CPH 931 — Fall 2008 — Dr. Charnigo

Written Assignment 3 Solutions

1a. Parameter estimates, standard errors, and p-values are as follows.

Parameter	Estimate	Standard Error	p-value
α	-1.7753	0.4472	<0.0001
β_1	0.9781	0.4118	0.0175
β_2	0.8898	0.5258	0.0906
β_3	-0.2151	0.4127	0.6022
β_4	-0.7282	0.3778	0.0539
α^*	-1.2663	0.3779	0.0008
β_1^*	0.2866	0.3286	0.3831
β_2^*	-1.7932	1.0462	0.0865
β_3^*	0.1202	0.3762	0.7494
β_4^*	-0.0587	0.3425	0.8639

1b. To answer this question we test $H_0 : \beta_2 = \beta_2^* = 0$ against the complementary alternative. We obtain $\chi^2 = 6.54$ with p-value 0.0380, so $H_0 : \beta_2 = \beta_2^* = 0$ is rejected and we cannot reasonably remove SMK from the model.

1c. Such a woman has $x_1 = x_2 = x_3 = 0$ and $x_4 = 1$. We calculate

$$\frac{\widehat{p}_{1|\mathbf{x}}}{\widehat{p}_{0|\mathbf{x}}} = \exp[-1.7753 - 0.7282] = 0.0818$$

and

$$\frac{\widehat{p}_{2|\mathbf{x}}}{\widehat{p}_{0|\mathbf{x}}} = \exp[-1.2663 - 0.0587] = 0.2658,$$

from which we obtain

$$1 + \frac{\widehat{p}_{1|\mathbf{x}}}{\widehat{p}_{0|\mathbf{x}}} + \frac{\widehat{p}_{2|\mathbf{x}}}{\widehat{p}_{0|\mathbf{x}}} = 1 + 0.0818 + 0.2658 = 1.3476 = \frac{1}{\widehat{p}_{0|\mathbf{x}}}.$$

Finally, $\widehat{p}_{0|\mathbf{x}} = 1/1.3476 = 0.7421$, $\widehat{p}_{1|\mathbf{x}} = 0.7421 \times 0.0818 = 0.0607$, and $\widehat{p}_{2|\mathbf{x}} = 0.7421 \times 0.2658 = 0.1972$, so that the estimated probability of adenosquamous SUBTYPE for such a woman is 0.0607.

1d. Parameter estimates, standard errors, and p-values are as follows.

Parameter	Estimate	Standard Error	p-value
α	-1.5178	0.2947	<0.0001
α^*	0.2782	0.2784	0.3177
β_1	0.3150	0.2439	0.1965
β_2	0.3062	0.3905	0.4329
β_3	0.4421	0.2733	0.1058
β_4	-0.7689	0.2503	0.0021

1e. Such a woman has $x_1 = x_2 = x_3 = 0$ and $x_4 = 1$. We calculate

$$\widehat{O}_{2|\mathbf{x}} = \exp[-1.5178 - 0.7689] = 0.1016$$

and

$$\widehat{O}_{12|\mathbf{x}} = \exp[0.2782 - 0.7689] = 0.6122,$$

from which we obtain

$$\widehat{p}_{2|\mathbf{x}} = \frac{\widehat{O}_{2|\mathbf{x}}}{1+\widehat{O}_{2|\mathbf{x}}} = \frac{0.1016}{1+0.1016} = 0.0922,$$

$$\widehat{p}_{1|\mathbf{x}} = \left(\widehat{p}_{1|\mathbf{x}} + \widehat{p}_{2|\mathbf{x}}\right) - \widehat{p}_{2|\mathbf{x}} = \left(\frac{\widehat{O}_{12|\mathbf{x}}}{1+\widehat{O}_{12|\mathbf{x}}}\right) - \widehat{p}_{2|\mathbf{x}} = \frac{0.6122}{1+0.6122} - 0.0922 = 0.2875,$$

and

$$\widehat{p}_{0|\mathbf{x}} = 1 - \widehat{p}_{1|\mathbf{x}} - \widehat{p}_{2|\mathbf{x}} = 1 - 0.0922 - 0.2875 = 0.6203,$$

so the estimated probability of well differentiated GRADE for such a woman is 0.6203.

2a. Letting X_1 denote TIME and using otherwise obvious notation, we find that the model-based estimate of the incidence rate ratio comparing day $x_1 + 10$ to day x_1 is

$$\frac{\exp[\widehat{\alpha} + \widehat{\beta}_1(x_1 + 10)]}{\exp[\widehat{\alpha} + \widehat{\beta}_1 x_1]} = \exp[10\widehat{\beta}_1] = \exp[0.136] = 1.146.$$

This estimate does not depend on x_1, so we obtain 1.146 both in comparing day 55 to day 45 and in comparing day 85 to day 75.

2b. [Display the plot of DAILYINF versus TIME.] The answers to part a do not seem reasonable:

• Daily incidence is between 0 and 4 from day 40 to day 50 but between 10 and 25 from day 51 to day 60. Ideally we would have a model-based estimate of no less than $10/4 = 2.500$ for comparing day 55 to day 45.

• Daily incidence is between 11 and 25 from day 70 to day 82 but between 0 and 9 from day 83 to day 90. Ideally we would have a model-based estimate of no more than $9/11 = 0.818$ for comparing day 85 to day 75.

2c. Using obvious notation, the model-based estimate of the incidence rate ratio comparing day $x_1 + 10$ to day x_1 is

$$\frac{\exp[\widehat{\alpha} + \widehat{\beta}_1(x_1 + 10) + \widehat{\beta}_2(x_1 + 10)^2]}{\exp[\widehat{\alpha} + \widehat{\beta}_1 x_1 + \widehat{\beta}_2 x_1^2]} = \exp[10\widehat{\beta}_1 + 100\widehat{\beta}_2 + 20\widehat{\beta}_2 x_1] = \exp[2.611 - 0.042x_1].$$

With $x_1 = 45$ we obtain $\exp[0.721] = 2.056$ and with $x_1 = 75$ we obtain $\exp[-0.539] = 0.583$. These answers are much more reasonable than those to part a, though even the quadratic model understates the escalation in incidence around day 51.

2d. The quantity

$$\exp[10\beta_1 + 100\beta_2 + 20\beta_2 x_1]$$

is free of x_1 if and only if $\beta_2 = 0$, so we must test $H_0 : \beta_2 = 0$. We have $\chi^2 = 291.87$ with p-value less than 0.0001, so $H_0 : \beta_2 = 0$ is rejected and we declare significantly different the 2.056 and 0.583.

2e. With a multiplicative adjustment of $\sqrt{2.4221}$ to the standard errors in Poisson regression, we still reject $H_0 : \beta_2 = 0$ because $\chi^2 = 120.51$ with p-value less than 0.0001. With negative binomial regression, we reject $H_0 : \beta_2 = 0$ because $\chi^2 = 107.02$ with p-value less than 0.0001.