Quiz2

Ningyuan Wang

10/20/2020

1.

a. The residuals does not sum to zero, but the weighted residuals sum to zero.
b. Based on the result of F-test, since $p < 0.05$, reject H0 (i.e. weighted reduced model with only the intercept). According to the ANOVA table, SST is 97.019, SSR is 69.028, and SSE is 27.990, so SSE + SSR = SST. The codes for SST and SSR confirm the above results.

```
# a.
m1 = gls(dist ~ speed, data = cars, weight = varConstPower(1, form=~speed))
summary(m1)

## Generalized least squares fit by REML
##   Model: dist ~ speed
##   Data: cars
##        AIC      BIC    logLik
##   412.8352 422.1912 -201.4176
##
## Variance function:
##  Structure: Constant plus power of variance covariate
##  Formula: ~speed
##  Parameter estimates:
##     const     power
## 3.160444 1.022368
##
## Coefficients:
##                 Value Std.Error   t-value p-value
## (Intercept) -11.085378  4.052378 -2.735524  0.0087
## speed         3.484162  0.320237 10.879947  0.0000
##
##  Correlation:
##       (Intr)
## speed -0.9
##
## Standardized residuals:
##        Min         Q1        Med         Q3        Max
## -1.4520579 -0.6898209 -0.1308277  0.6375029  3.0757014
##
## Residual standard error: 0.7636833
## Degrees of freedom: 50 total; 48 residual

sum(resid(m1)) # 20.46435, ???is different if 0.76^2*48

## [1] 20.46435
```

```
myweights = (3.1604+cars$speed^1.0224)^(-2)
sum(myweights*resid(m1)) # 4.878433e-06
```

```
## [1] 4.878433e-06
```

```
# b.
m2 = lm(dist ~ speed, data = cars, weight = myweights)
m0 = lm(dist ~ 1, data = cars, weight = myweights)
anova(m0, m2)
```

```
## Analysis of Variance Table
##
## Model 1: dist ~ 1
## Model 2: dist ~ speed
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     49 97.019
## 2     48 27.990  1    69.028 118.38 1.49e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# SST, SSR
y_bar = coef(m0) # 28.60139
sst = sum(myweights*((cars$dist - y_bar)^2)) # 97.01857
ssr = sum(myweights*((fitted(m1) - y_bar)^2)) # 69.02904
```

2.

a.   The model result shows below.
b.   The results of the parameter estimates and standard errors calculated based on the Fisher scoring algorithm are same to the model results.

```
m3 <- glm(acc ~ truck + road + time + area + road*time +
road*area,family=poisson, offset=logtrav, data=dataf)
summary(m3)
```

```
##
## Call:
## glm(formula = acc ~ truck + road + time + area + road * time +
##     road * area, family = poisson, data = dataf, offset = logtrav)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.6388  -0.5644  -0.1062   0.4557   2.2073
##
## Coefficients:
##              Estimate Std. Error  z value Pr(>|z|)
## (Intercept) -13.89100    0.06673 -208.180  < 2e-16 ***
## truck2        0.09039    0.08861    1.020  0.30772
## road2         0.70600    0.08857    7.971 1.57e-15 ***
## road3         1.92451    0.09271   20.758  < 2e-16 ***
## time2         0.36547    0.11540    3.167  0.00154 **
## area2        -0.47201    0.10255   -4.603 4.17e-06 ***
```

```
## road2:time2     0.17539       0.16662      1.053  0.29252
## road3:time2     0.61055       0.19121      3.193  0.00141 **
## road2:area2     0.28262       0.14726      1.919  0.05496 .
## road3:area2    -0.63468       0.14424     -4.400 1.08e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 711.086  on 23  degrees of freedom
## Residual deviance:  18.982  on 14  degrees of freedom
## AIC: 156.06
##
## Number of Fisher Scoring iterations: 4

obsrate = acc/trav
acc1 = ifelse(acc==0, 0.5,acc)
obsrate1  = acc1/trav
mu = obsrate1
eta = log(mu)
z = eta
w = acc1

for (i in 1:5){
  mod = lm(z ~ truck + road + time + area + road*time + road*area,
           weights = w)
  eta = fitted(mod)
  mu = exp(eta)
  z = eta + (obsrate - mu)/mu
  w = trav*mu
}
coef(mod)

##  (Intercept)        truck2         road2         road3         time2
area2
## -13.89099812    0.09038607    0.70600305    1.92450739    0.36547438     -
0.47200876
##  road2:time2   road3:time2   road2:area2   road3:area2
##    0.17538832    0.61054921    0.28262247   -0.63468355

varbeta = summary(mod)$cov.unscaled
sqrt(diag(varbeta))

## (Intercept)        truck2         road2         road3         time2         area2
##  0.06672591    0.08861241    0.08856863    0.09271285    0.11540369    0.10254527
## road2:time2 road3:time2 road2:area2 road3:area2
##  0.16662365    0.19120943    0.14726351    0.14423577
```