# GLM

*Alok Kumar*

*4 February 2018*

## 1 GLM Introduction:

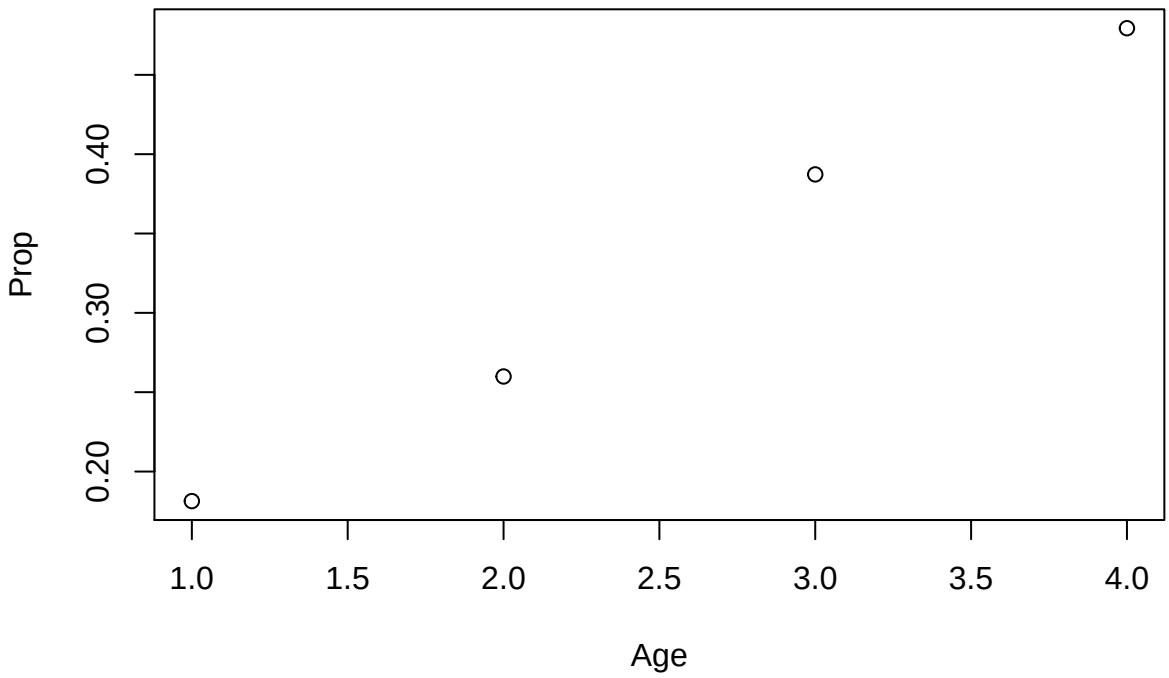
In statistics, the generalized linear model (GLM) is a flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution.

GLM have three components: 1. Distribution of Responce variable. (y ~ exponential family) 2. Linear predictor 3. Link function

## 2 Experiment Details

The data taken show the distribution of 1607 women interviewed in fiji fertility survey of 1975, classified by current age , level of education, desire for more children, and contraceptive use.

```
data = read.table(".//Data//Cont by Age.txt" , head(T))
attach(data)
Prop = Using/Total
plot(Age,Prop)
```



Converting data trom using and Total to using and not using. We are doing this for odd ratio. cbind - bind the two column and create matrix.

```
y = cbind(Using, Total-Using)
y
```

```
##      Using
## [1,]    72 325
## [2,]   105 299
## [3,]   237 375
## [4,]    93 101
```

# 3 One factor Model

GLM in which only one factor exist.

Logistic regression model for each age group using the y(parameter estimates calculated)

```
g1 = "Group 1 (age < 25)"
g2 = "Group 2 (age 25 - 29)"
g3 = "Group 3 (age 30 - 39)"
g4 = "Group 4 (age 40 - 49)"
```

```
glmOddvsAgeFactor = glm(y~factor(Age) , family = binomial(link = logit))
```

Systematic component: g(u) = log(p/1-p) = b0 + b1x1 + b2x2 + b3x3

Random component : yi | x1,x2,x3 ~ B(n,p) where x1 if Age = 2 (25-29) x2 if Age = 3 (30-39) x3 if Age = 4 (40-49)

```
summary(glmOddvsAgeFactor)
```

```
##
## Call:
## glm(formula = y ~ factor(Age), family = binomial(link = logit))
##
## Deviance Residuals:
## [1]  0  0  0  0
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.5072     0.1303 -11.571  < 2e-16 ***
## factor(Age)2   0.4607     0.1727   2.667  0.00765 **
## factor(Age)3   1.0483     0.1544   6.788 1.14e-11 ***
## factor(Age)4   1.4246     0.1940   7.345 2.06e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 7.9192e+01  on 3  degrees of freedom
## Residual deviance: 2.1405e-13  on 0  degrees of freedom
## AIC: 32.647
##
## Number of Fisher Scoring iterations: 3
```

## 3.1 Odds of using contraception for each group

Calculation logit(p) = coefficient

logit(p) = log(p/1-p) = log(odd) = coefficient

odd = exp(coefficient)

### 3.1.1 Odd interpretation for Group 1 (age < 25)

```
ageGroup1Odd = exp(glmOddvsAgeFactor$coefficients[1])
ageGroup1Odd
```

```
## (Intercept)
##   0.2215385
```

Therefore odd of someone using contraceptive in Group 1 (age < 25) is 22.1538462%.

### 3.1.2 Odd interpretation for Group 2 (age 25 - 29)

```
ageGroup2Odd = exp(glmOddvsAgeFactor$coefficients[1] + glmOddvsAgeFactor$coefficients[2])
ageGroup2Odd
```

```
## (Intercept)
##   0.3511706
```

Therefore odd of someone using contraceptive in Group 2 (age 25 - 29) is 35.1170569%.

### 3.1.3 Odd interpretation for Group 3 (age 30 - 39)

```
ageGroup3Odd = exp(glmOddvsAgeFactor$coefficients[1] + glmOddvsAgeFactor$coefficients[3])
ageGroup3Odd
```

```
## (Intercept)
##       0.632
```

Therefore odd of someone using contraceptive in Group 3 (age 30 - 39) is 63.2%.

### 3.1.4 Odd interpretation for Group 4 (age 40 - 49)

```
ageGroup4Odd = exp(glmOddvsAgeFactor$coefficients[1] + glmOddvsAgeFactor$coefficients[4])
ageGroup4Odd
```

```
## (Intercept)
##   0.9207921
```

Therefore odd of someone using contraceptive in Group 4 (age 40 - 49) is 92.0792079%.

## 3.2 Inter group comparison

### 3.2.1 Group 2 (age 25 - 29) in compare of Group 1 (age < 25)

```
oddg2vsg1 = ageGroup2Odd/ageGroup1Odd
oddg2vsg1Per = (oddg2vsg1 - 1)*100
```

Group 2 (age 25 - 29) is 58.5144928 % more likely to use contraceptive then Group 1 (age < 25)

### 3.2.2 Group 3 (age 30 - 39) in compare of Group 2 (age 25 - 29)

```
oddg3vsg2 = ageGroup3Odd/ageGroup2Odd
oddg3vsg2Per = (oddg3vsg2 - 1)*100
```

Group 3 (age 30 - 39) is 79.9695238 % more likely to use contraceptive then Group 2 (age 25 - 29)