**Guannan Shen HW7**

**Question: how do air pollution effects compare with the common cold? How can you manipulate the slope values so that you have an apples-to-apples comparison?**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Solutions for Fixed Effects** |  |  |  |  |  |  |
| **Effect** | **asthma** | **Estimate** | **Standard**  **Error** | **DF** | **t Value** | **Pr > |t|** |
| **Intercept** | 5 | -4.7495 | 0.6584 | 57 | -7.21 | <.0001 |
| **Intercept** | 4 | -3.8535 | 0.6288 | 57 | -6.13 | <.0001 |
| **Intercept** | 3 | -3.0251 | 0.6165 | 57 | -4.91 | <.0001 |
| **Intercept** | 2 | -2.2276 | 0.6107 | 57 | -3.65 | 0.0006 |
| **l1mmaxpm25** |  | 0.002118 | 0.005429 | 1220 | 0.39 | 0.6966 |
| **temperature** |  | -0.00497 | 0.009680 | 1220 | -0.51 | 0.6081 |
| **humidity** |  | -0.00260 | 0.006389 | 1220 | -0.41 | 0.6843 |
| **l1cold** |  | 1.8464 | 0.2205 | 1220 | 8.37 | <.0001 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Label** | **Mean** | **Std Dev** | **Lower Quartile** |
| l1mmaxpm25  temperature  humidity  l1cold | temp  humidity | 22.5505703  41.8995519  38.7578728  0.1574746 | 17.2674677  11.4984365  16.4713041  0.3643799 | 12.0000000  33.3333333  27.9583333  0 |

PROC GLIMMIX DATA= asthma method= RMPL;

class id;

model asthma(desc) = l1mmaxpm25 temperature humidity l1cold/ solution dist = multinomial link= cumlogit;

random id;

run;

PROC MEANS DATA = asthma MEAN STD Q1;

var l1mmaxpm25 temperature humidity l1cold;

where asthma^=. and l1cold^=.;

title3 'Step 3: Selecting Statistics';

run;

The model is ordinal logistic regression with random intercepts for subjects. And the cumulative odds were calculated for P(Y ≥ k), where k = 1, 2, 3, 4, and 5. The air pollution effect (l1mmaxpm25) is not significantly associated with the asthma symptoms (p = 0.6966), while the common cold is significantly associated with the asthma symptoms (p < 0.0001). One unit increase in l1mmaxpm25 will increase P(Y ≥ k)/(1 - P(Y ≥ k)) (odds) by exp(0.002118). Similarly, with cold compared to without cold, P(Y ≥ k)/(1 - P(Y ≥ k)) will increase by exp(1.8464). Generally, both air pollution and cold will make asthma patients sicker. However, only the effect of cold is significant.

To make the effects of air pollution and cold comparable, we can use standard deviation or inter quartile range (IQR) to standardize the units. For instance, standardized effect of l1mmaxpm25 is exp((1/17.2674677) \* 0.002118)

**Determine P(*Y*=*y*) for *y*=1, 2, 3, 4 and 5. Note that your answer will depend on levels of the covariates, so use average values of the 4 covariates. What if you had used, say, the first quartile of each of the covariates? How would your answer change, generally?**

By restricting the data so that it like data used in the OLR model, I got average values and the 1st quartiles for l1mmaxpm25, temperature, humidity and l1cold. To make the patterns between the 1st quartiles and mean comparable, I used l1cold as 1 both for the average and the 1st quartiles (The l1cold is the only significant covariate, the status of cold impacts the outcome a lot).

For l1mmaxpm25, temperature, humidity and l1cold, I used 22.5505703, 41.8995519, 38.7578728, 1 as the average values and 12.0000000, 33.3333333, 27.9583333, 1 as the 1st quartile.

P(Y=y) for y=1, 2, 3, 4 and 5, using the averages: 0.65530688, 0.15313989, 0.09777481, 0.05324882 and 0.04052959.

P(Y=y) for y=1, 2, 3, 4 and 5, using the 1st quartile: 0.64431502, 0.15653928, 0.10118063, 0.05551471 and 0.04245035.

The pattern reflects the overall effects of air pollution, temperature and humidity, as all these values increase, the asthma symptoms tend to be less severe.

Code

############### l1mmaxpm25 temperature humidity l1cold ################

beta0 <- c(-2.2276, -3.0251, -3.8535, -4.7495)

# covariates

betas <- c(0.002118, -0.00497, -0.00260, 1.8464)

# predictors

mean\_sas <- c(22.5505703, 41.8995519, 38.7578728, 1)

q1\_sas <- c(12.0000000, 33.3333333, 27.9583333, 1)

## p >= 1 == 1

cum\_mean <- exp(beta0)\*exp(sum(betas\*mean\_sas) )/ (1 + exp(beta0)\*exp(sum(betas\*mean\_sas) ))

## individual p

p\_mean <- c(1, cum\_mean) - c(cum\_mean, 0)

p\_mean

sum(p\_mean)

## for q1

cum\_q1 <- exp(beta0)\*exp(sum(betas\*q1\_sas) )/ (1 + exp(beta0)\*exp(sum(betas\*q1\_sas) ))

p\_q1 <- c(1, cum\_q1) - c(cum\_q1, 0)

p\_q1

sum(p\_q1)