E3AhmadSharif

AhmadSharif

2023-02-08

```
getwd()
```

```
## [1] "/home/ahmad/Desktop/tampere_sda/2022_2023/Semister_II/Statistical_Modeling"
```

```
data<-read.table("Datasets/ozone.txt", header=TRUE, sep="", dec=".")
names(data)
```

```
## [1] "rad"   "temp"  "wind"  "ozone"
```

```
rad = data$rad
temp = data$temp
wind = data$wind
```

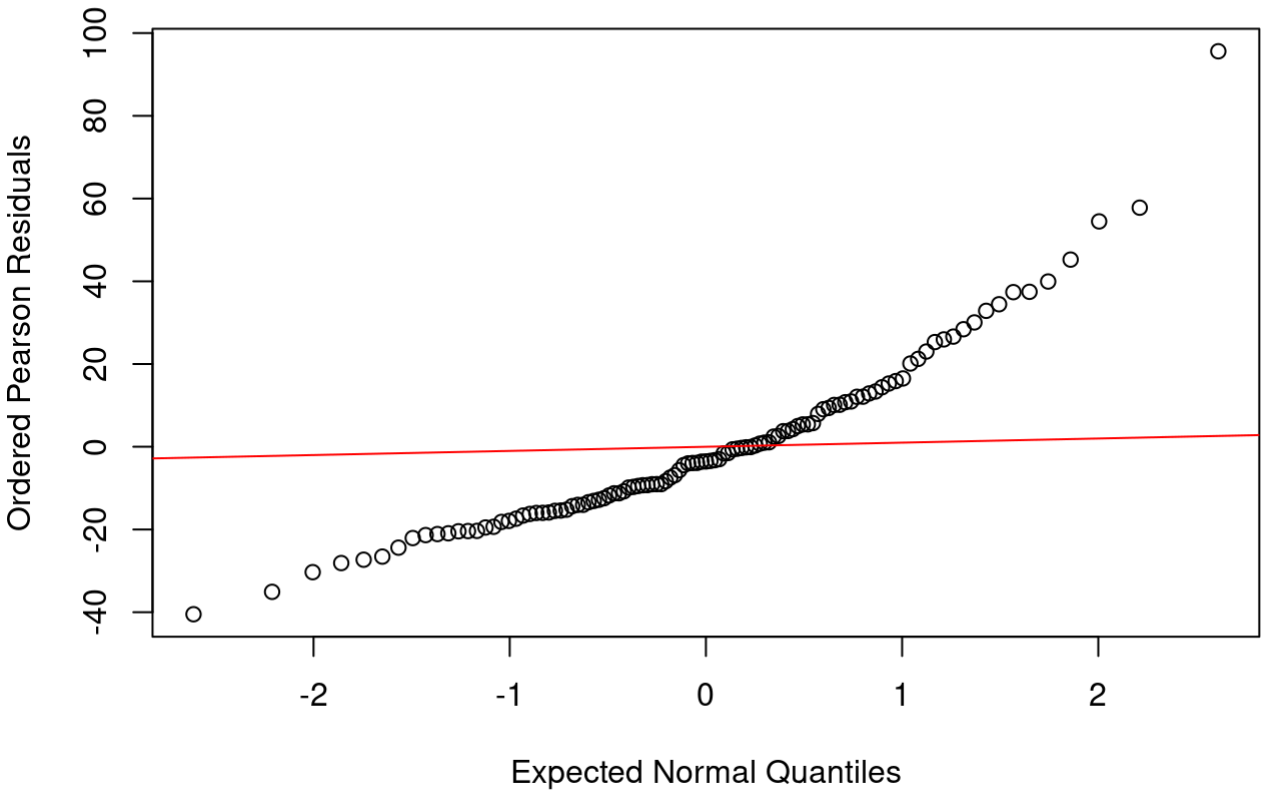
Including Plots

You can also embed plots, for example:

```
gaussian.model <- glm(data$ozone~ rad + temp + wind, family=gaussian(link="identit
y"), data=data)
pearson_residuals <- residuals(gaussian.model, type = "pearson")
pearson_residuals
```

```
##           1           2           3           4           5           6
##   7.93791864   0.98983440 -12.81334437  -0.47694390  -9.27244006  25.94974829
##           7           8           9          10          11          12
##  37.45797129 -16.63463560 -20.38291476 -14.05465969  26.63150128  -9.03510006
##          13          14          15          16          17          18
##  10.94580672  32.86085354   1.07763476   2.60081893  -0.29285957  -9.03706201
##          19          20          21          22          23          24
##  -1.61182705  30.05823714  15.87496671  10.14556711  54.47620246 -16.24489965
##          25          26          27          28          29          30
## -17.38650218  15.28124989 -21.35511066 -30.31615871  -3.60197435  34.45099416
##          31          32          33          34          35          36
##   5.39697689 -13.09937604 -22.07466848  57.79557769 -11.24438190 -20.92198465
##          37          38          39          40          41          42
##  -3.92972202 -15.21466350  -3.55647632  14.38198415  12.08043972  16.50830452
##          43          44          45          46          47          48
## -14.36489507  -3.25005085 -16.00716111 -14.03359126 -18.17412750  -9.48470248
##          49          50          51          52          53          54
##   5.41667955  12.10641339 -19.34645280  13.35118193  45.24543018 -27.30718978
##          55          56          57          58          59          60
##   9.37601377  12.88721989 -19.51778900   0.75137213   5.00166630 -12.44949229
##          61          62          63          64          65          66
## -15.89203567 -35.07311826  37.37669904  25.30708270  39.94614588  -6.87047367
##          67          68          69          70          71          72
## -21.10921318  20.12199834 -10.77926311  10.76387416 -20.43816388 -11.77282263
##          73          74          75          76          77          78
##   2.45366593  -5.66562328  -0.07993266  -1.51563817  95.60035477   9.07259974
##          79          80          81          82          83          84
##   0.30079590  21.24065914  -3.42893198   3.80354492  23.01546746  -4.43733115
##          85          86          87          88          89          90
## -17.92785313   5.72103959 -13.40550632   3.76204938 -26.55361559 -20.34024915
##          91          92          93          94          95          96
## -15.98203886 -15.41885831  10.10232182  -4.01441548 -28.11578090  -9.05899076
##          97          98          99         100         101         102
##  -8.31858308  -9.70463296   4.26543970 -15.49697986  -3.90452017 -40.48540868
##         103         104         105         106         107         108
##  -0.62321441  -0.13493503  -7.45030184 -11.25406296  28.41416682  -9.86389509
##         109         110         111
##  -9.30211732 -24.39236001  -3.00244547
```

```
plot(qnorm(ppoints(length(pearson_residuals))), sort(pearson_residuals),
     xlab = "Expected Normal Quantiles", ylab = "Ordered Pearson Residuals")
abline(0, 1, col = "red")
```



Inverse Gaussian

```
inverse_gaussian.model <- glm(data$ozone~ rad + temp + wind, family = inverse.gaussia
n(link = "identity"), data = data)
inverse_gaussian.model
```

```
##
## Call:  glm(formula = data$ozone ~ rad + temp + wind, family = inverse.gaussian(lin
k = "identity"),
##     data = data)
##
## Coefficients:
## (Intercept)          rad         temp         wind
##     -72.321        0.121        1.149        0.471
##
## Degrees of Freedom: 110 Total (i.e. Null);  107 Residual
## Null Deviance:       3.178
## Residual Deviance: 1.482     AIC: 983.4
```

```
inverse_gaussian_pearson_residuals <- residuals(inverse_gaussian.model, type = "pears
on")
```

Gamma

```
gama.model <- glm(data$ozone~ rad + temp + wind, family=Gamma(link="log"), data=data)
gama.model
```

```
##
## Call:  glm(formula = data$ozone ~ rad + temp + wind, family = Gamma(link = "log"),
##     data = data)
##
## Coefficients:
## (Intercept)          rad         temp         wind
##    0.452088     0.002103     0.043022    -0.065914
##
## Degrees of Freedom: 110 Total (i.e. Null);  107 Residual
## Null Deviance:       71.95
## Residual Deviance: 25.85     AIC: 925.9
```

```
gamma_pearson_residuals <- residuals(gama.model, type = "pearson")
```

```
summary(gaussian.model)
```

```
##
## Call:
## glm(formula = data$ozone ~ rad + temp + wind, family = gaussian(link = "identit
y"),
##     data = data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -40.485  -14.210   -3.556   10.124   95.600
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -64.23208   23.04204  -2.788  0.00628 **
## rad           0.05980    0.02318   2.580  0.01124 *
## temp          1.65121    0.25341   6.516 2.43e-09 ***
## wind         -3.33760    0.65384  -5.105 1.45e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 448.2628)
##
##     Null deviance: 121802  on 110  degrees of freedom
## Residual deviance:  47964  on 107  degrees of freedom
## AIC: 998.63
##
## Number of Fisher Scoring iterations: 2
```

```
summary(gama.model)
```

```
##
## Call:
## glm(formula = data$ozone ~ rad + temp + wind, family = Gamma(link = "log"),
##     data = data)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.70970  -0.40806  -0.09134   0.24151   1.17971
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.4520877  0.5315684   0.850  0.39696
## rad          0.0021026  0.0005347   3.933  0.00015 ***
## temp         0.0430223  0.0058461   7.359 3.98e-11 ***
## wind        -0.0659143  0.0150839  -4.370 2.89e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.2385663)
##
##     Null deviance: 71.950  on 110  degrees of freedom
## Residual deviance: 25.853  on 107  degrees of freedom
## AIC: 925.9
##
## Number of Fisher Scoring iterations: 7
```

```
summary(inverse_gaussian.model)
```

```
##
## Call:
## glm(formula = data$ozone ~ rad + temp + wind, family = inverse.gaussian(link = "id
entity"),
##     data = data)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.29570  -0.11982  -0.03303   0.04140   0.26836
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -72.32103    9.54246  -7.579 1.33e-11 ***
## rad           0.12103    0.01845   6.561 1.96e-09 ***
## temp          1.14921    0.15467   7.430 2.79e-11 ***
## wind          0.47097    0.21640   2.176   0.0317 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for inverse.gaussian family taken to be 0.01101699)
##
##     Null deviance: 3.1782  on 110  degrees of freedom
## Residual deviance: 1.4817  on 107  degrees of freedom
## AIC: 983.4
##
## Number of Fisher Scoring iterations: 10
```

ii. Yi ~ Gamma(μi, φ) Gamma is better.

ii. Yi ~ Gamma(μi, φ) Gamma is better.

c) Regardless of your solutions to (a) and (b), let us assume Yi ~ Gamma(μi, φ)

Fit the GLM model

```
gama.model <- glm(data$ozone~ rad + temp + wind, data = data, family = gaussian(link
= "log"))
coef_sum <- coef(gama.model)[2] + coef(gama.model)[3]
coef_sum
```

```
##        rad
## 0.03766163
```

```
se_sum <- sqrt(vcov(gama.model)[2,2] + vcov(gama.model)[3,3] - 2 * vcov(gama.model)
[2,3])
se_sum
```

```
## [1] 0.00576034
```

```
# Calculate the t-statistic
t_statistic <- coef_sum / se_sum
t_statistic
```

```
##      rad
## 6.538092
```

```
p_value <- 2 * (1 - pt(abs(t_statistic), df = df.residual(gama.model)))
p_value
```

```
##          rad
## 2.184854e-09
```

Reject null hypothesis as it is less

02

```
weld_data<-read.table("Datasets/weld.txt", sep="\t", dec=".", header=TRUE)
names(weld_data)
```

```
## [1] "Drying"     "Material"   "Thickness"  "Angle"      "Opening"
## [6] "Preheating" "Strength"
```

```
Strength <- weld_data$Strength
Material <- weld_data$Material
Drying <- weld_data$Drying
Drying
```

```
##  [1] 0 0 1 1 1 1 0 0 1 1 0 0 0 0 1 1
```

```
head(weld_data)
```

```
##   Drying Material Thickness Angle Opening Preheating Strength
## 1      0        0         0     0       0          0     43.7
## 2      0        1         1     1       1          1     40.2
## 3      1        1         0     0       0          0     42.4
## 4      1        0         1     1       1          1     44.7
## 5      1        1         0     0       1          1     42.4
## 6      1        0         1     1       0          0     45.9
```

```
gaussian.model <- glm(weld_data$Strength~ factor(Drying) + factor(Material), family=g
aussian(link="identity"), data=weld_data)
gaussian.model
```

```
##
## Call:  glm(formula = weld_data$Strength ~ factor(Drying) + factor(Material),
##     family = gaussian(link = "identity"), data = weld_data)
##
## Coefficients:
##       (Intercept)    factor(Drying)1  factor(Material)1
##             43.44              2.15             -3.10
##
## Degrees of Freedom: 15 Total (i.e. Null);  13 Residual
## Null Deviance:       60.8
## Residual Deviance: 3.867     AIC: 30.69
```

```
case1 <- data.frame(Drying = 0, Material = 0, Thickness = 0, Angle = 0, Opening = 0,
Preheating = 0)
case2 <- data.frame(Drying = 0, Material = 1, Thickness = 0, Angle = 0, Opening = 0,
Preheating = 0)
pred1 <- predict(gaussian.model, newdata = case1)
pred2 <- predict(gaussian.model, newdata = case2)
d_value <- pred2 - pred1
d_value
```

```
##    1
## -3.1
```