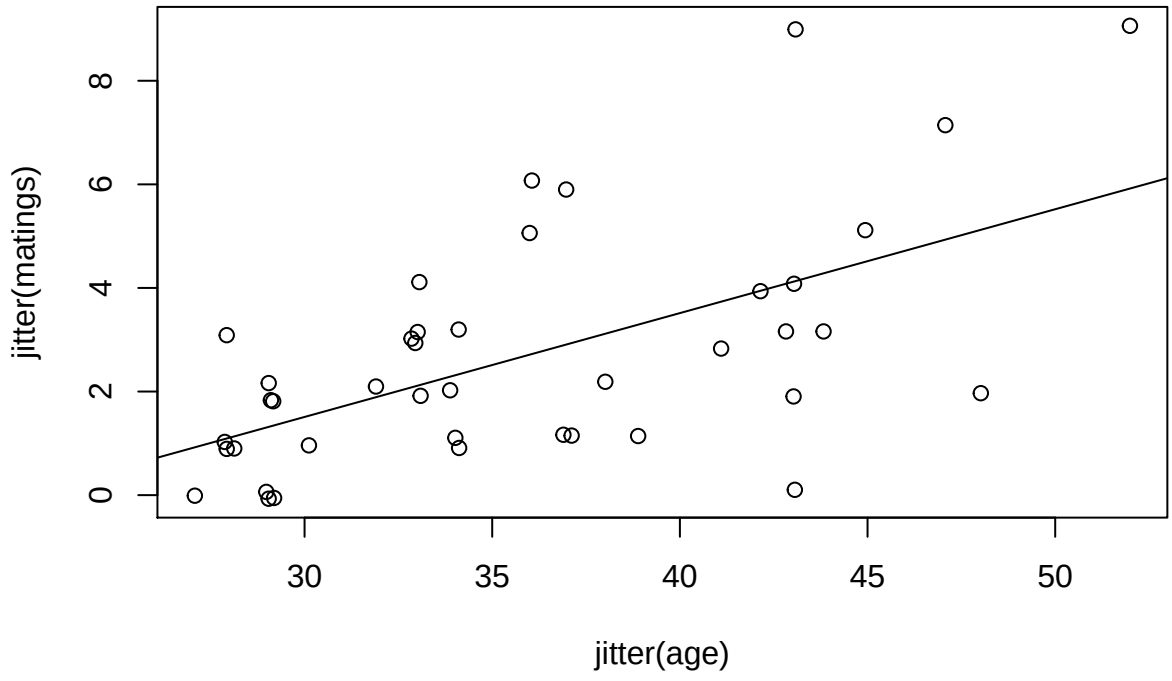
Elephants Example: Poisson Regression

```
Elephants <- read.csv("C:/hess/STAT512/RNotes/MultReg5/MR5_Elephants.csv", header=TRUE)
str(Elephants)
```

```
## 'data.frame':    41 obs. of  2 variables:
##  $ age    : int  27 28 28 28 28 29 29 29 29 29 ...
##  $ matings: int  0 1 1 1 3 0 0 0 2 2 ...
```

```
plot(jitter(matings) ~ jitter(age), data = Elephants)
abline(lm(matings ~ age, data = Elephants))
```

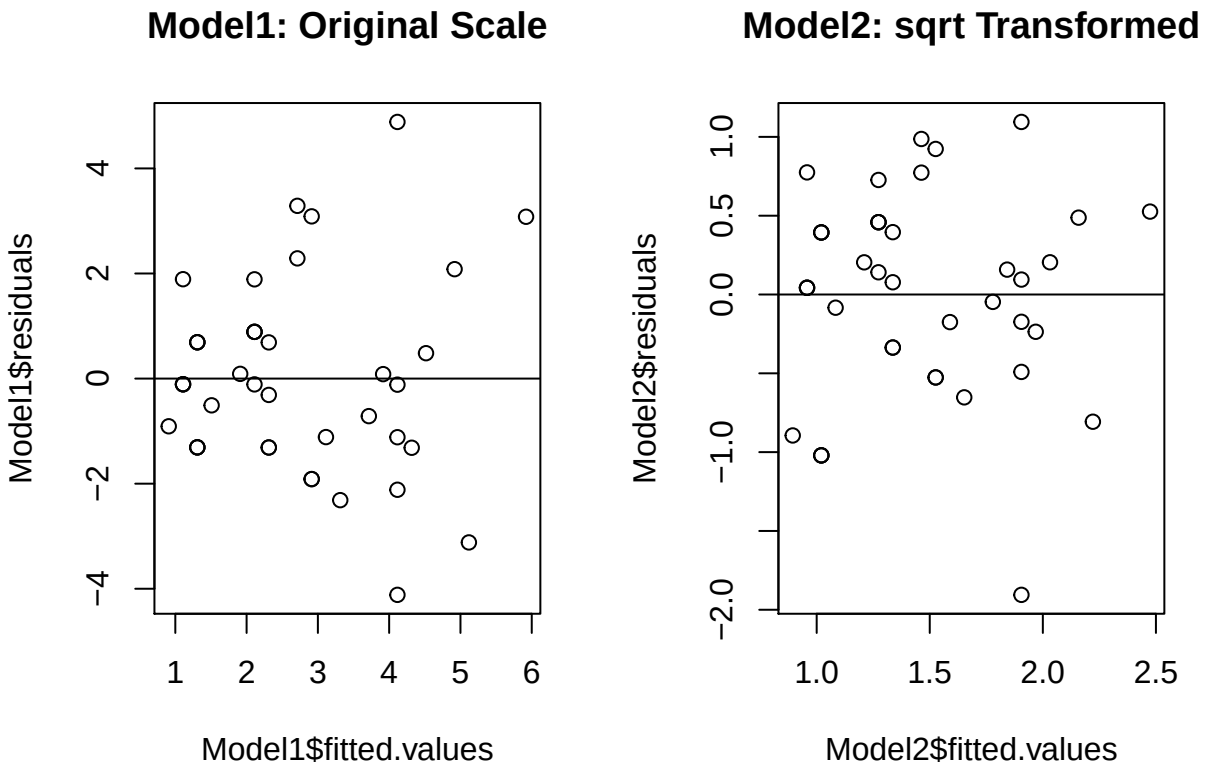


Simple Linear Regression

We try original and sqrt() transformed response.

```
Model1 <- lm(matings ~ age, data = Elephants)
Model2 <- lm(sqrt(matings) ~ age, data = Elephants)
par(mfrow = c(1, 2))
plot(Model1$residuals ~ Model1$fitted.values)
abline(h=0)
```

```r
title("Model1: Original Scale")
plot(Model2$residuals ~ Model2$fitted.values)
abline(h=0)
title("Model2: sqrt Transformed")
```



Poisson Regression with Identity link

```r
Model3 <- glm(matings ~ age, family = poisson(link = "identity"), data = Elephants)
summary(Model3)
```

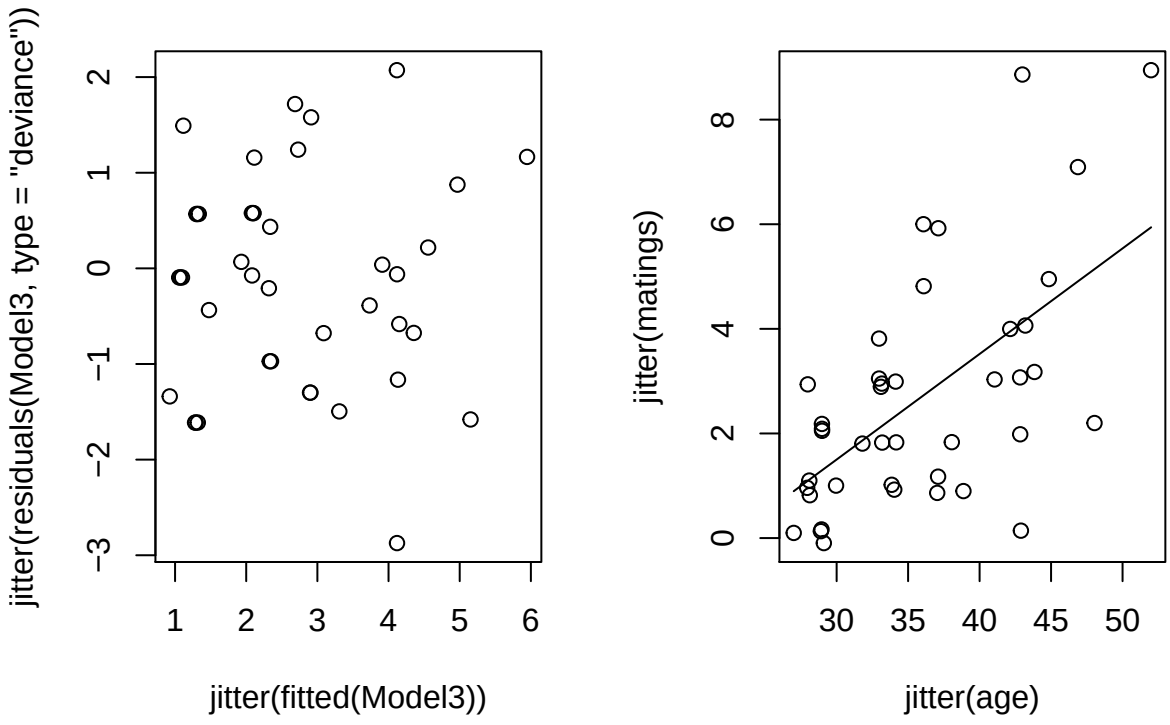
```
##
## Call:
## glm(formula = matings ~ age, family = poisson(link = "identity"),
##     data = Elephants)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.87228  -0.97171  -0.09509   0.57794   2.07192
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.55205    1.33916  -3.399 0.000676 ***
```

```
## age          0.20179    0.04023   5.016 5.29e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 75.372  on 40  degrees of freedom
## Residual deviance: 50.058  on 39  degrees of freedom
## AIC: 155.5
##
## Number of Fisher Scoring iterations: 5
```

```
#Checking for overdispersion
sum(residuals(Model3, type = "pearson")^2)/df.residual(Model3)
```

```
## [1] 1.126969
```

```
par(mfrow=c(1, 2))
plot(jitter(residuals(Model3, type = "deviance")) ~ jitter(fitted(Model3)))
#create a vector of ages for plotting
xage <- seq(27, 52, by = 0.25)
yhat <- predict(Model3, list(age = xage))
plot(jitter(matings) ~ jitter(age), data = Elephants)
lines(yhat ~ xage)
```



Poisson Regression with Log link

The default for "predict" is the linear predictor in the scale that is transformed by the link function. Transform back to original scale using type="response.

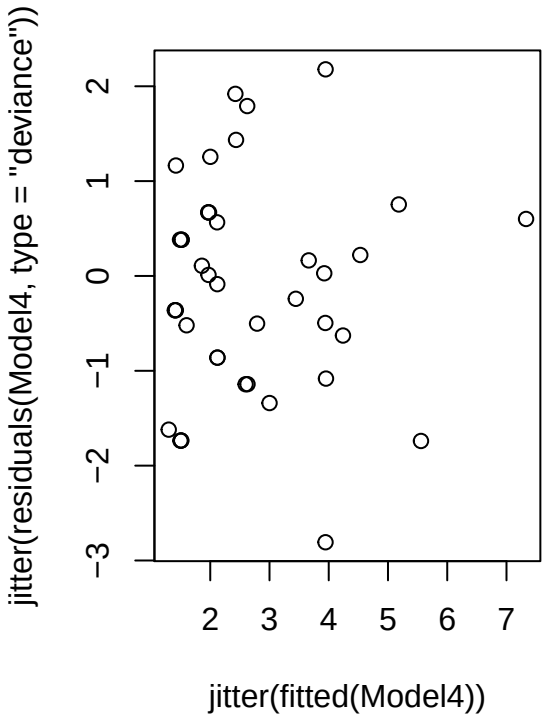
```
Model4 <- glm(matings ~ age, family = poisson(link = "log"), data = Elephants)
summary(Model4)
```

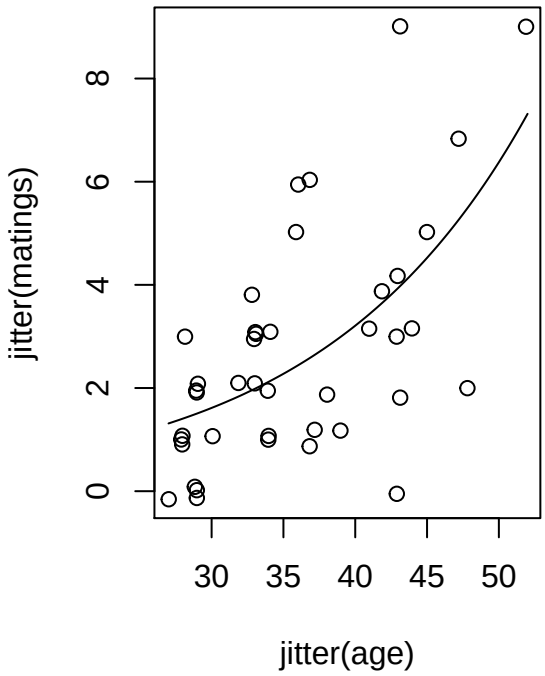
```
##
## Call:
## glm(formula = matings ~ age, family = poisson(link = "log"),
##     data = Elephants)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.80798  -0.86137  -0.08629   0.60087   2.17777
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.58201    0.54462  -2.905  0.00368 **
## age          0.06869    0.01375   4.997 5.81e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 75.372  on 40  degrees of freedom
## Residual deviance: 51.012  on 39  degrees of freedom
## AIC: 156.46
##
## Number of Fisher Scoring iterations: 5
```

```
#Checking for overdispersion
sum(residuals(Model4, type = "pearson")^2)/df.residual(Model4)
```

```
## [1] 1.157334
```

```
par(mfrow = c(1, 2))
plot(jitter(residuals(Model4, type="deviance")) ~ jitter(fitted(Model4)))
yhat <- predict(Model4, list(age = xage), type = "response")
plot(jitter(matings) ~ jitter(age), data = Elephants)
lines(yhat ~ xage)
```


jitter(residuals(Model4, type = "deviance"))
−3
−2
−1
0
1
2
2
3
4
5
6
7
jitter(fitted(Model4))


jitter(matings)
0
2
4
6
8
30
35
40
45
50
jitter(age)