

Untitled

untitled

```
# Input the Fleming and Harrington Count Data (fhcnt) and name the covariates
fhcnt = read.table("https://biostatcenter.gwu.edu/sites/biostatcenter.gwu.edu/files/Lachin%20Files/fhcnt")
names(fhcnt)=c("id", "z1", "z2", "z3", "z4", "z5", "z6", "z7", "z8", "z9", "nevents", "fuptime")
```

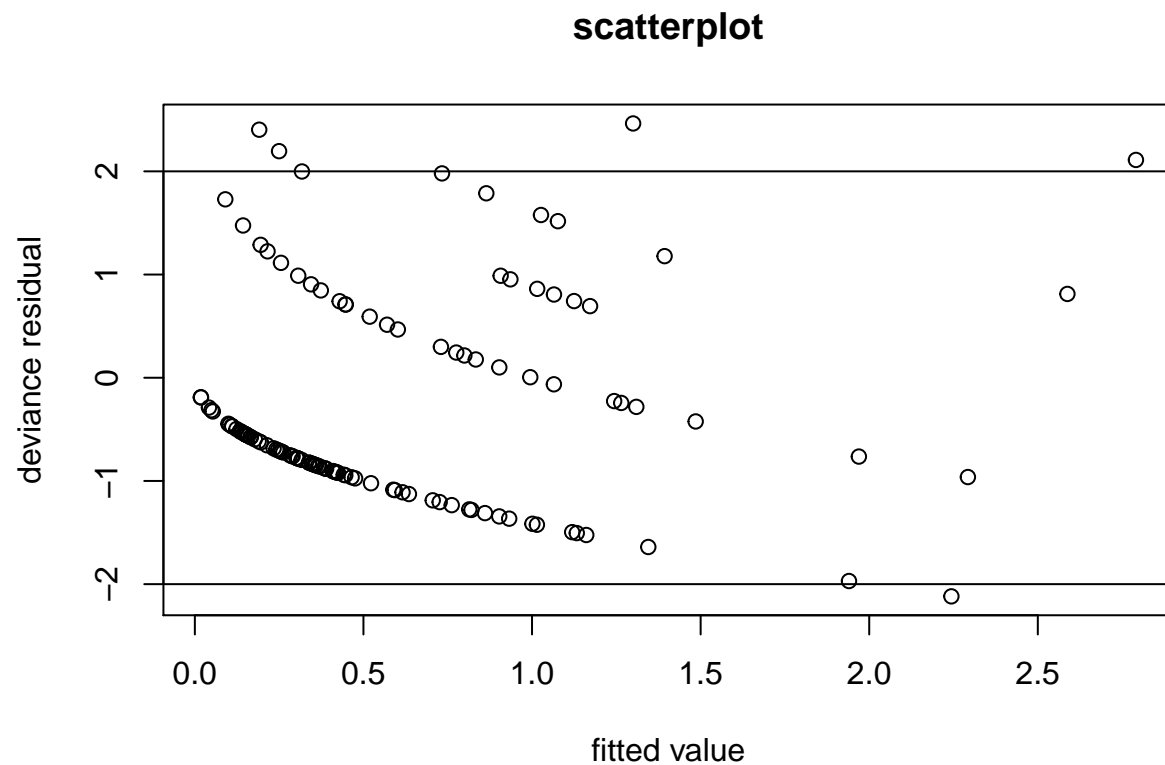
a)

```
model <- glm(nevents~as.factor(z1)+as.factor(z2)+z3+z4+z5+as.factor(z6)+as.factor(z7)+as.factor(z8)+as.factor(z9)+offset(log(fuptime)), family = poisson(link = "log"), data = fhcnt)
summary(model)
```

```
##
## Call:
## glm(formula = nevents ~ as.factor(z1) + as.factor(z2) + z3 +
##      z4 + z5 + as.factor(z6) + as.factor(z7) + as.factor(z8) +
##      as.factor(z9) + offset(log(fuptime)), family = poisson(link = "log"),
##      data = fhcnt)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1183  -0.8879  -0.5864   0.2579   2.4641
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -5.348270   1.058252  -5.054 4.33e-07 ***
## as.factor(z1)2  1.156140   0.277853   4.161 3.17e-05 ***
## as.factor(z2)2  0.754567   0.287943   2.621 0.00878 **
## z3            -0.083938   0.035200  -2.385 0.01710 *
## z4              0.007649   0.010556   0.725 0.46869
## z5              0.010345   0.016127   0.641 0.52122
## as.factor(z6)2 -1.958265   0.597411  -3.278 0.00105 **
## as.factor(z7)2  0.668546   0.343399   1.947 0.05155 .
## as.factor(z8)2 -0.856499   0.393411  -2.177 0.02947 *
## as.factor(z9)2 -0.110733   0.326942  -0.339 0.73484
## as.factor(z9)3 -0.955742   0.487365  -1.961 0.04987 *
## as.factor(z9)4 -0.788223   0.494094  -1.595 0.11065
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 179.79  on 127  degrees of freedom
## Residual deviance: 133.27  on 116  degrees of freedom
## AIC: 260.3
##
## Number of Fisher Scoring iterations: 6
```

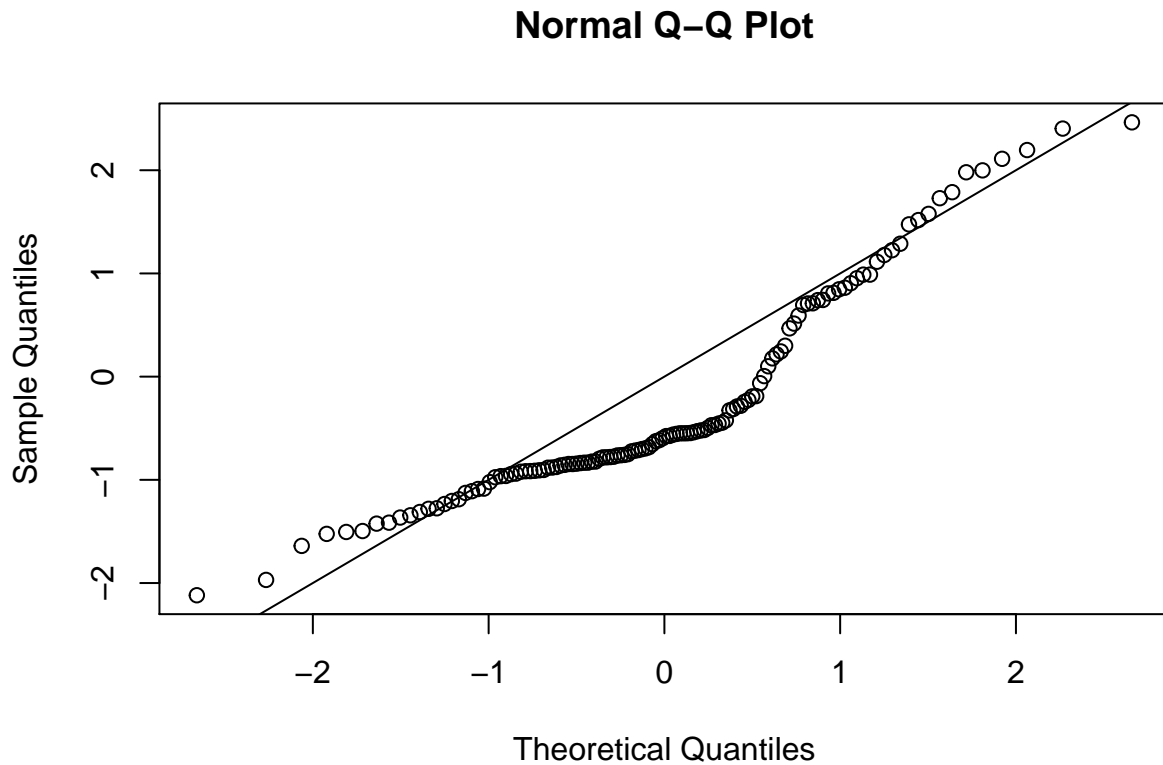
b)

```
rd <- residuals.glm(model,"deviance")
fv <- model$fitted.values
plot(fv,rd,xlab="fitted value",ylab="deviance residual",main="scatterplot")
abline(h=2)
abline(h=-2)
```



We can see the there are some points lies above the two lines, thus the model is not fitted.

```
qqnorm(rd)
abline(a=0,b=1)
```



We can see that in the middle of the graph, points are not lie along with the line, thus the model is not fitted

c)

The estimated for treatment group is 1.156140

```
exp(1.156140)
```

```
## [1] 3.177644
```

Thus the relative rate is  $3.177644 = \frac{\lambda_2}{\lambda_1}$  which  $\lambda_2$  represent take placebo and  $\lambda_1$  represent take interferon. Thus  $\lambda_2 > \lambda_1$ , take interferon actually decrease the rate. Thus treatment with interferon effective in reducing serious infections among children with CGD

```
LOG_CI = 1.156140 + c(-1,1)*1.96*0.277853
exp(LOG_CI)
```

```
## [1] 1.843283 5.477955
```

95% confidence interval of relative rate is [1.843283 ,5.477955]

d)

```
x = as.matrix(c(0,0,0,0,0,0,0,0,0,0,1,-1),ncol=1)
v = summary(model)$cov.unscaled
se<-sqrt(t(x) %*% v %*% x)
mle = -0.955742-(-0.788223)
CI <- mle + c(-1,1)*1.96*se[1]
exp(CI)
```

```
## [1] 0.2761451 2.5903447
```

Thus the 95% Confidence Interval of relative rate is [0.2761451 2.5903447]

e)

$$\hat{\mu} = \lambda \times t = \exp(\log(\hat{\lambda}) + \log(t))$$

$$\log(\hat{\lambda}) = x^T \hat{\beta}$$

```
beta = c(-5.348270,1.156140,0.754567,-0.083938,0.007649,0.010345,-1.958265,0.668546,
          -0.856499,-0.110733,-0.955742,-0.788223)
```

```
x = c(1,0,0,12,142,34,0,1,0,1,0,0)
log_hat = sum(beta*x)
mu = exp(log_hat+log(365.25))
mu
```

```
## [1] 4.668123
```

Estimated number of infections are 4.668123

f)

The purpose of this study was to find the different variables that would influence the results of a serious infections among children with chronic granulomatous disease.

The treatment group: interferon versus placebo (Odds Ratio(OR) = 3.18) and Inheritance pattern: X-linked versus autosomal recessive (OR=2.127) were associated with increased odds the response

Age (OR=0.919), Corticosteroid use on entry: yesversus no(OR=0.1411) Gender: male versus female (OR=0.4246) and Type of hospital: Amsterdam (OR = 0.3845) were associated with decreased odds the response

```
# Characteristics siginificantly associated with increased odds of response
exp(model$coeff)[model$coeff>0 & summary(model)$coeff[,4]<0.05]
```

```
## as.factor(z1)2 as.factor(z2)2
##          3.177645          2.126691
```

```
# Characteristics significantly associated with decreased odds of response
exp(model$coeff)[model$coeff<0 & summary(model)$coeff[,4]<0.05]
```

```
##      (Intercept)              z3 as.factor(z6)2 as.factor(z8)2 as.factor(z9)3
##      0.004756373      0.919487882      0.141102991      0.424646137      0.384526752
```