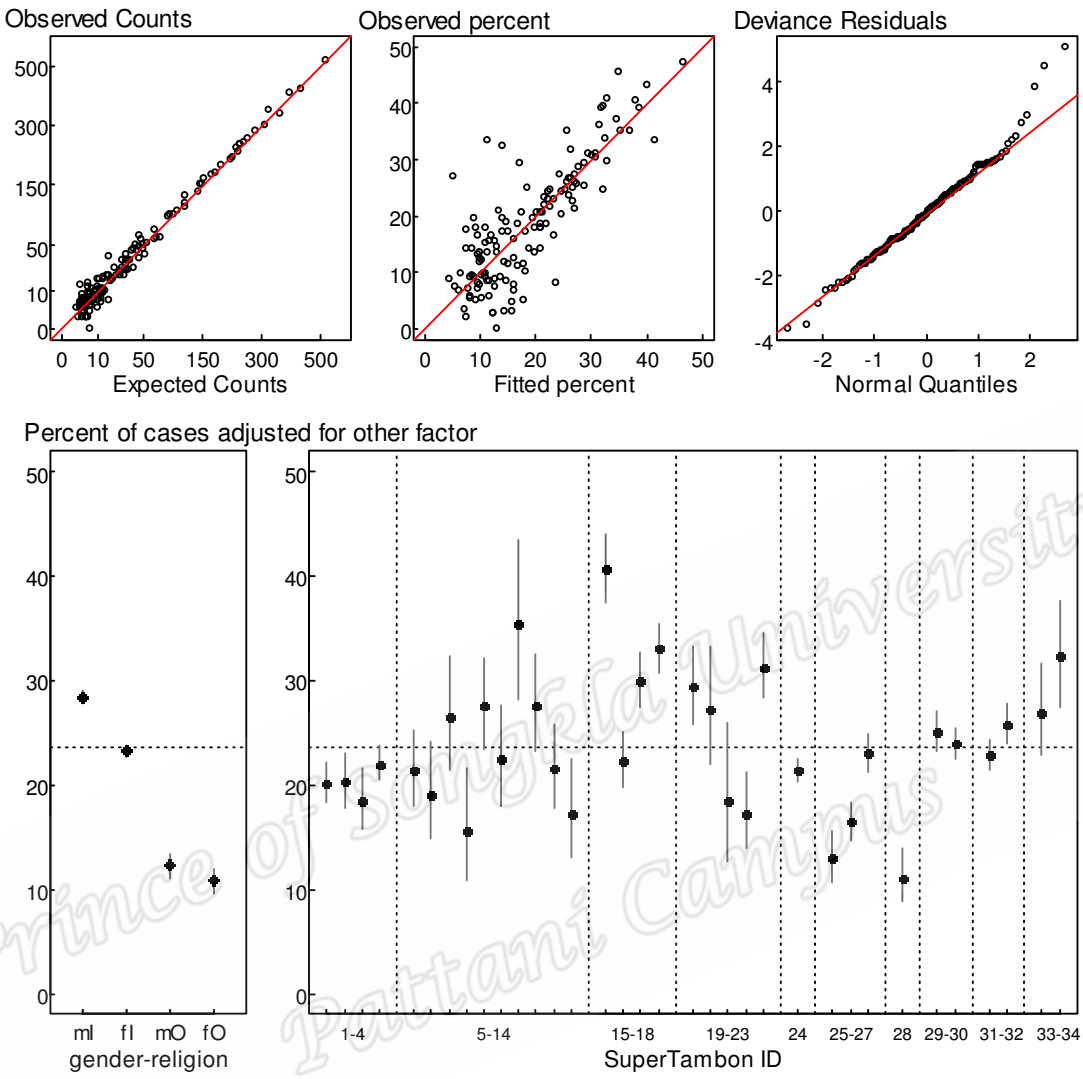

Figure 4.1: The logistic model results for 15-17 year olds who were 'at risk', in Pattani super-tambons

Table 4.2 shows logistic regression fitted to the non-participation rates. For Songkhla the model gave a residual deviance of 530.7 with 153 degrees of freedom, and in Figure 4.2 the residual plot indicated no departure from linearity. The Muslim male group had the highest non-participation rate (21.2 percent), followed by Muslim females (16.1 percent), other males (13.8 percent) and other females (9.7 percent). Non-participation rate were high overall in Phawong+Koyo (4), NSatingPa+Krasin (5), Paching (8), ThaMoSai (11), Khu+Khae (13), ThaMuang (21), BanNot (25),

Ranot (28), PadngBesa (36), KlongHae+KlongUtapao (42), NBangklam (47), NSingHaNakorn (49), HuaKhao (51) and MungNgam (52).

Table 4.2: Regression coefficients and standard errors from additive model in Songkhla

Factors	Coef.	SE	Factors	Coef.	SE
(Intercept)	-1.28	0.05	Ranot (28)	0.30	0.04
			ERattaphum (29)	-0.03	0.06
BoYang (1)	0.07	0.04	THaChamuang (30)	-0.05	0.09
KhaoRupChang (2)	-0.25	0.07	KhaoPhra (31)	0.15	0.09
KoTaeo+Tungwang (3)	-0.14	0.11	Sadao (32)	0.15	0.09
Phawong+Koyo (4)	0.34	0.08	Prik (33)	-0.44	0.11
NSatingPa+Krasasin (5)	0.23	0.05	NSadao (34)	-0.12	0.08
SSatingPa+BangKiat (6)	-0.10	0.08	SamNakTaeo (35)	-0.43	0.13
BanNa (7)	-0.48	0.11	PadangBesa (36)	0.21	0.07
PaChing (8)	0.45	0.15	SamnakKham (37)	0.19	0.11
SapanMaiKaen+Sakom (9)	-0.34	0.11	HatYai (38)	-0.28	0.06
NaWa (10)	-0.82	0.21	KhuanLang (39)	-0.04	0.06
ThaMoSai (11)	0.63	0.14	NHatYai+NNamom (40)	0.10	0.06
NaThap+Chanong (12)	0.01	0.08	KhoHong (41)	-0.09	0.04
Khu+Khae (13)	0.45	0.11	KhoHae+KhongUtapao (42)	0.36	0.08
KlongPia+TalingChan (14)	-0.12	0.11	Chalung (43)	-0.80	0.22
NamKhao+KhunTatWai (15)	-0.40	0.18	TungTamSao+WKhk (44)	-0.10	0.09
ENatawi (16)	0.15	0.08	EKhk+SHatyai+SNamom(45)	-0.06	0.06
W-CNathawi (17)	0.12	0.10	KhuanNiang (46)	-0.32	0.07
SNathawi (18)	-0.08	0.08	NBangklam (47)	0.33	0.12
Tepa+PakBang+KoSaba (19)	-0.02	0.07	ThaChang (48)	-0.01	0.09
LamPlai+WangYai (20)	-0.41	0.10	NSingHaNakorn (49)	0.60	0.06
ThaMuang (21)	0.27	0.09	SathingMo (50)	0.04	0.09
SaKom (22)	-0.07	0.13	HuaKhao (51)	0.43	0.08
CSabaYoi (23)	0.10	0.07	MuangNgam (52)	0.31	0.10
Pian (24)	0.12	0.12	Muslim male	0.44	0.03
BanNot (25)	0.48	0.13	Muslim female	0.10	0.03
Khuha (26)	-0.07	0.14	Other male	-0.07	0.02
KaoDaeng+BaHoi (27)	-0.54	0.16	Other female	-0.47	0.02
Null deviance: 1865.68 df 207			Res. deviance: 530.68 df 153		

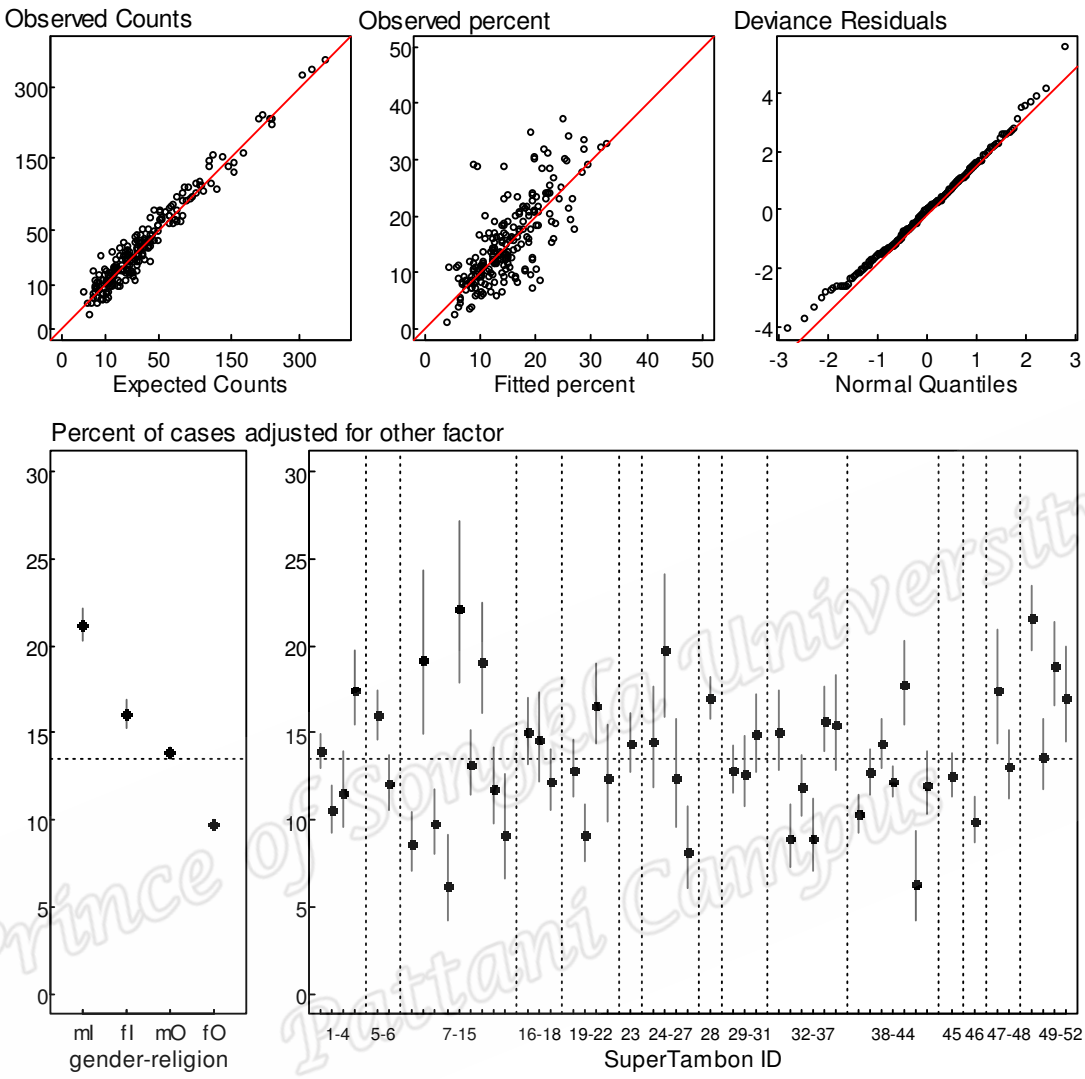

Figure 4.2: The logistic model results for 15-17 year olds who were 'at risk', in Songkhla super-tambons

Although the residual plots indicate that the model is plausible when fitting the data for each province, the residual deviances from these additive models are much higher than their corresponding numbers of degrees of freed, so these models do not account adequately for the variation in the data for both Pattani and Songkhla, and further analysis is needed.

4.2 Multiplicative logistic regression models

The area plots in Pattani and Songkhla show that both Muslim males and females had higher non-participation rates than non-Muslims in most areas. But in some areas non-Muslims had higher rates. This indicates that the patterns are different in regions.

The further analysis involved fitting the model that allows for interaction between the demographic factor and region. So in further analysis we used multiplicative models.

Table 4.3 shows the coefficients and standard errors from the multiplicative model and corresponding plots are shown in Figure 4.3. Note that the residual deviance is reduced by more than a factor of three, although still higher than the number of degrees of freedom. This indicates that the multiplicative model, though not ideal, is substantially better than the additive model.

Table 4.3: Regression coefficients and standard errors from multiplicative model in Pattani

Factors	Coef.	SE	Factors	Coef.	SE
(Intercept)	-1.79	0.07			
Sabarang+Anuru (1)	0.00	0.00	gramma:Sabarang+Anuru (1)	0.71	0.08
Bana (2)	0.20	0.11	gramma:Bana (2)	0.16	0.10
Jabangtiko+Talubo (3)	-0.36	0.23	gramma:Jabangtiko+Talubo (3)	0.96	0.25
Rusamilae+ (4)	0.12	0.10	gramma:Rusamilae+ (4)	0.62	0.09
KhokPho (5)	-0.10	0.18	gramma:KhokPho (5)	0.95	0.19
Makrut+Bangkro (6)	-0.03	0.17	gramma:Makrut+Bangkro (6)	0.53	0.18
Pabon+Changhaitok (7)	0.24	0.22	gramma:Pabon+Changhaitok (7)	0.84	0.24
Saikhao (8)	-0.06	0.22	gramma:Saikhao (8)	-0.02	0.24
Napradu (9)	0.52	0.14	gramma:Napradu (9)	0.40	0.14
Pahlo (10)	0.53	0.17	gramma:Pahlo (10)	-0.12	0.18
ThungPhala (11)	0.81	0.18	gramma:ThungPhala (11)	0.67	0.21
ThaRua (12)	-0.11	0.24	gramma:ThaRua (12)	1.44	0.26
NaKet (13)	0.17	0.16	gramma:NaKet (13)	0.48	0.16
KhuanNori (14)	0.02	0.19	gramma:KhuanNori (14)	0.09	0.20
Tuyong+BangTawa (15)	0.79	0.14	gramma:Tuyong+BangTawa(15)	0.94	0.14
CentralNongChick (16)	0.30	0.17	gramma:CentralNongChick (16)	0.41	0.18
SouthNongChick (17)	0.26	0.20	gramma:SouthNongChick (17)	0.95	0.22
WestNongChick (18)	0.41	0.14	gramma:WestNongChick (18)	0.96	0.15
PanareCity (19)	0.63	0.17	gramma:PanareCity (19)	0.45	0.18
ThaKam+Bn (20)	0.23	0.21	gramma:ThaKam+Bn (20)	0.96	0.22
Don (21)	-0.01	0.25	gramma:Don (21)	0.41	0.26
Khuan+Thanum (22)	-0.28	0.22	gramma:Khuan+Thanum (22)	0.77	0.24
NorthPanare (23)	0.81	0.13	gramma:NorthPanare (23)	0.32	0.13
Mayo+TYD (24)	0.16	0.14	gramma:Mayo+TYD (24)	0.52	0.14
Taluban (25)	-0.34	0.18	gramma:Taluban (25)	0.35	0.19
WestSaiburi+Karubi (26)	-0.37	0.23	gramma:WestSaiburi+Karubi(26)	0.77	0.24
EastSaiburi (27)	0.07	0.17	gramma:EastSaiburi (27)	0.74	0.18
MaiKaen (28)	-0.15	0.16	gramma:MaiKaen (28)	-0.36	0.17
SouthYaring (29)	0.13	0.18	gramma:SouthYaring (29)	0.81	0.19
NorthYaring (30)	-0.04	0.17	gramma:NorthYaring (30)	0.93	0.18
NorthYarang (31)	0.31	0.14	gramma:NorthYarang (31)	0.44	0.14
SouthYarang (32)	0.50	0.13	gramma:SouthYarang (32)	0.40	0.13
Maelan+MungTia (33)	0.42	0.14	gramma:Maelan+MungTia (33)	0.55	0.14
Parai (34)	0.22	0.22	gramma:Parai (34)	1.33	0.24
Null deviance: 1513.56 df 135			Residual deviance: 136.63 df 68		

Figure 4.3 shows plots of observed counts and percentages against fitted values as well as residuals against normal quantiles in its upper panels, with a scatter plot of the non-participation rates versus disparity indexes for the regions in the lower panel.

The lower panel shows a scatter plot of the non-participation rates and the disparity indexes for the 34 statistical regions in Pattani province. On the x-axis, a score of zero for the disparity index would mean that, for the particular region, all four gender-religion groups would have the same non-participation levels, a positive value indicates greater non-participation rates for Muslim youth, and a negative value indicates greater non-participation rates for non-Muslims. On the y- axis, a placement near the dotted line would indicate that, for that particular region, the overall level of non-participation would be at the overall average for all regions. Regions with points above and below the dotted line have levels that are higher and lower, respectively, than the average level of non-participation for 15-17 year olds in Pattani province.

The plot also shows confidence intervals that can be used to make valid statistical conclusions. The vertical line segments denote a 95 percent confidence interval for the non-participation rates (when compared with the overall mean), whereas the horizontal line segments denote a 95 percent confidence intervals for the disparity indexes (compared to zero disparity). The negative value of disparity index reflects that Muslim had a low proportion of non-participation referred to as positive disparity for Muslim. To reduce clutter the confidence intervals are shown only for selected regions.

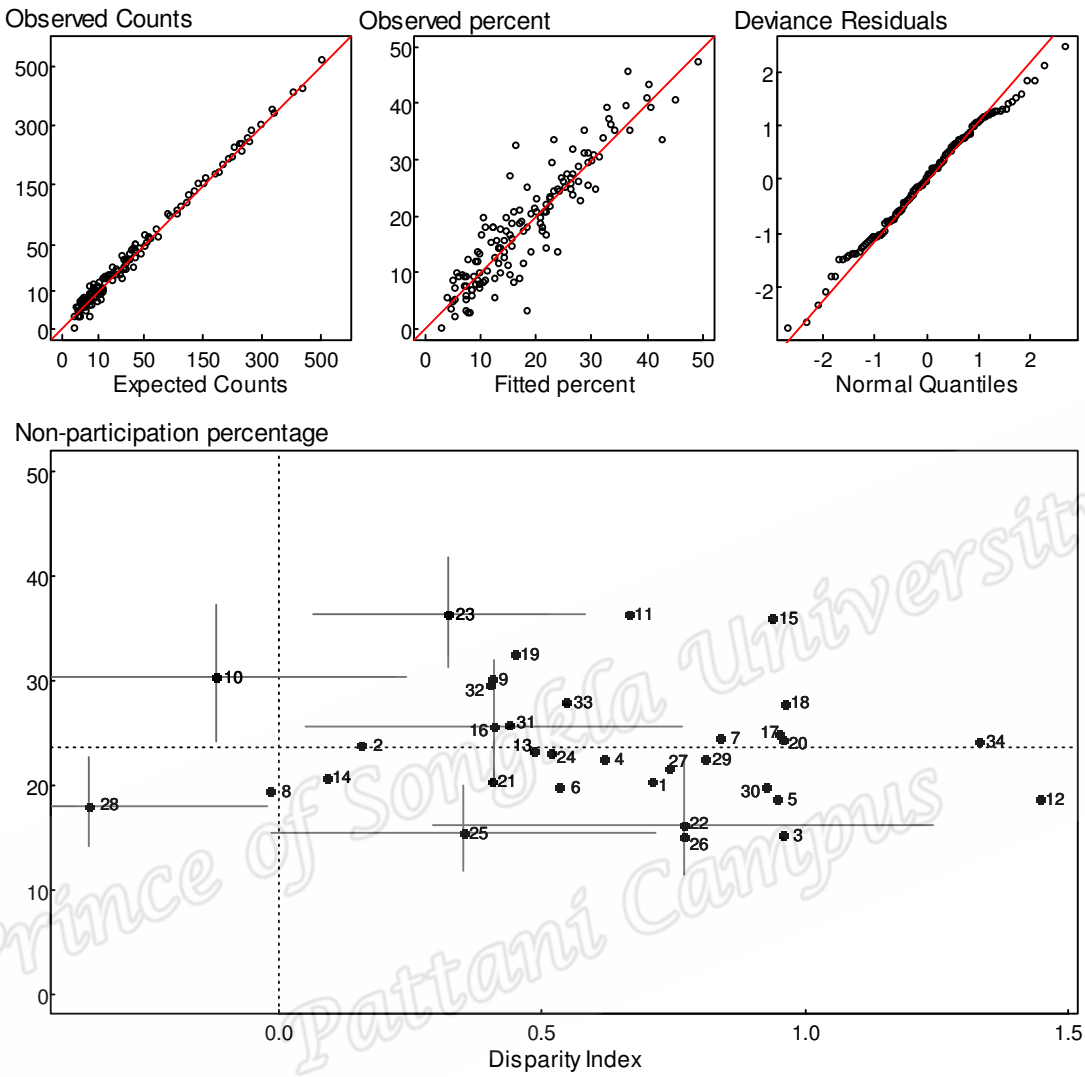

Figure 4.3: Multiplicative model results for youth non-participation rates in Pattani regions

Using the confidence intervals, the regions may be classified into clusters, with the following results:

1: Higher than average non participation, positive disparity (0 region).

2: Higher than average non participation, no evidence of disparity (1 region).

3: Higher than average non participation, negative disparity (7 regions).

4: Average non-participation, positive disparity (0 region).

5: Average non-participation, no evidence of disparity (4 regions).

6: Average non-participation, negative disparity (17 regions).

7: Lower than average non-participation, positive disparity (1 region).

8: Lower than average non-participation, no evidence of disparity (1 region).

9: Lower than average non-participation, negative disparity (3 regions).

Figure 4.4 shows a corresponding thematic map, where disparity is labelled as Muslim- for Muslim disadvantage and Muslim+ for Muslim advantage. A high disparity occurred in Napradu (9), Thung Phala (11), Tuyoung+BangTawa (15), Panare City (19), North Panare (23), South Yarang (32) and Maelan+Mung Tia (33) with Muslims having a higher than average non-participation, negative disparity. In Phalo (10) there is no evidence that the disparity index is higher than average non participation. In contrast, Muslims in the MaiKaen (28) region have a lower than average non-participation, positive disparity.

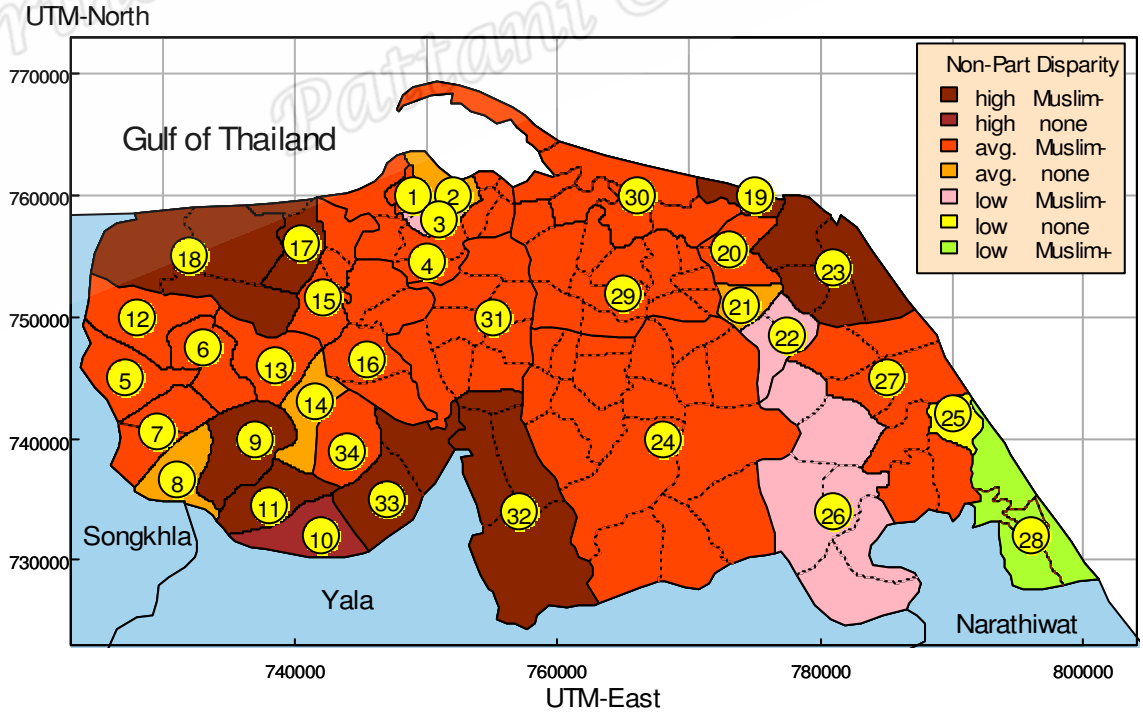

Figure 4.4: Map of multiplicative model for youth non-participation rates in Pattani regions

Table 4.4 shows the coefficients and standard errors from the multiplicative model fitted to the non-participation rate for each super-tambon in Songkhla Province, giving a residual deviance of 277.9 with 104 degree of freedom.

Table 4.4: Regression coefficients and standard errors from multiplicative model in Songkhla

Factors	Coef.	SE	Factors	Coef.	SE
(Intercept)	-1.61	0.06			
BoYang (1)	0.00	0.00	gramma:BoYang (1)	0.55	0.07
KhaoRupChang (2)	-0.31	0.11	gramma:KhaoRupChang (2)	0.58	0.12
KoTaeo+Tungwang (3)	-0.34	0.13	gramma:KoTaeo+Tungwang (3)	0.09	0.13
Phawong+Koyo (4)	0.16	0.10	gramma:Phawong+Koyo (4)	0.35	0.09
NSatingPa+Krasasin (5)	0.15	0.10	gramma:NSatingPa+Krasasin (5)	0.52	0.09
SSatingPa+BangKiat (6)	-0.13	0.11	gramma:SSatingPa+BangKiat (6)	0.62	0.12
BanNa (7)	-0.56	0.14	gramma:BanNa (7)	0.16	0.15
PaChing (8)	0.23	0.18	gramma:PaChing (8)	0.49	0.20
SapanMaiKaen+Sakom (9)	-0.45	0.14	gramma:SapanMaiKaen+Sakom (9)	0.23	0.15
NaWa (10)	-1.06	0.23	gramma:NaWa (10)	0.67	0.27
ThaMoSai (11)	0.40	0.16	gramma:ThaMoSai (11)	0.60	0.18
NaThap+Chanong (12)	-0.24	0.12	gramma:NaThap+Chanong (12)	0.51	0.13
Khu+Khae (13)	-0.09	0.18	gramma:Khu+Khae (13)	0.96	0.20
KlongPia+TalingChan (14)	-0.36	0.14	gramma:KlongPia+TalingChan (14)	0.49	0.15
NamKhao+KhunTatWai (15)	-0.57	0.19	gramma:NamKhao+KhunTatWai (15)	0.35	0.22
ENatawi (16)	-0.02	0.10	gramma:ENatawi (16)	0.44	0.10
W-CNathawi (17)	-0.08	0.12	gramma:W-CNathawi (17)	0.65	0.13
SNathawi (18)	-0.10	0.11	gramma:SNathawi (18)	0.92	0.12
Tepa+PakBang+KoSaba (19)	-0.24	0.12	gramma:Tepa+PakBang+KoSaba(19)	0.44	0.12
LamPlai+WangYai (20)	-0.63	0.12	gramma:LamPlai+WangYai (20)	0.68	0.13
ThaMuang (21)	-0.17	0.13	gramma:ThaMuang (21)	0.91	0.14
SaKom (22)	-0.09	0.14	gramma:SaKom (22)	-0.45	0.14
CSabaYoi (23)	-0.19	0.12	gramma:CSabaYoi (23)	0.56	0.12
Pian (24)	-0.11	0.16	gramma:Pian (24)	0.46	0.19
BanNot (25)	0.42	0.15	gramma:BanNot (25)	0.04	0.17
Khuha (26)	-0.35	0.17	gramma:Khuha (26)	0.71	0.19
KaoDaeng+BaHoi (27)	-0.64	0.17	gramma:KaoDaeng+BaHoi (27)	0.79	0.21
Ranot (28)	0.24	0.08	gramma:Ranot (28)	0.54	0.07
ERattaphum (29)	-0.27	0.09	gramma:ERattaphum (29)	0.16	0.08
THaChamuang (30)	-0.21	0.11	gramma:THaChamuang (30)	0.29	0.11
KhaoPhra (31)	-0.03	0.11	gramma:KhaoPhra (31)	0.38	0.11
Sadao (32)	-0.10	0.12	gramma:Sadao (32)	0.15	0.13
Prik (33)	-0.73	0.14	gramma:Prik (33)	0.76	0.16

Table 4.4: Regression coefficients and standard errors from multiplicative model in Songkhla (continued).

Factors	Coef.	SE	Factors	Coef.	SE
NSadao (34)	-0.49	0.13	gramma:NSadao (34)	-0.01	0.13
SamNakTaeo (35)	-0.53	0.14	gramma:SamNakTaeo (35)	-0.12	0.15
PadangBesa (36)	0.02	0.09	gramma:PadangBesa (36)	0.42	0.09
SamnakKham (37)	-0.06	0.13	gramma:SamnakKham (37)	-0.15	0.13
HatYai (38)	-0.57	0.11	gramma:HatYai (38)	0.17	0.10
KhuanLang (39)	-0.23	0.09	gramma:KhuanLang (39)	0.30	0.08
NHatYai+NNamom (40)	-0.04	0.09	gramma:NHatYai+NNamom (40)	0.45	0.08
KhoHong (41)	-0.26	0.09	gramma:KhoHong (41)	0.36	0.08
KhoHae+KhongUtapao (42)	0.04	0.11	gramma:KhoHae+KhongUtapao(42)	-0.15	0.10
Chalung (43)	-0.87	0.22	gramma:Chalung (43)	-0.19	0.24
TungTamSao+WKhk (44)	-0.30	0.12	gramma:TungTamSao+WKhk (44)	0.31	0.12
EKhk+SHatyai+SNamom(45)	-0.29	0.10	gramma:EKhk+SHatyai+SNamom(45)	0.26	0.10
KhuanNiang (46)	-0.62	0.11	gramma:KhuanNiang (46)	0.10	0.11
NBangklam (47)	0.12	0.14	gramma:NBangklam (47)	0.23	0.15
ThaChang (48)	-0.18	0.11	gramma:ThaChang (48)	0.36	0.10
NSingHaNakorn (49)	0.54	0.09	gramma:NSingHaNakorn (49)	0.62	0.08
SathingMo (50)	-0.22	0.11	gramma:SathingMo (50)	0.74	0.12
HuaKhao (51)	0.27	0.11	gramma:HuaKhao (51)	0.32	0.10
MuangNgam (52)	0.14	0.11	gramma:MuangNgam (52)	0.17	0.11
Null deviance: 1865.68 df 207			Residual deviance: 277.89 df 104		

Figure 4.5 shows plots of observed counts and percentages against fitted values as well as residuals against normal quantiles in its upper panels, with a scatter plot of the non-participation rates versus disparity indexes for the regions in the lower panel.

The lower panel shows a scatter plot of the non-participation rates and the disparity indexes for the 52 statistical regions in Songkhla province. On the x-axis, a score of zero for the disparity index would mean that, for the particular region, all four gender-religion groups would have the same non-participation levels, a positive value indicates greater non-participation rates for Muslim youth, and a negative value indicates greater non-participation rates for non-Muslims. On the y-axis, a placement near the dotted line would indicate that, for that particular region, the overall level of

non-participation would be close to the overall average for all regions. Regions with points above or below the dotted line have non-participation rates that are higher or lower, respectively, than average.

The plot also shows confidence intervals that can be used to make valid statistical conclusions, The vertical line segments denote 95 percent confidence intervals for the non-participation rates (when compared with the overall mean), whereas the horizontal line segments denote 95 percent confidence intervals for the disparity indexes (compared to zero disparity). To reduce clutter, the confidence intervals are shown only for selected regions. Using these confidence intervals, for each of the two outcome measures, the regions may be classified into nine groups according to whether (a) the confidence interval for the non-participation percentage is wholly above the mean, crosses the mean, or is wholly below the mean, and (b) whether the confidence interval for the disparity index is wholly above 0, contains 0, or is below 0. Thus, for region 49 the confidence interval for its non-participation percentage is wholly above the overall mean non-participation percentage for all regions, and the confidence interval for its disparity index is wholly greater than 0, so the region is classified as higher than average non participation and Muslim disadvantage with respect to disparity. On the other hand, for region 10 the confidence interval for its non-participation percentage is wholly below the overall mean non-participation percentage and the confidence interval for its disparity index is wholly greater than 0, so this region is classified as lower than average non-participation and Muslim disadvantage with respect to disparity. For region 3, the confidence interval for the non-participation percentage crosses the overall mean and the confidence interval for

the disparity index contains 0, so this region is classified as average with respect to both non-participation and disparity.

This classification thus gives rise to nine possible clusters as follows.

1: Higher than average non participation, Muslim disadvantage (7 regions).

2: Higher than average non participation, no evidence of disparity (3 regions).

3: Higher than average non participation, Muslim advantage (no region).

4: Average non-participation, Muslim disadvantage (21 regions).

5: Average non-participation, no evidence of disparity (7 regions).

6: Average non-participation, Muslim advantage (1 region).

7: Lower than average non-participation, Muslim disadvantage (4 regions).

8: Lower than average non-participation, no evidence of disparity (9 regions).

9: Lower than average non-participation, Muslim advantage (no region).

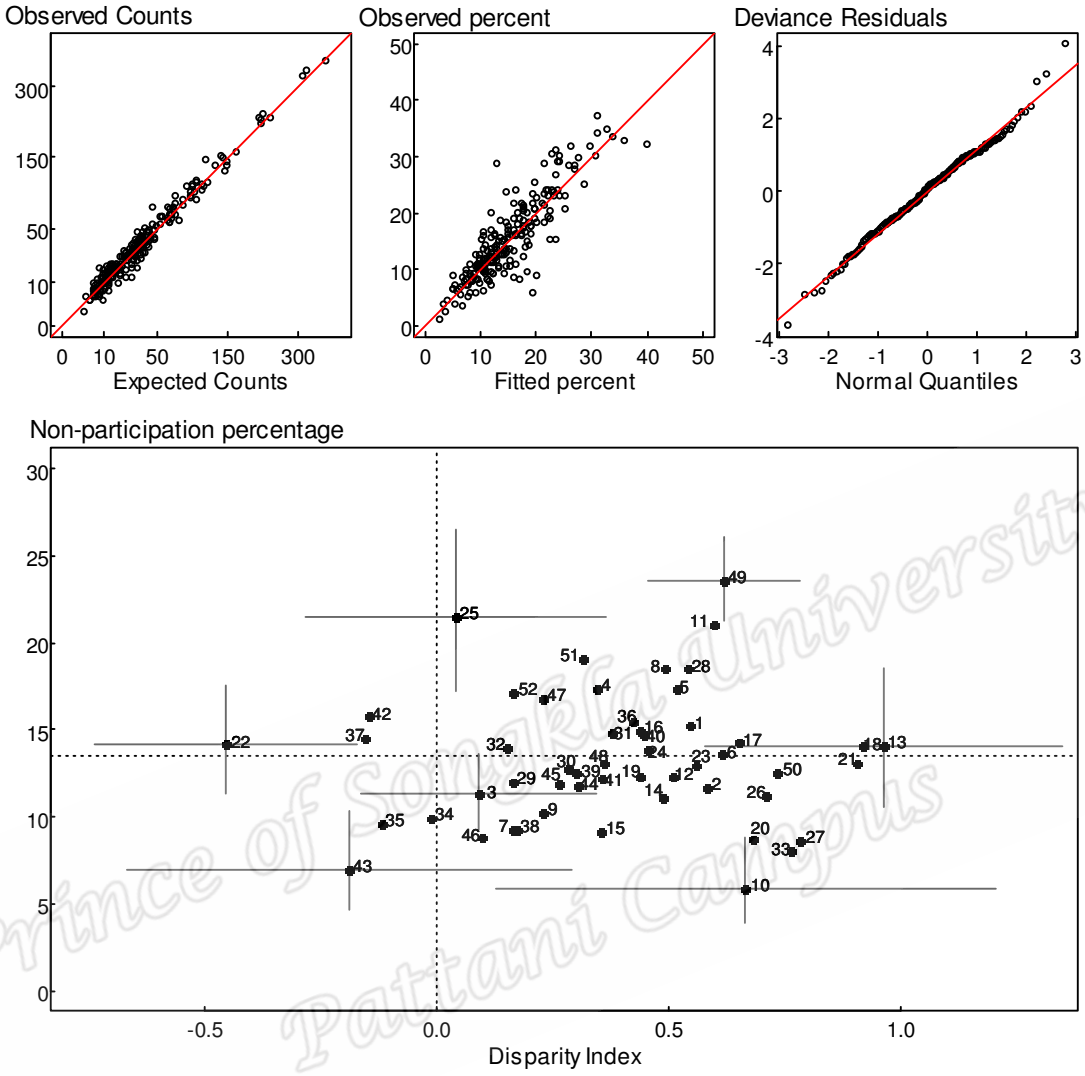

Figure 4.5: Multiplicative model results for youth non-participation rates in Songkhla regions

Figure 4.6 shows a corresponding thematic map, where disparity is labeled as MD for Muslim disadvantage and MA for Muslim advantage.

In seven regions (Phawong+Koyo (4), North SatingPa+Krasasin (5), Paching (8), ThaMoSai (11), Ranot (28), North SinghaNakorn (49) and HuaKhao (51)) Muslims had higher than average non-participation and disparity disadvantage. In three regions (BanNot (25), North Bangklam (47) and HuaKhao (51)) Muslims had higher than average non-participation but no evidence of disparity. Muslims in the Sakom (22)

region had average non-participation and a high negative disparity index (i.e. a disparity advantage).

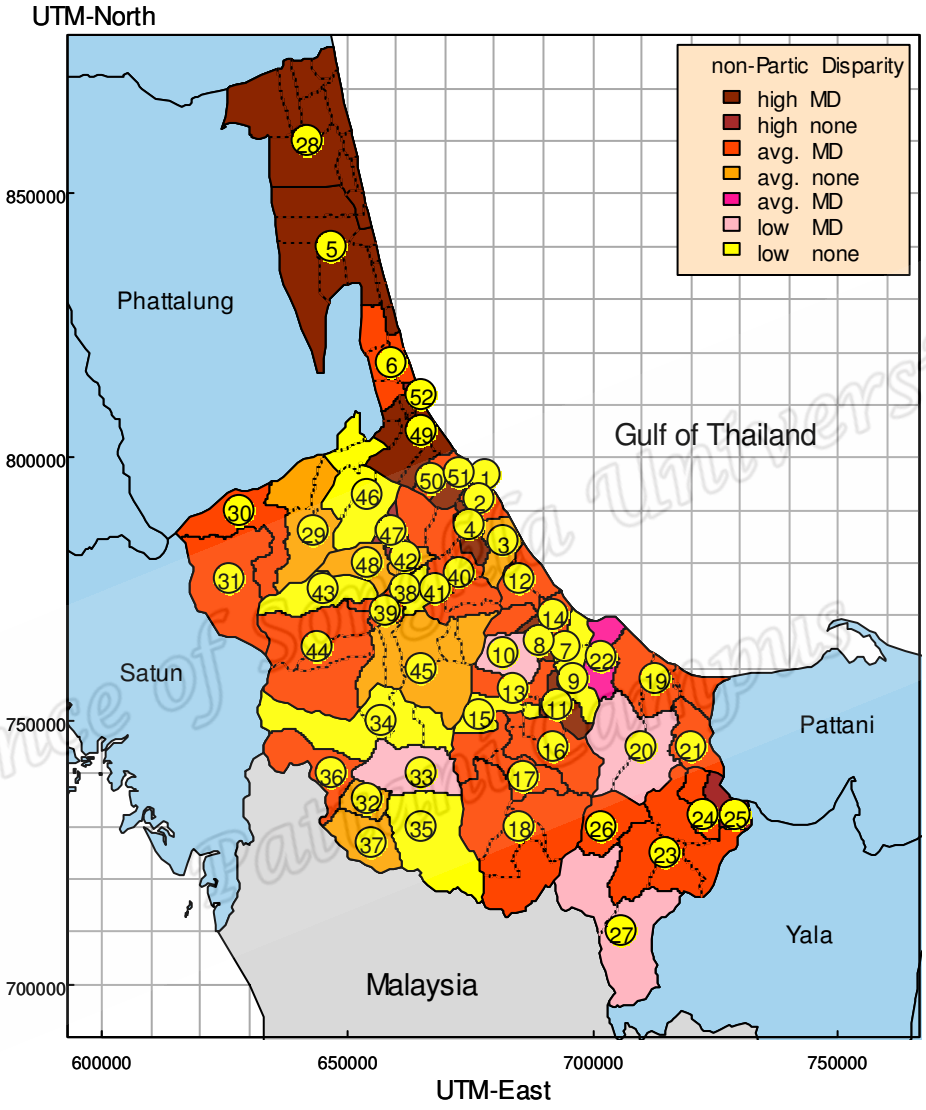

Figure 4.6: Map of multiplicative model for youth non-participation rates in Songkhla regions