

# ConversionIndex

*Jordi Valero and Marta Pérez Casan*

*11/20/2018*

```
setwd("~/Desktop/PiE2")  
library(car)
```

```
## Loading required package: carData
```

```
library(tables)
```

```
## Loading required package: Hmisc
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
## Loading required package: ggplot2
```

```
##
```

```
## Attaching package: 'Hmisc'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      format.pval, units
```

```
library(RcmdrMisc)
```

```
## Loading required package: sandwich
```

```
##
```

```
## Attaching package: 'RcmdrMisc'
```

```
## The following object is masked from 'package:Hmisc':
```

```
##
```

```
##      Dotplot
```

```
library(car)
```

```
#library(Hmisc)
```

```
#library(lattice)
```

```
#library(survival)
```

```
#library(Formula)
```

```
#library(colorspace)
```

```
#library(ggplot2)
```

```
library(emmeans)
```

```
## NOTE: As of emmeans versions > 1.2.3,
```

```
##      The 'cld' function will be deprecated in favor of 'CLD'.
```

```
##      You may use 'cld' only if you have package:multcomp attached.
```

```
library(multcompView)
```

We first read the dataset

```
conversionindex <- read.csv2("./Dades/CI.csv")
head(conversionindex)
```

```
##   SWEET LOT      CI
## 1   D00   1 2.099792
## 2   D00   6 2.192982
## 3   D00  12 2.152174
## 4   D00  19 2.099476
## 5   D00  25 2.048093
## 6   D08   4 2.050633
```

```
summary(conversionindex)
```

```
##   SWEET      LOT      CI
## D00:5   Min.    : 1   Min.    :1.893
## D08:5   1st Qu.: 7   1st Qu.:1.979
## D15:5   Median :13   Median :2.051
## D20:5   Mean    :13   Mean    :2.050
## D30:5   3rd Qu.:19   3rd Qu.:2.103
##          Max.    :25   Max.    :2.211
```

```
dim(conversionindex)
```

```
## [1] 25  3
```

We have just one explanatory variable which is the sweetener dose (categorical variable with 5 levels).

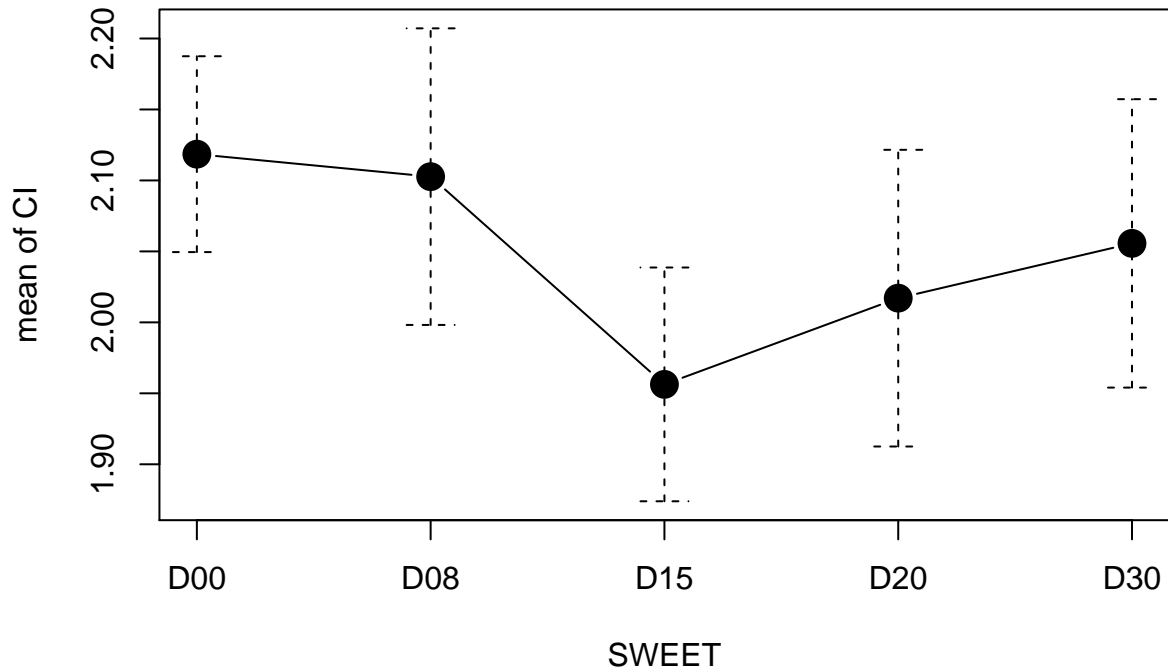
## Descriptive Statistics

```
cv<-function(x) {sd(x)/mean(x)}
tabular(CI*SWEET~((n=1)+mean+sd+cv),conversionindex)
```

	SWEET	n	mean	sd	cv
CI	D00	5	2.119	0.05557	0.02623
	D08	5	2.103	0.08415	0.04002
	D15	5	1.956	0.06632	0.03390
	D20	5	2.017	0.08418	0.04173
	D30	5	2.056	0.08183	0.03981

```
with(conversionindex,plotMeans(CI,SWEET,error.bars="conf.int",level=0.95))
```

## Plot of Means



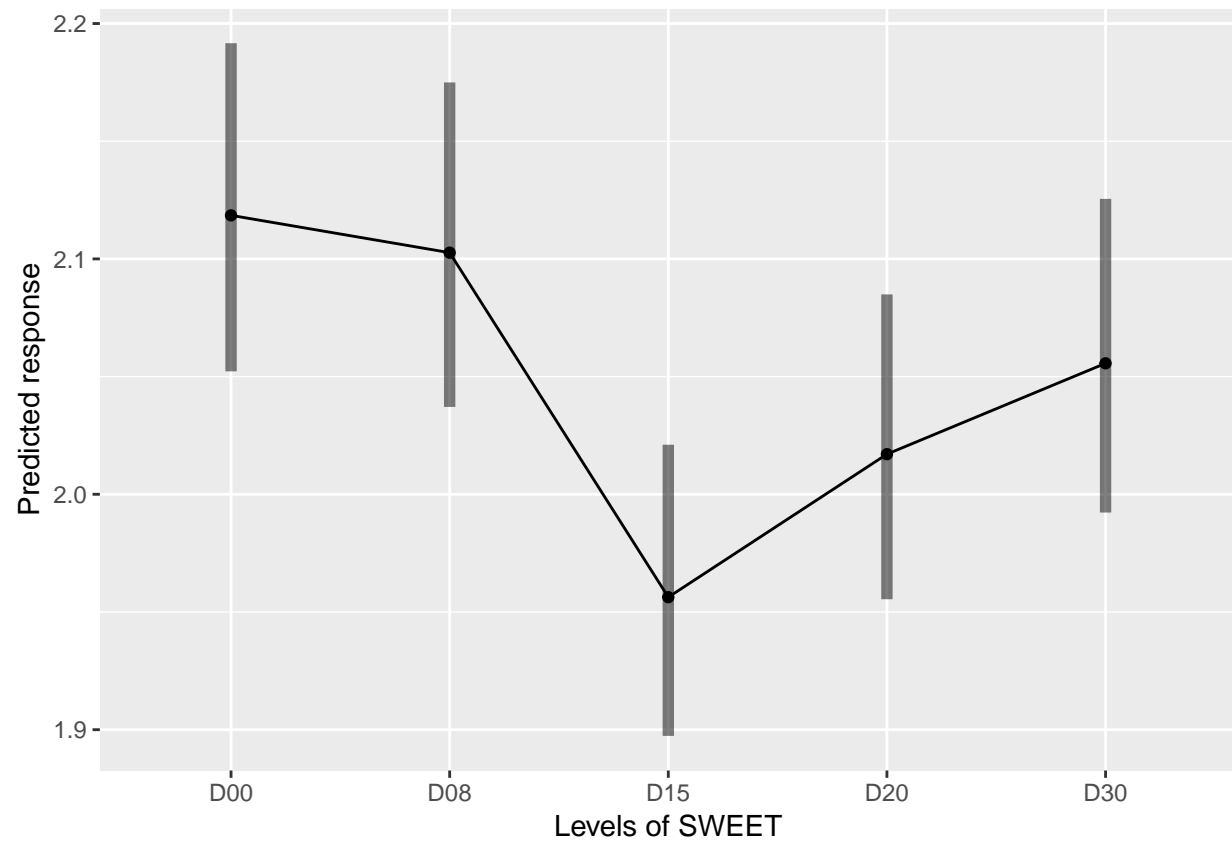
### Model 1: GLM Inverse Gaussian with canonical link

```
m1<-glm(CI~SWEET,family=inverse.gaussian,data=conversionindex)
summary(m1)
```

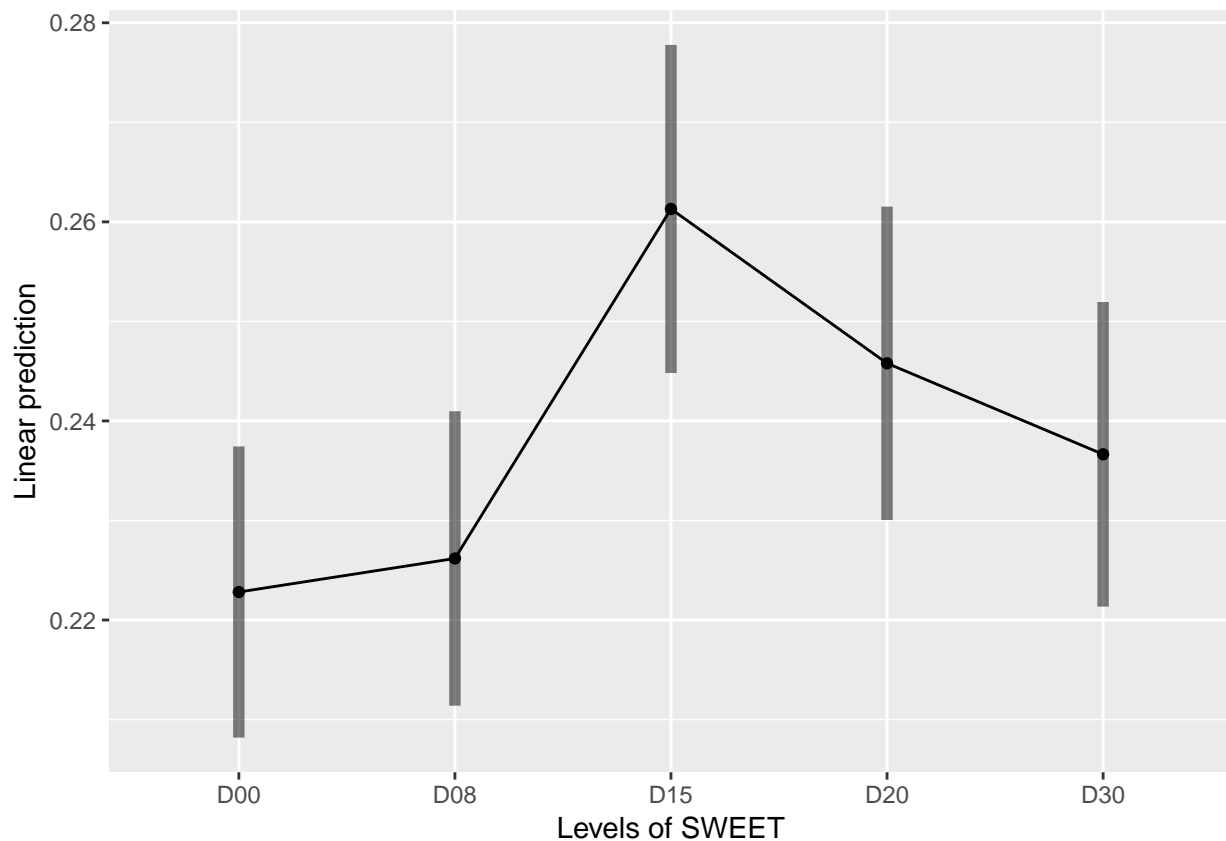
```
##
## Call:
## glm(formula = CI ~ SWEET, family = inverse.gaussian, data = conversionindex)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.042407  -0.021036  -0.000818   0.016905   0.034593
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.222814   0.007462  29.861 < 2e-16 ***
## SWEETD08     0.003370   0.010612   0.318  0.75408
## SWEETD15     0.038484   0.011242   3.423  0.00269 **
## SWEETD20     0.022978   0.010963   2.096  0.04901 *
## SWEETD30     0.013835   0.010799   1.281  0.21479
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for inverse.gaussian family taken to be 0.0006617132)
##
## Null deviance: 0.023507  on 24  degrees of freedom
## Residual deviance: 0.013303  on 20  degrees of freedom
## AIC: -51.752
```

```
##  
## Number of Fisher Scoring iterations: 3
```

```
emmip(m1, ~SWEET, CIs=T, type="response")
```



```
emmip(m1, ~SWEET, CIs=T)
```



```
summary(m1)$family
```

```
##
## Family: inverse.gaussian
## Link function: 1/mu^2
```

```
logLik(m1)
```

```
## 'log Lik.' 31.87621 (df=6)
```

```
scale<-sqrt(summary(m1)$disp)
```

### Residual calculus and plots

```
sum(residuals(m1,type="pearson")^2)
```

```
## [1] 0.01323407
```

```
sum(residuals(m1,type="pearson")^2)/m1$df.residual
```

```
## [1] 0.0006617036
```

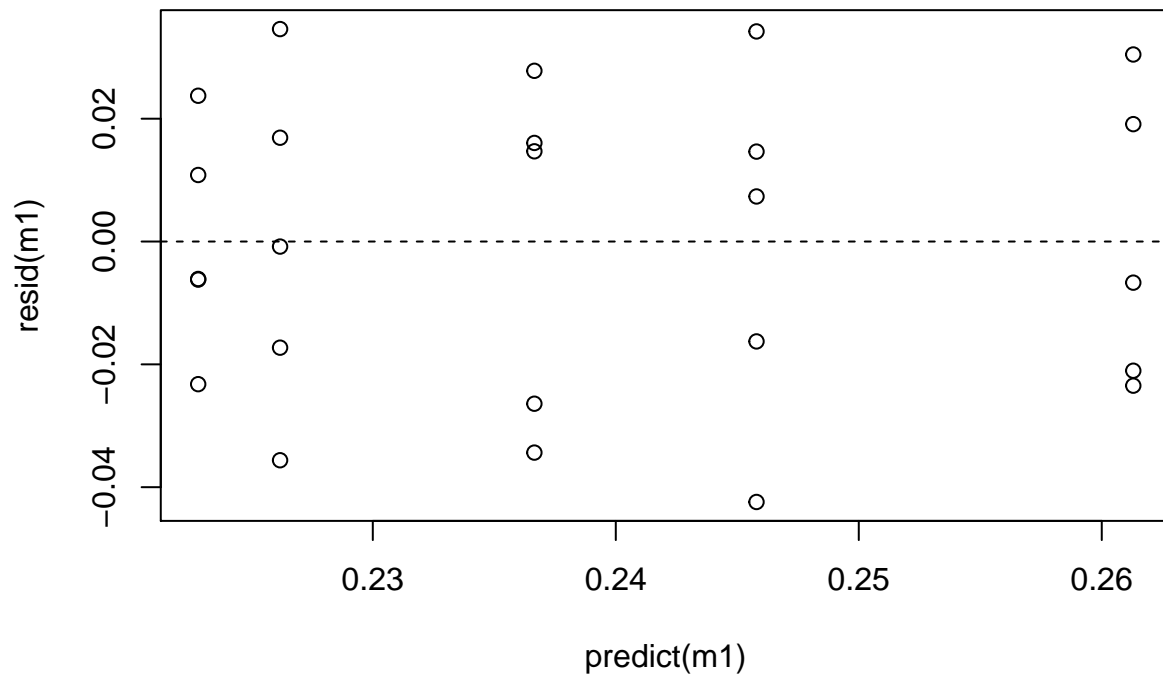
```
sum(residuals(m1,type="deviance")^2)
```

```
## [1] 0.01330279
```

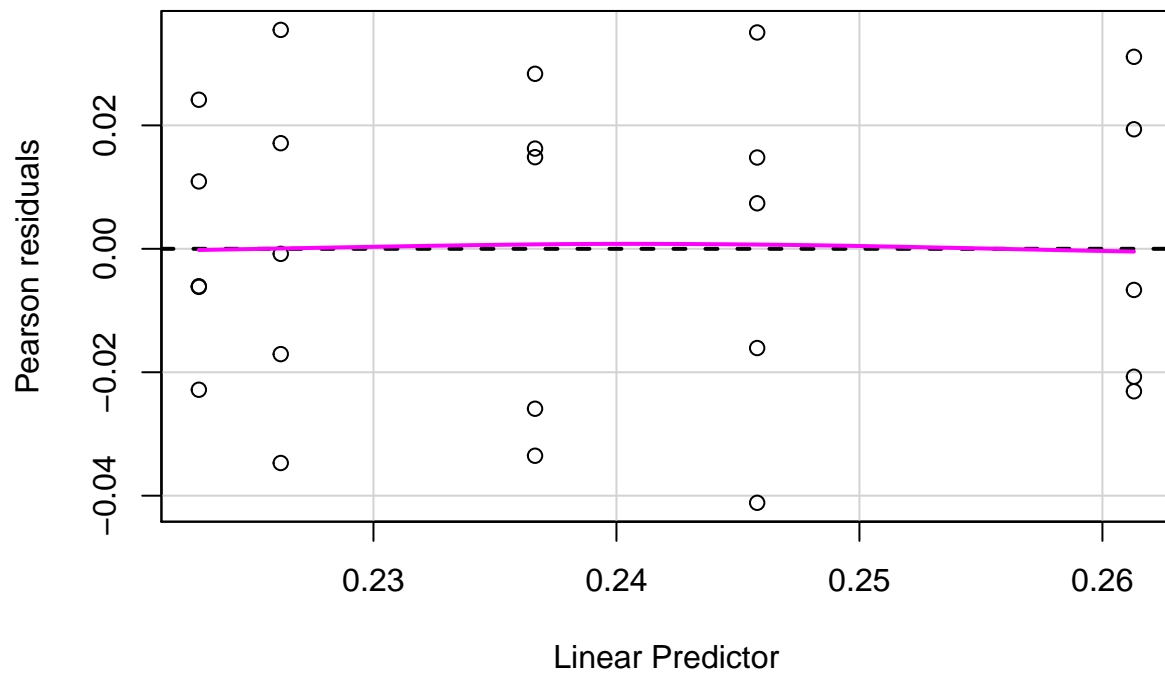
```
sum(residuals(m1,type="deviance")^2)/m1$df.residual
```

```
## [1] 0.0006651397
```

```
plot(predict(m1),resid(m1))
abline(h=0,lty=2)
```



```
residualPlot(m1)
```



```
Anova(m1,test="F")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: CI
## Error estimate based on Pearson residuals
```

```
##
##              Sum Sq Df F value  Pr(>F)
## SWEET      0.010204  4  3.8554 0.01763 *
## Residuals 0.013234 20
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(em1<-emmeans(m1,~SWEET))

## SWEET      emmean      SE df asymp.LCL asymp.UCL
## D00      0.2228136 0.007461659 Inf 0.2081890 0.2374381
## D08      0.2261840 0.007546103 Inf 0.2113939 0.2409741
## D15      0.2612971 0.008408703 Inf 0.2448164 0.2777779
## D20      0.2457914 0.008031574 Inf 0.2300498 0.2615330
## D30      0.2366490 0.007806474 Inf 0.2213486 0.2519494
##
## Results are given on the 1/mu^2 (not the response) scale.
## Confidence level used: 0.95

summary(em1,ty="response")

## SWEET response      SE df asymp.LCL asymp.UCL
## D00      2.118504 0.03547260 Inf  2.052224  2.191650
## D08      2.102660 0.03507518 Inf  2.037112  2.174972
## D15      1.956288 0.03147727 Inf  1.897366  2.021063
## D20      2.017050 0.03295495 Inf  1.955405  2.084918
## D30      2.055643 0.03390532 Inf  1.992248  2.125503
##
## Confidence level used: 0.95
## Intervals are back-transformed from the 1/mu^2 scale

#pairs(em1)
CLD(em1)

## SWEET      emmean      SE df asymp.LCL asymp.UCL .group
## D00      0.2228136 0.007461659 Inf 0.2081890 0.2374381  1
## D08      0.2261840 0.007546103 Inf 0.2113939 0.2409741  1
## D30      0.2366490 0.007806474 Inf 0.2213486 0.2519494 12
## D20      0.2457914 0.008031574 Inf 0.2300498 0.2615330 12
## D15      0.2612971 0.008408703 Inf 0.2448164 0.2777779  2
##
## Results are given on the 1/mu^2 (not the response) scale.
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 5 estimates
## significance level used: alpha = 0.05

CLD(em1,ty="response")

## SWEET response      SE df asymp.LCL asymp.UCL .group
## D15      1.956288 0.03147727 Inf  1.897366  2.021063  1
## D20      2.017050 0.03295495 Inf  1.955405  2.084918 12
## D30      2.055643 0.03390532 Inf  1.992248  2.125503 12
## D08      2.102660 0.03507518 Inf  2.037112  2.174972  2
## D00      2.118504 0.03547260 Inf  2.052224  2.191650  2
##
## Confidence level used: 0.95
## Intervals are back-transformed from the 1/mu^2 scale
```

```
## P value adjustment: tukey method for comparing a family of 5 estimates
## significance level used: alpha = 0.05
```

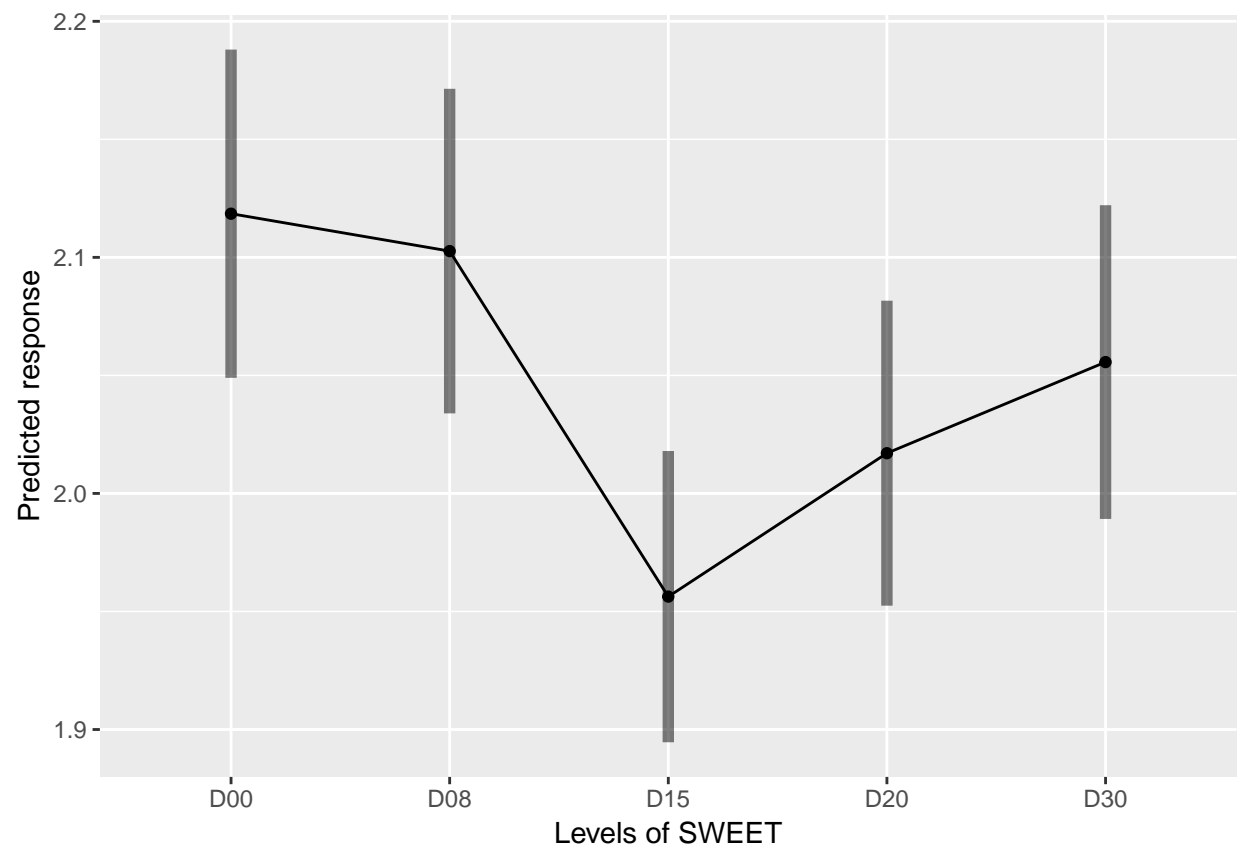
## Model 2: GLM Inverse Gaussian with identity link

```
m2<-glm(CI~SWEET,family=inverse.gaussian(link=identity),data=conversionindex)
```

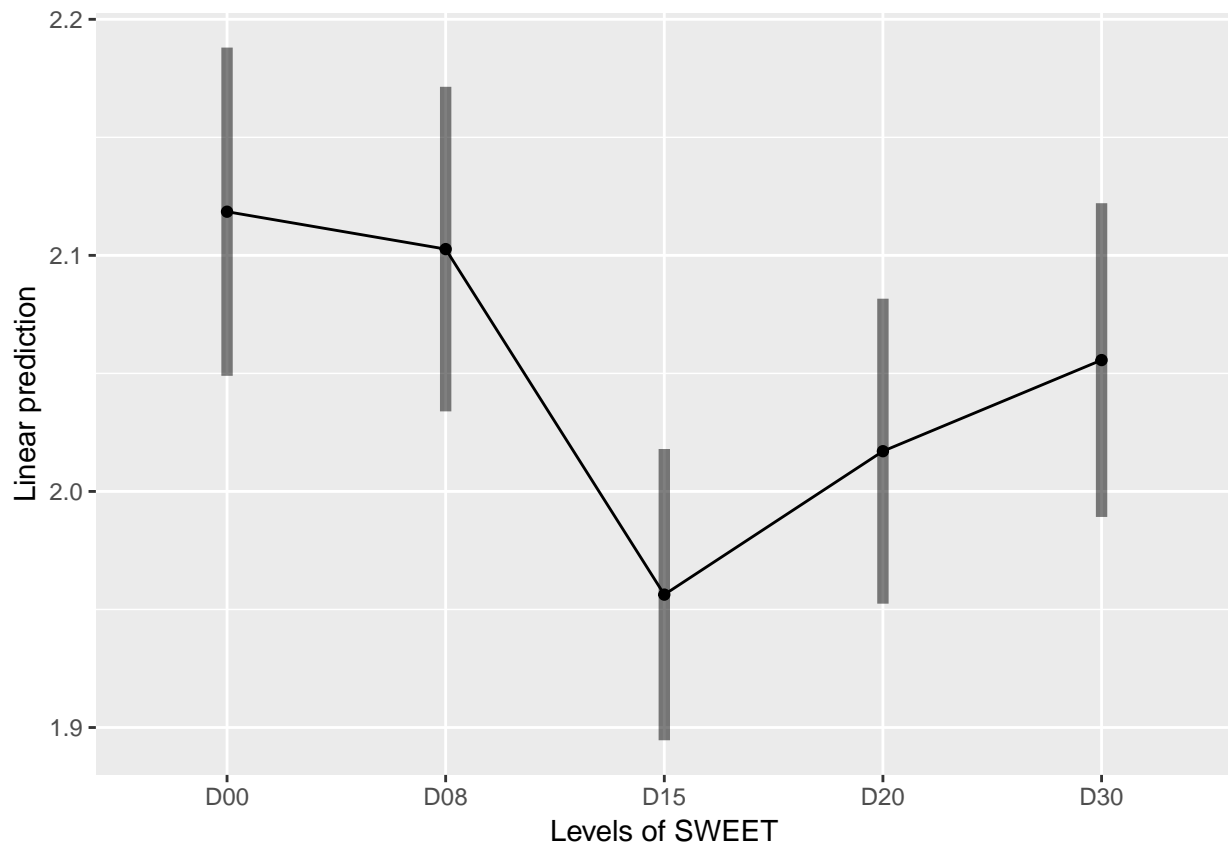
```
summary(m2)
```

```
##
## Call:
## glm(formula = CI ~ SWEET, family = inverse.gaussian(link = identity),
##      data = conversionindex)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.042407  -0.021036  -0.000818   0.016905   0.034593
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.11850    0.03547  59.723 < 2e-16 ***
## SWEETD08      -0.01584    0.04989  -0.318  0.75408
## SWEETD15      -0.16222    0.04742  -3.420  0.00271 **
## SWEETD20      -0.10145    0.04842  -2.095  0.04907 *
## SWEETD30      -0.06286    0.04907  -1.281  0.21483
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for inverse.gaussian family taken to be 0.0006617036)
##
## Null deviance: 0.023507  on 24  degrees of freedom
## Residual deviance: 0.013303  on 20  degrees of freedom
## AIC: -51.752
##
## Number of Fisher Scoring iterations: 3
emmip(m2,~SWEET,CIs=T,type="response")
```





```
emmip(m2, ~SWEET, CIs=T)
```



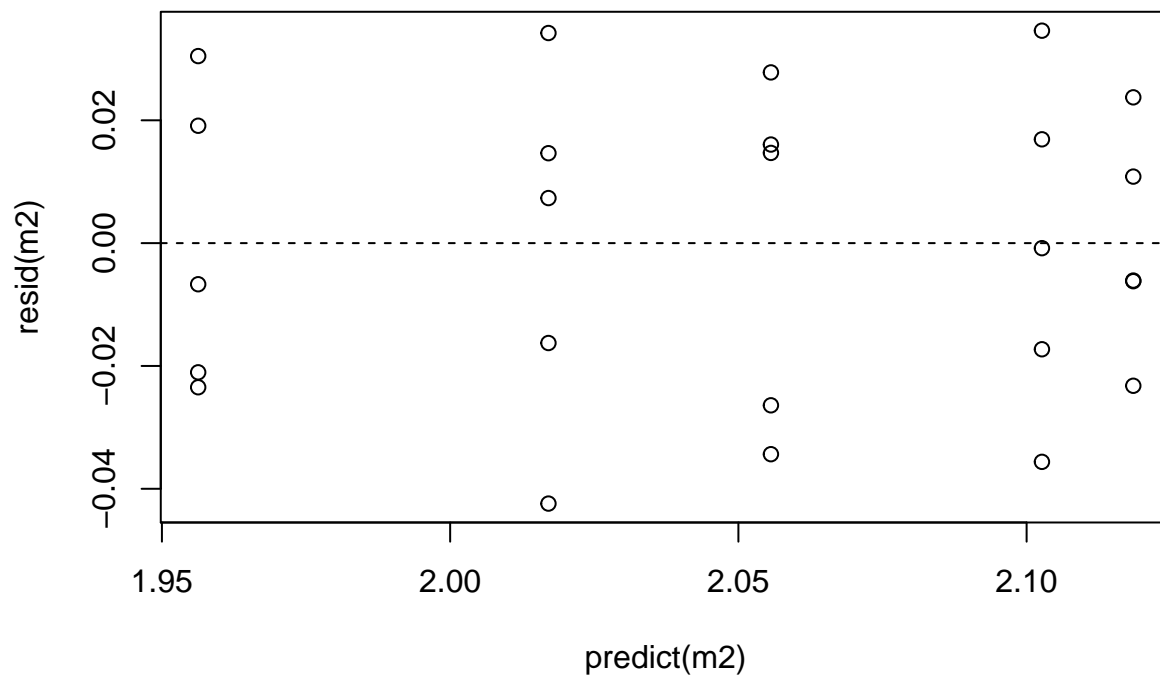
```
summary(m2)$family
```

```
##
## Family: inverse.gaussian
## Link function: identity
```

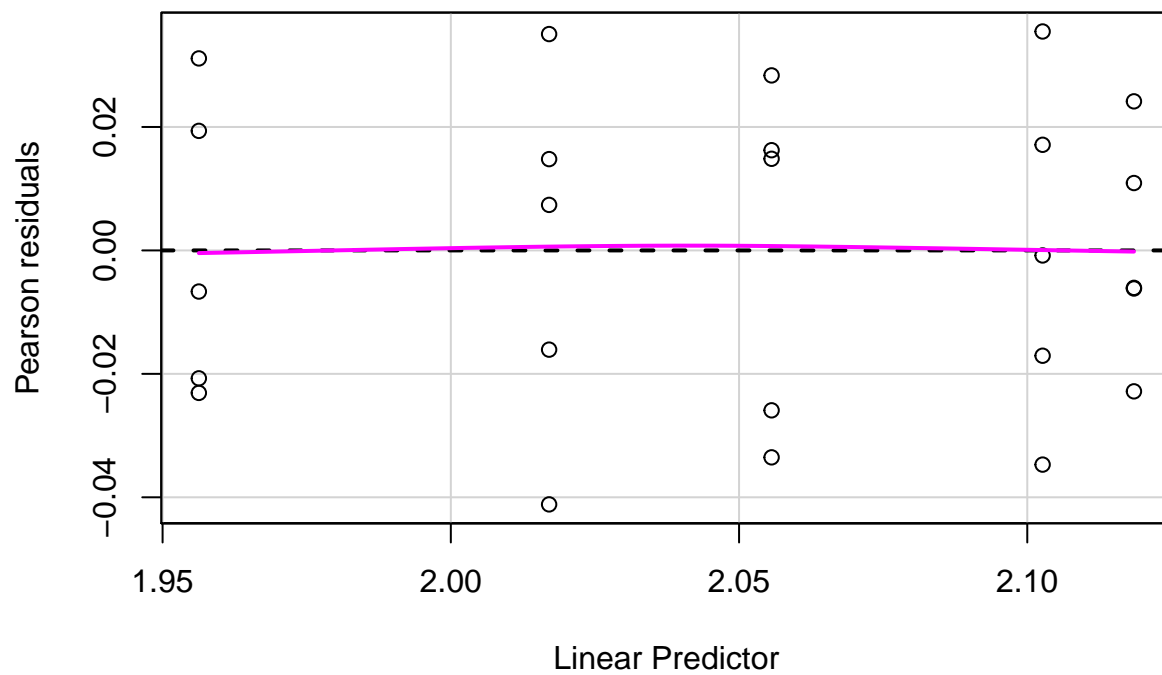
```
logLik(m2)
```

```
## 'log Lik.' 31.87621 (df=6)
```

```
scale<-sqrt(summary(m2)$disp)
plot(predict(m2),resid(m2))
abline(h=0,lty=2)
```



```
residualPlot(m2)
```



```
Anova(m2, test="F")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: CI
## Error estimate based on Pearson residuals
##
##           Sum Sq Df F value    Pr(>F)
## SWEET      0.010204  4   3.8554 0.01763 *
```

```
## Residuals 0.013234 20
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(em2<-emmeans(m2,~SWEET))

## SWEET    emmean      SE df asymp.LCL asymp.UCL
## D00    2.118504 0.03547240 Inf  2.048979  2.188028
## D08    2.102660 0.03507521 Inf  2.033914  2.171406
## D15    1.956288 0.03147718 Inf  1.894593  2.017982
## D20    2.017050 0.03295503 Inf  1.952459  2.081641
## D30    2.055643 0.03390535 Inf  1.989190  2.122096
##
## Confidence level used: 0.95

summary(em2,ty="response")

## SWEET    emmean      SE df asymp.LCL asymp.UCL
## D00    2.118504 0.03547240 Inf  2.048979  2.188028
## D08    2.102660 0.03507521 Inf  2.033914  2.171406
## D15    1.956288 0.03147718 Inf  1.894593  2.017982
## D20    2.017050 0.03295503 Inf  1.952459  2.081641
## D30    2.055643 0.03390535 Inf  1.989190  2.122096
##
## Confidence level used: 0.95

#pairs(em2)
CLD(em2)

## SWEET    emmean      SE df asymp.LCL asymp.UCL .group
## D15    1.956288 0.03147718 Inf  1.894593  2.017982    1
## D20    2.017050 0.03295503 Inf  1.952459  2.081641   12
## D30    2.055643 0.03390535 Inf  1.989190  2.122096   12
## D08    2.102660 0.03507521 Inf  2.033914  2.171406    2
## D00    2.118504 0.03547240 Inf  2.048979  2.188028    2
##
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 5 estimates
## significance level used: alpha = 0.05

CLD(em2,ty="response")

## SWEET    emmean      SE df asymp.LCL asymp.UCL .group
## D15    1.956288 0.03147718 Inf  1.894593  2.017982    1
## D20    2.017050 0.03295503 Inf  1.952459  2.081641   12
## D30    2.055643 0.03390535 Inf  1.989190  2.122096   12
## D08    2.102660 0.03507521 Inf  2.033914  2.171406    2
## D00    2.118504 0.03547240 Inf  2.048979  2.188028    2
##
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 5 estimates
## significance level used: alpha = 0.05
```

What we can see is that what really change are the parameter estimations, but not the predicted values, neither the goodness of fit of the model.