Example of Analysis

Panfeng Liang

June 22, 2020

## Analysis of Death Rate through Logistic Model :

fit\_full=glm(formula = survival ~ factor(gender)+factor(occupation), data = clean.data, family = binomial(link = logit))  
summary(fit\_full)

##   
## Call:  
## glm(formula = survival ~ factor(gender) + factor(occupation),   
## family = binomial(link = logit), data = clean.data)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -3.7048 0.0457 0.0697 0.1543 0.2295   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 6.8618 0.7087 9.682 < 2e-16 \*\*\*  
## factor(gender)M -0.8425 0.4295 -1.961 0.04982 \*   
## factor(occupation)OTHER -1.5945 0.8687 -1.835 0.06645 .   
## factor(occupation)PHYSICIAN -2.3956 0.7926 -3.022 0.00251 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 356.92 on 3622 degrees of freedom  
## Residual deviance: 320.03 on 3619 degrees of freedom  
## AIC: 328.03  
##   
## Number of Fisher Scoring iterations: 9

* Among all medical occupations, females are more likely to die from the infection of Covid19 according to the data.
* Among all medical occupations, the estimated conditional odds ratio between the death of Covid19 and gender is .
* Among male medical works, the estimated odds of death of Covid19 is just of that in female ones.
* Among both genders of infected medical workers, nurses are more likely to die than physicians and other medical staffs.
* Among a specific gender, the estimated odds that physicians die from Covid19 infection is times the estimated odds for nurses.
* Among a specific gender, the estimated odds that other medical staffs die from Covid19 infection is times the estimated odds for nurses.

## Analysis of Infected Cases through General Linear Model

count.data=clean.data %>% count(gender, occupation, province)  
  
fit\_infected=glm(formula = n ~ factor(gender) + factor(occupation)+factor(province), data = count.data, family = poisson)  
summary(fit\_infected)

##   
## Call:  
## glm(formula = n ~ factor(gender) + factor(occupation) + factor(province),   
## family = poisson, data = count.data)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -24.4260 -0.2459 0.0000 0.6092 17.0167   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.42376 0.35405 4.021 5.79e-05 \*\*\*  
## factor(gender)M -0.87591 0.03655 -23.964 < 2e-16 \*\*\*  
## factor(occupation)OTHER -1.04924 0.04641 -22.609 < 2e-16 \*\*\*  
## factor(occupation)PHYSICIAN -0.42457 0.03745 -11.338 < 2e-16 \*\*\*  
## factor(province)四川 -0.12328 1.06130 -0.116 0.908   
## factor(province)天津 -0.12328 1.06130 -0.116 0.908   
## factor(province)安徽 -0.99919 1.06085 -0.942 0.346   
## factor(province)山东 -0.16375 0.57021 -0.287 0.774   
## factor(province)广西 -1.42376 1.06082 -1.342 0.180   
## factor(province)江苏 0.55962 0.44320 1.263 0.207   
## factor(province)江西 -0.65422 0.79082 -0.827 0.408   
## factor(province)河南 -0.56154 0.61256 -0.917 0.359   
## factor(province)海南 -0.25677 0.79105 -0.325 0.745   
## factor(province)湖北 5.71563 0.35409 16.142 < 2e-16 \*\*\*  
## factor(province)湖南 -1.42376 1.06082 -1.342 0.180   
## factor(province)甘肃 -0.12328 1.06130 -0.116 0.908   
## factor(province)福建 -1.42376 1.06082 -1.342 0.180   
## factor(province)重庆 -1.42376 1.06082 -1.342 0.180   
## factor(province)陕西 -0.99919 1.06085 -0.942 0.346   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for poisson family taken to be 1)  
##   
## Null deviance: 13811.3 on 30 degrees of freedom  
## Residual deviance: 1277.9 on 12 degrees of freedom  
## AIC: 1419.8  
##   
## Number of Fisher Scoring iterations: 5

* According to the data, male medical works are slightly more likely to be infected by Covid19.
* In each category of occupation and location, the number of infected male medical workers is more than that of females.
* Nurses are slightly less likely to be infected by Covid19 than physicians and other medical staffs.
* Among both genders, there are more physicians infected than nurses.
* There are more other medical staffs infected than nurses.
* In Hubei province, medical workers are more likely to be infected.
* There are more medical workers infected in Hubei than in other provinces.