W271 Spring 18: Lab 2

Alyssa Eisenberg, Jeffrey Hsu, Gerard Kelly

Alcohol Consumption, Self-Esteem and Romantic Interactions

Introduction

The researchers stated the hypothesis as follow: *"We hypothesized that negative interactions with romantic partners would be associated with alcohol consumption (and an increased desire to drink). We predicted that people with low trait self-esteem would drink more on days they experienced more negative relationship interactions compared with days during which they experienced fewer negative relationship interactions. The relation between drinking and negative relationship interactions should not be evident for individuals with high trait self-esteem."*

EDA

```
library(car); require(dplyr); library(Hmisc); library(mcprofile); library(ggplot2); library(gridExtra);
dehart <- read.table(file="DeHartSimplified.csv", header=TRUE, sep=",")
describe(dehart) #with a 10-page limit, should we include this type of output?
```

```
## dehart
##
##  13  Variables      623  Observations
## ---------------------------------------------------------------------------
## id
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      623        0       89        1    75.89    56.82      7.0     16.2
##      .25      .50      .75      .90      .95
##     33.0     60.0    123.0    147.2    153.0
##
## lowest :   1   2   4   5   7, highest: 153 154 155 156 160
## ---------------------------------------------------------------------------
## studyday
##        n  missing distinct     Info     Mean      Gmd
##      623        0        7     0.98        4    2.289
##
## Value          1     2     3     4     5     6     7
## Frequency     89    89    89    89    89    89    89
## Proportion 0.143 0.143 0.143 0.143 0.143 0.143 0.143
## ---------------------------------------------------------------------------
## dayweek
##        n  missing distinct     Info     Mean      Gmd
##      623        0        7     0.98        4    2.289
##
## Value          1     2     3     4     5     6     7
## Frequency     89    89    89    89    89    89    89
## Proportion 0.143 0.143 0.143 0.143 0.143 0.143 0.143
## ---------------------------------------------------------------------------
## numall
```

```
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      622        1       18     0.97    2.524    2.636     0.00     0.00
##      .25      .50      .75      .90      .95
##     1.00     2.00     3.75     6.00     8.00
##
## Value          0     1     2     3     4     5     6     7     8     9
## Frequency    141   112   132    81    49    43    24     6     9     7
## Proportion 0.227 0.180 0.212 0.130 0.079 0.069 0.039 0.010 0.014 0.011
##
## Value         10    11    12    13    14    15    18    21
## Frequency      7     4     2     1     1     1     1     1
## Proportion 0.011 0.006 0.003 0.002 0.002 0.002 0.002 0.002
## ---------------------------------------------------------------------------
## nrel
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      623        0       33    0.551    0.359   0.6252        0        0
##      .25      .50      .75      .90      .95
##        0        0        0        1        2
##
## lowest : 0.0000000 0.2000000 0.2500000 0.3333333 0.4000000
## highest: 5.0000000 5.5000000 5.8333333 6.0000000 9.0000000
## ---------------------------------------------------------------------------
## prel
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      623        0       68    0.982    2.583    2.613   0.0000   0.0000
##      .25      .50      .75      .90      .95
##   0.4167   2.0000   4.0000   6.0000   7.8683
##
## lowest : 0.0000000 0.2000000 0.2500000 0.3333333 0.5000000
## highest: 8.1666667 8.3333333 8.5000000 8.6666667 9.0000000
## ---------------------------------------------------------------------------
## negevent
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      623        0      131    0.996   0.4414   0.4123   0.0000   0.0000
##      .25      .50      .75      .90      .95
##   0.1583   0.3500   0.6292   1.0000   1.1500
##
## lowest : 0.00000000 0.02500000 0.03333333 0.05000000 0.07500000
## highest: 1.70000000 1.93000000 1.95000000 2.01666667 2.37666667
## ---------------------------------------------------------------------------
## posevent
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      623        0      216        1    1.048   0.7077    0.200    0.300
##      .25      .50      .75      .90      .95
##    0.600    0.950    1.378    1.938    2.200
##
## lowest : 0.00000000 0.04000000 0.05000000 0.06666667 0.10000000
## highest: 3.23333333 3.25000000 3.30000000 3.40000000 3.88333333
## ---------------------------------------------------------------------------
## gender
##        n  missing distinct     Info     Mean      Gmd
##      623        0        2    0.739    1.562   0.4932
##
## Value          1     2
```

```
## Frequency    273   350
## Proportion 0.438 0.562
## ---------------------------------------------------------------------------
## rosn
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      623        0       17    0.993    3.436   0.4663      2.7      2.9
##      .25      .50      .75      .90      .95
##      3.2      3.5      3.8      3.9      4.0
##
## Value        2.1   2.4   2.5   2.7   2.8   2.9   3.0   3.1   3.2   3.3
## Frequency      7     7    14     7    21    35    42    21    28    42
## Proportion 0.011 0.011 0.022 0.011 0.034 0.056 0.067 0.034 0.045 0.067
##
## Value        3.4   3.5   3.6   3.7   3.8   3.9   4.0
## Frequency     35    84    63    49    63    49    56
## Proportion 0.056 0.135 0.101 0.079 0.101 0.079 0.090
## ---------------------------------------------------------------------------
## age
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      623        0       89        1    34.29     5.18    26.24    27.82
##      .25      .50      .75      .90      .95
##    30.53    34.57    38.19    40.15    40.56
##
## lowest : 24.43258 25.57700 26.05613 26.14100 26.23682
## highest: 40.56400 40.58864 40.68720 40.82957 42.27789
## ---------------------------------------------------------------------------
## desired
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      620        3       22    0.996    4.465    1.921    1.333    2.000
##      .25      .50      .75      .90      .95
##    3.333    4.667    5.667    6.667    7.333
##
## lowest : 1.000000 1.333333 1.666667 2.000000 2.333333
## highest: 6.666667 7.000000 7.333333 7.666667 8.000000
## ---------------------------------------------------------------------------
## state
##        n  missing distinct     Info     Mean      Gmd      .05      .10
##      620        3       25    0.993    3.966   0.4894    3.222    3.333
##      .25      .50      .75      .90      .95
##    3.667    4.000    4.222    4.556    4.556
##
## lowest : 2.333333 2.444444 2.555556 2.666667 2.777778
## highest: 4.555556 4.666667 4.777778 4.888889 5.000000
## ---------------------------------------------------------------------------
```

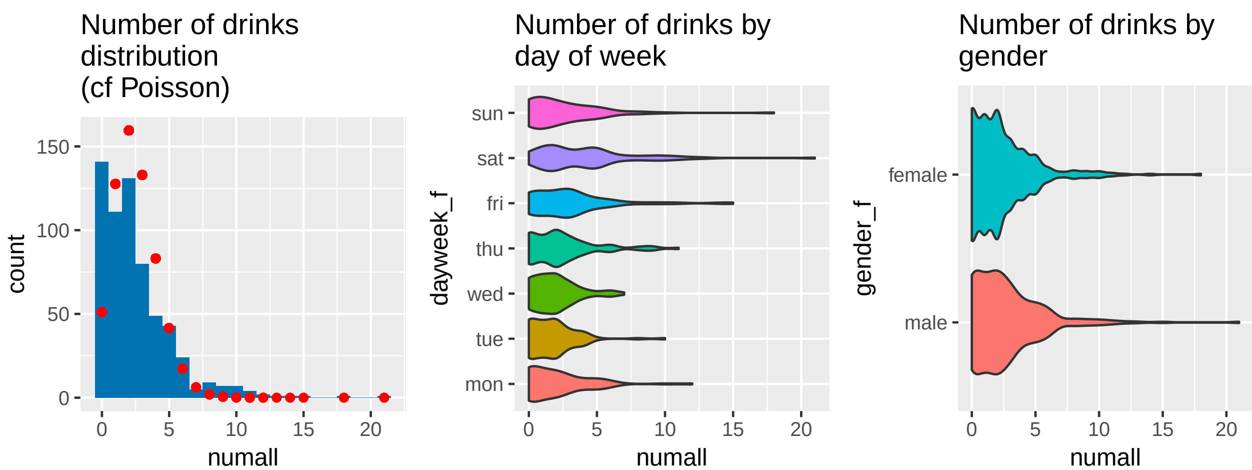
The dataset contains 623 observations of 13 variables representing entries in records kept by study participants. The variable **id** is a numeric identifier for each of the 89 study participants. Each participant recorded entries for seven consecutive days, indexed by the **studyday** variable, with the **dayweek** variable indicating which days of the week these correspond to (Monday = 1). The variable **gender** takes on one of two values according to whether the participant is male (1) or female (2); about 56% of the participants are female.

```
dehart$dayweek_f <- factor(dehart$dayweek); levels(dehart$dayweek_f) = c("mon","tue","wed","thu","fri",
dehart$gender_f <- factor(dehart$gender); levels(dehart$gender_f) = c("male","female")
```

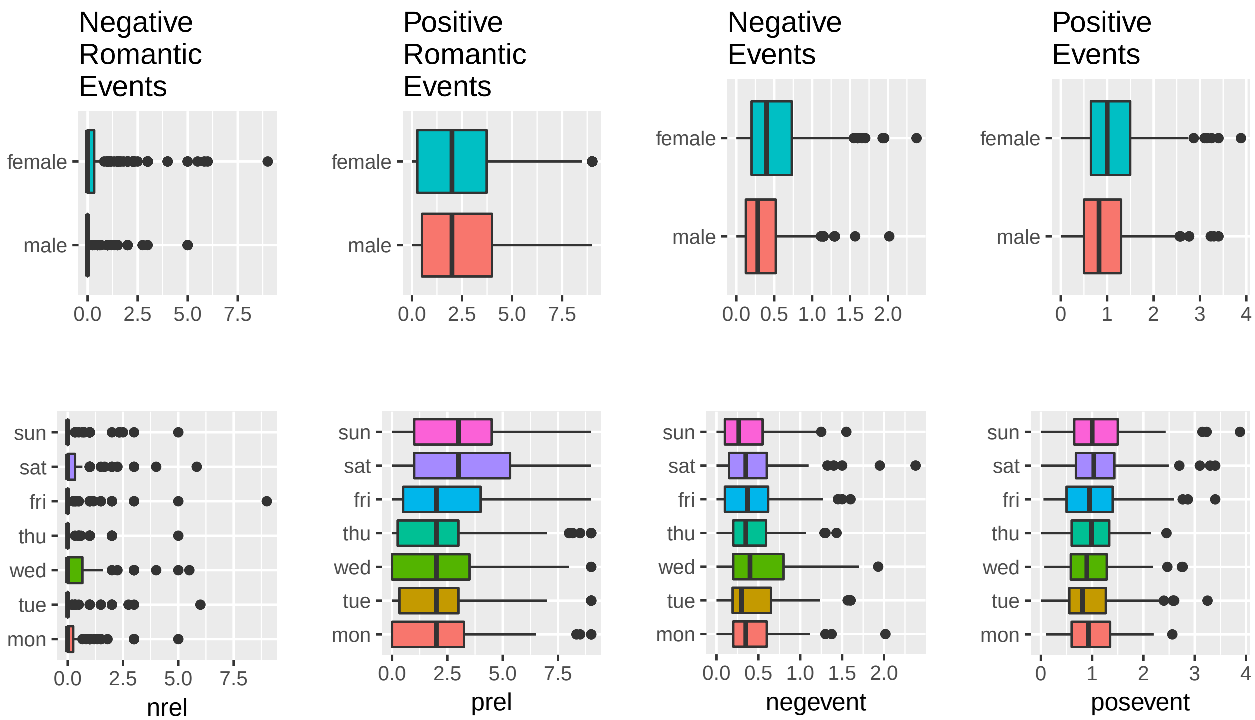
For each of the seven days, participants record the number of drinks consumed with the integer **numall** count variable. There is one missing value. Values range from 0 to 15 with single outliers at 19 and 21. Observations are concentrated in the range 0 to 5. The sample mean and sample variance are 2.52 and 2.66 respectively. The median number of drinks is 3 for Friday, 4 for Saturday and 2 for all other days. There is a pronounced positive skew to the number of drinks for Fridays, Saturdays and Sundays.

The variables **nrel** and **prel** are index measures for the negative and positive romantic-relationship events experienced by the subject each day (ranging from 0 to around 9), while **negevent** and **posevent** are index values combining the total number and intensity of negative and positive events experienced that day (ranging from 0 to around 4). There are no missing values for these variables. Distributions (and scales) for these measurements are very different, although tend to be similar for males and females, and similar across days of the week, although **prel** has signifiantly higher average (mean and median) values on Saturdays and Sundays. Distributions for all these variables are strongly positively-skewed, but this is particularly extreme for **nrel**, where a small fraction of outlying observations accounts for almost all of the variation.

```
#mean(dehart$numall, na.rm=TRUE); sd(dehart$numall, na.rm=TRUE)
#aggregate(numall ~ dayweek_f, data = dehart, FUN = function(x) c(m = mean(x), n = median(x)))
#The distribution of 622 times a Poisson random variable with parameter lambda is overlaid on the histo
lambda = 2.5
p1 <- ggplot(na.omit(dehart), aes(x = numall)) + geom_histogram(aes(y = ..count..), binwidth = 1, fill=
  geom_point(aes(y = 622*dpois(x = numall,lambda)), color = "red")+ ggtitle("Number of drinks\ndistribu
p2<-ggplot(na.omit(dehart), aes(dayweek_f, numall)) + geom_violin(aes(fill = dayweek_f)) + ggtitle("Numb
p3<-ggplot(na.omit(dehart), aes(gender_f, numall)) + geom_violin(aes(fill = gender_f)) + ggtitle("Number
grid.arrange(p1, p2, p3, ncol = 3)
```




```
p1a<-ggplot(dehart, aes(gender_f, nrel)) + geom_boxplot(aes(fill = gender_f)) + labs(x = "", y = "") + g
p1b<-ggplot(dehart, aes(dayweek_f, nrel)) + geom_boxplot(aes(fill = dayweek_f)) + labs(x = "") + ggtitle
p2a<-ggplot(dehart, aes(gender_f, prel)) + geom_boxplot(aes(fill = gender_f)) + labs(x = "", y = "") + g
p2b<-ggplot(dehart, aes(dayweek_f, prel)) + geom_boxplot(aes(fill = dayweek_f)) + labs(x = "") + ggtitle
p3a<-ggplot(dehart, aes(gender_f, negevent)) + geom_boxplot(aes(fill = gender_f)) + labs(x = "", y = "")
p3b<-ggplot(dehart, aes(dayweek_f, negevent)) + geom_boxplot(aes(fill = dayweek_f)) + labs(x = "") + gg
p4a<-ggplot(dehart, aes(gender_f, posevent)) + geom_boxplot(aes(fill = gender_f)) + labs(x = "", y = "")
p4b<-ggplot(dehart, aes(dayweek_f, posevent)) + geom_boxplot(aes(fill = dayweek_f)) + labs(x = "") + gg
grid.arrange(p1a, p2a, p3a, p4a, p1b, p2b, p3b, p4b, ncol = 4)
```




```
aggregate(nrel ~ dayweek_f, data = dehart, FUN = mean)
```

```
##   dayweek_f      nrel
## 1       mon 0.3471910
## 2       tue 0.2882022
## 3       wed 0.5541466
## 4       thu 0.2295880
## 5       fri 0.4014981
## 6       sat 0.4033708
## 7       sun 0.2893258
```

The **rosn** variable measures trait (long-term) self-esteem, a single measurement for each participant taken at the beginning of the study that does not change over the course of the seven days. This measurement ranges between 2 and 4, with a mean value around 3.4. Distributions for males and females differ, with a larger proportion of males recording lower values and a larger proportion of females recording higher values. The **age** variable measures age in years, ranging between 24.4 and 42.3 with a mean value of 34.3 and similar distributions for males and females but with a higher proportion of females recording higher values. Neither **rosn** nor **age** have any missing observations.

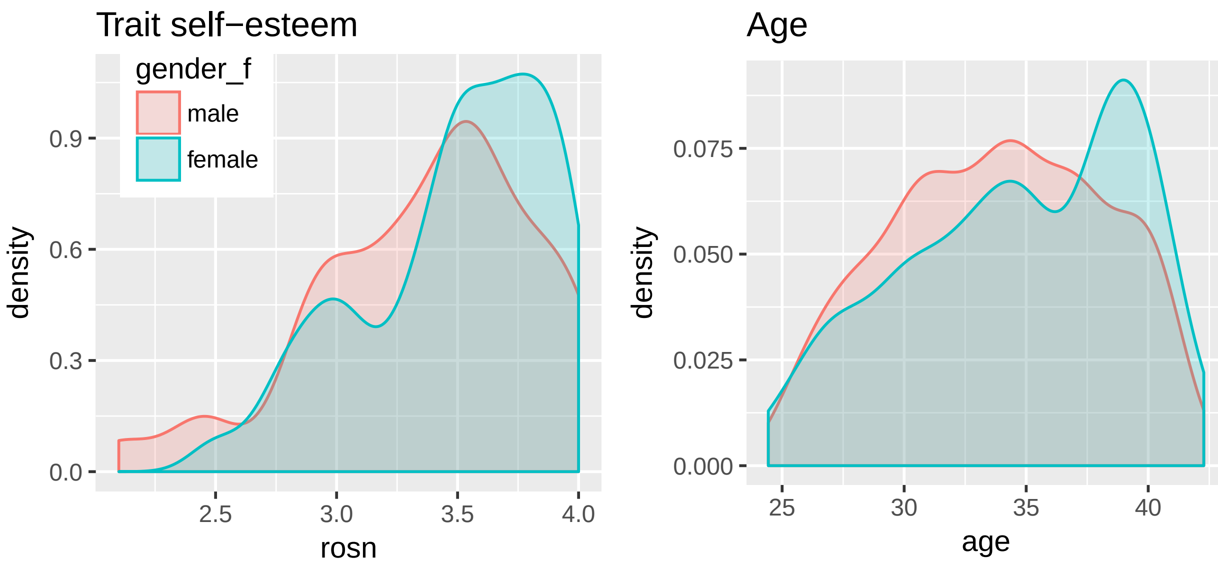
```
quantile(dehart$rosn[dehart$gender == 1])
```

```
##   0%  25%  50%  75% 100%
##  2.1  3.0  3.5  3.7  4.0
```

```
quantile(dehart$rosn[dehart$gender == 2])
```

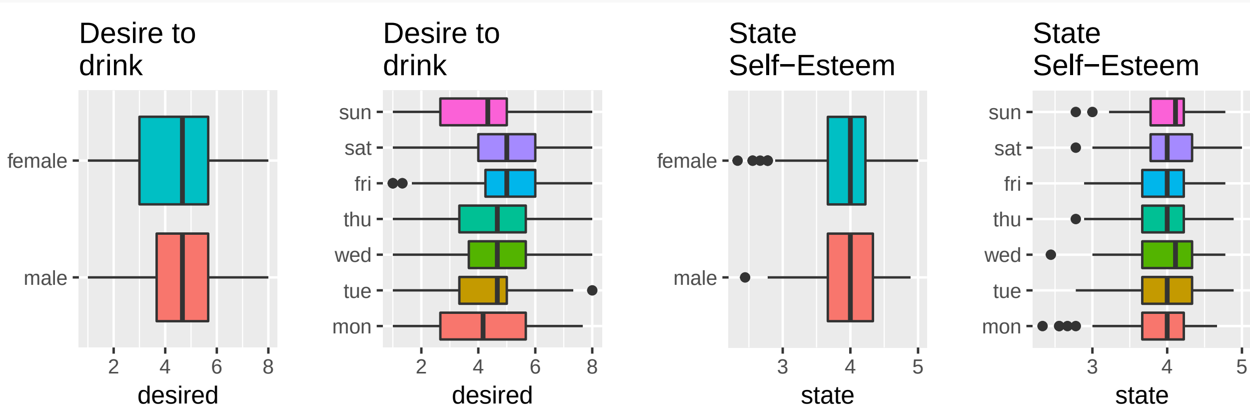
```
##   0%  25%  50%  75% 100%
## 2.50 3.30 3.55 3.80 4.00
```

```
p1<-ggplot(dehart, aes(x = rosn, fill = gender_f, colour = gender_f)) + geom_density(alpha=0.2)+ ggtitl
p2<-ggplot(dehart, aes(x = age, fill = gender_f, colour = gender_f)) + geom_density(alpha=0.2)+
  ggtitle("Age")+theme(legend.position="none")
grid.arrange(p1, p2, ncol = 2)
```



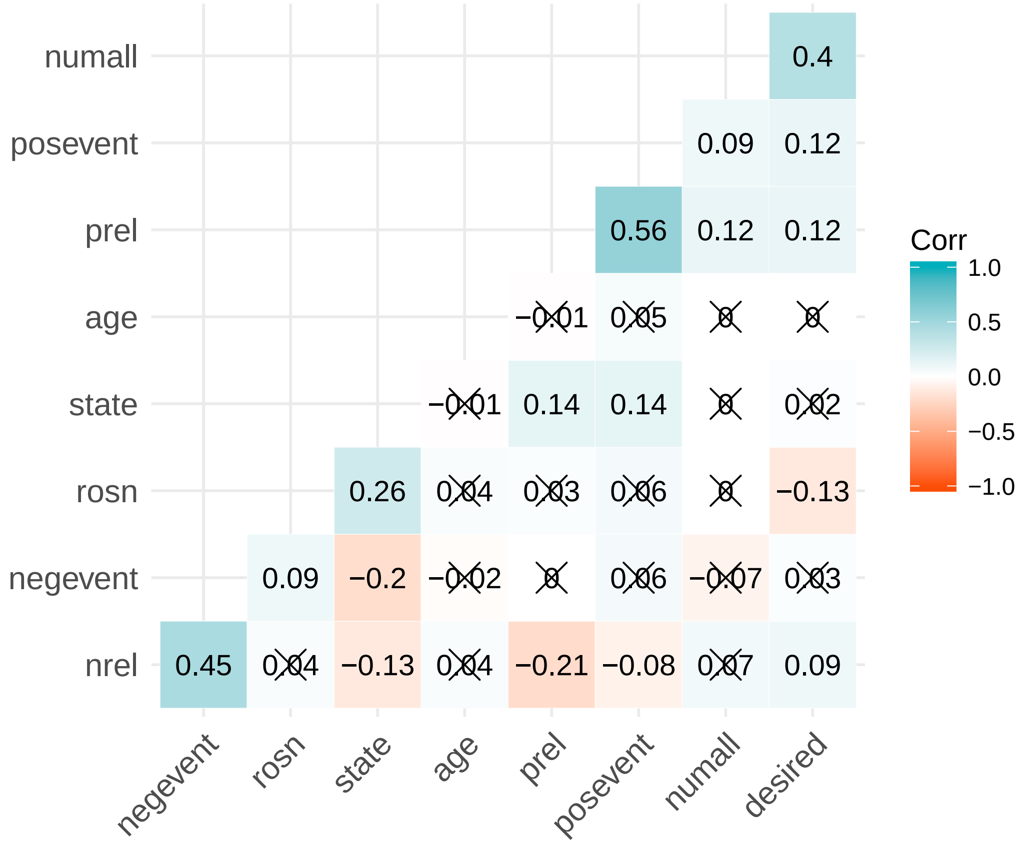
The **desired** variable is a measure of the participant's recorded desire to drink, with values ranging between 1 and 8, a mean of 4.5 and a fairly symmetric distribution. with a significant share of responses at minimum and maximum values. Average values are slightly higher on average for males than for females and are highest on Friday and Saturday and lowest on Sunday and Monday. The **state** variable is a record of the participant's state (short-term) self-esteem as it varies each day. This ranges between 2 and 5 with a mean of 3.97 and a moderately negative-skewed distribution, little difference between males and females and little difference between days of the week (although with more negative outliers on Mondays). The variables **desired** and **state** both contain three missing values, two of these being from the same observation.

```
# aggregate(desired ~ gender_f, data = dehart, FUN = mean)
# aggregate(desired ~ dayweek_f, data = dehart, FUN = mean)
# aggregate(state ~ gender_f, data = dehart, FUN = mean)
# aggregate(state ~ dayweek_f, data = dehart, FUN = mean)
p1<-ggplot(na.omit(dehart), aes(gender_f, desired)) + geom_boxplot(aes(fill = gender_f)) + labs(x = "")
  ggtitle("Desire to\ndrink") + theme(legend.position="none") + coord_flip()
p2<-ggplot(na.omit(dehart), aes(dayweek_f, desired)) + geom_boxplot(aes(fill = dayweek_f)) + labs(x = "
  ggtitle("Desire to\ndrink") + theme(legend.position="none") + coord_flip()
p3<-ggplot(na.omit(dehart), aes(gender_f, state)) + geom_boxplot(aes(fill = gender_f)) + labs(x = "") +
  ggtitle("State\nSelf-Esteem") + theme(legend.position="none") + coord_flip()
p4<-ggplot(na.omit(dehart), aes(dayweek_f, state)) + geom_boxplot(aes(fill = dayweek_f)) + labs(x = "")
  ggtitle("State\nSelf-Esteem") + theme(legend.position="none") + coord_flip()
grid.arrange(p1, p2, p3, p4, ncol = 4)
```



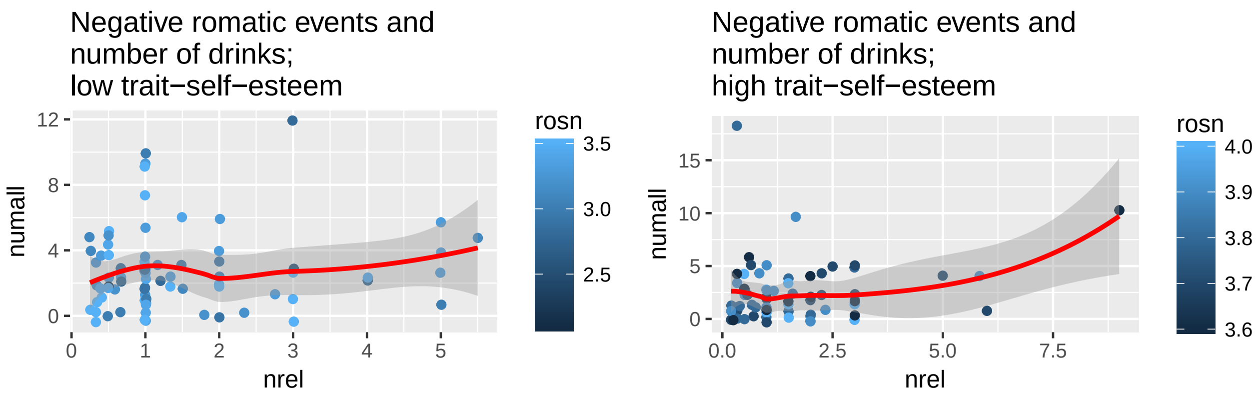
Bivarate relationships between these continuous variables are summarized in the below plot of correlation coefficients below. Most of these relationship are weak or insignificant. The strongest positive correlations are between **nrel** and **negevent** (negative events and negative romantic events), and between **prel** and **posevent** (positive events and positive romantic events). There is also a moderate positive correlation between the number of drinks and the desire to drink (**numall** and **desired**), and a mild positive correlation between trait and state self-esteem (**rosn** and **state**). State self-esteem has a weak positive correlation to positive event variables and a weak negative correlation to negative event variables, but no correlation with the number of drinks nor the desire to drink. Trait self-esteem has a weak negative correlation with the desire to drink but no correlation with the number of drinks. Age is uncorreated with any other variable.

```
data <- na.omit(dehart[,c(4,5,6,7,8,10,11,12,13)])
corr <- round(cor(data), 2)
ggcorrplot(corr, p.mat = cor_pmat(data),hc.order = TRUE, type = "lower", color = c("#FC4E07", "white",
```



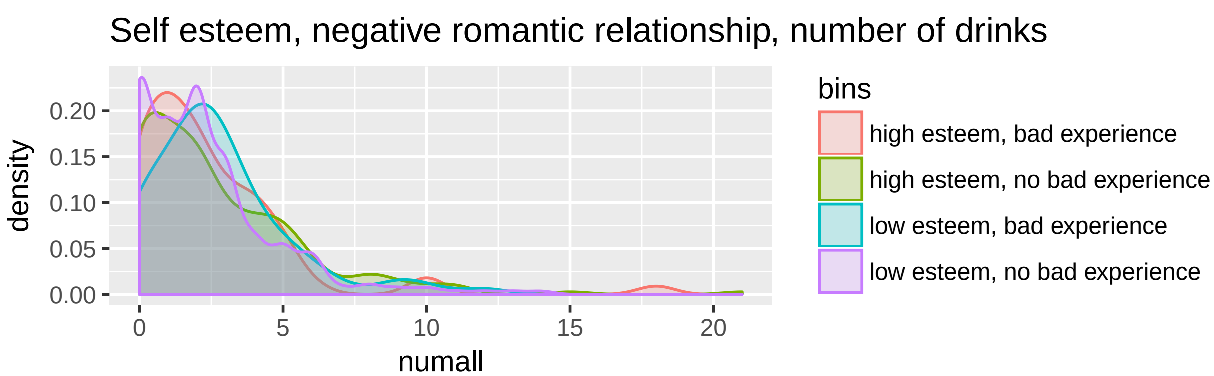
The heavily skewed distribution for **nrel** indicates that the negative romatic relationship events relevant to the hypothesis are relatively infrequent. A subset of the dataset can be created for participant-days involving a non-zero rating on this variable. The relationship between **nrel** and numall can then be compared for participants with below-median trait self-esteem and above-median trait self-esteem. The relationship between negative romatic relationship events and the number of drinks appears to be stronger for individuals with higher trait-self esteem, however the range of this is driven by a relatively small number of individuals recording high **nrel** values that do not appear in the lower trait-self-esteem subset.

```
dehart_nrel = dehart[which(dehart$nrel != 0),]
p1 <- ggplot(na.omit(dehart_nrel[which(dehart_nrel$rosn<=3.5),]), aes(nrel, numall)) + geom_jitter(aes(c
p2 <- ggplot(na.omit(dehart_nrel[which(dehart_nrel$rosn>3.5),]), aes(nrel, numall)) + geom_jitter(aes(co
grid.arrange(p1, p2,ncol = 2)
```

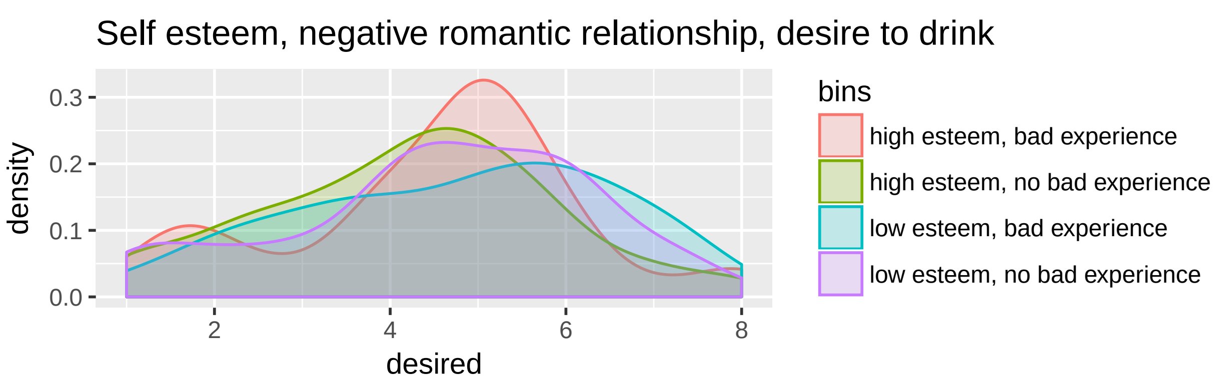



```
dehart$bins <- dehart$rosn
dehart$bins[dehart$rosn <= median(dehart$rosn) & dehart$nrel <= median(dehart$nrel)] = "low esteem, no b
dehart$bins[dehart$rosn <= median(dehart$rosn) & dehart$nrel > median(dehart$nrel)] = "low esteem, bad e
dehart$bins[dehart$rosn > median(dehart$rosn) & dehart$nrel <= median(dehart$nrel)] = "high esteem, no b
dehart$bins[dehart$rosn > median(dehart$rosn) & dehart$nrel > median(dehart$nrel)] = "high esteem, bad e
```

```
p1<- ggplot(na.omit(dehart), aes(numall, fill = bins, colour = bins)) +
  geom_density(alpha=0.2) +
  ggtitle("Self esteem, negative romantic relationship, number of drinks") + ggtitle("Self esteem, negat

p2<- ggplot(na.omit(dehart), aes(desired, fill = bins, colour = bins)) +
  geom_density(alpha=0.2) +
  ggtitle("Self esteem, negative romantic relationship, number of drinks") + ggtitle("Self esteem, negat

p1
```




```
p2
```



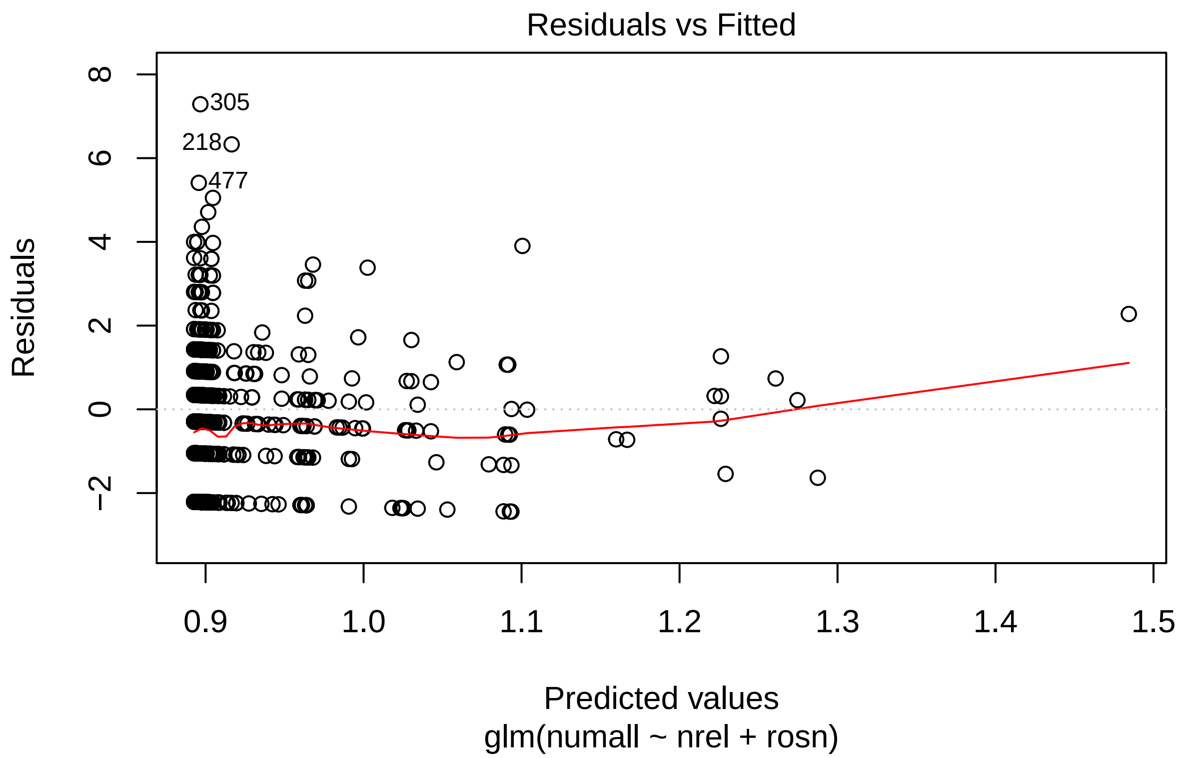
```r
# NUMBER OF DRINKS; POISSON MODEL
# significant interaction effect

# BASIC MODEL

N.mod.0 <- glm(formula = numall ~ nrel + rosn, family = poisson(link = "log"), data = na.omit(dehart))
summary(N.mod.0)
```

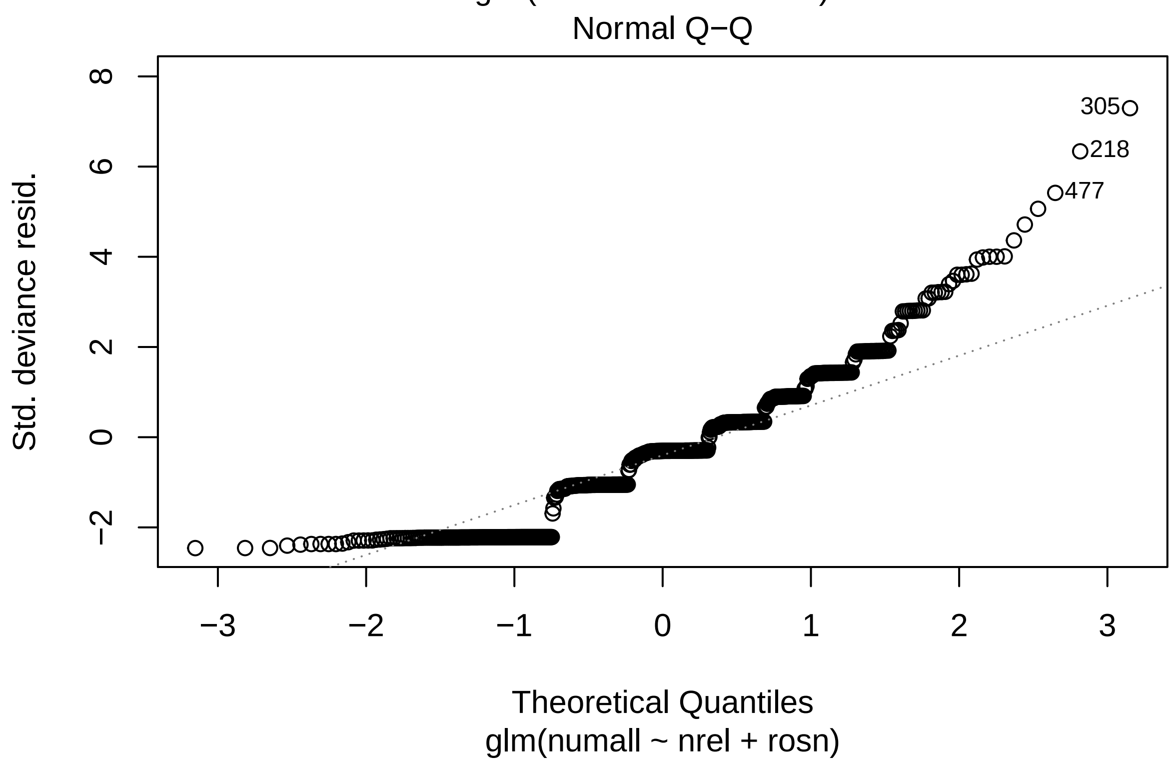
```
##
## Call:
## glm(formula = numall ~ nrel + rosn, family = poisson(link = "log"),
##     data = na.omit(dehart))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.4434  -1.1434  -0.3040   0.3448   7.2878
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.932497   0.209136   4.459 8.24e-06 ***
## nrel         0.065298   0.023671   2.759  0.00581 **
## rosn        -0.009943   0.060520  -0.164  0.86951
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1583.5  on 617  degrees of freedom
## Residual deviance: 1576.5  on 615  degrees of freedom
## AIC: 2944.5
##
## Number of Fisher Scoring iterations: 5
```

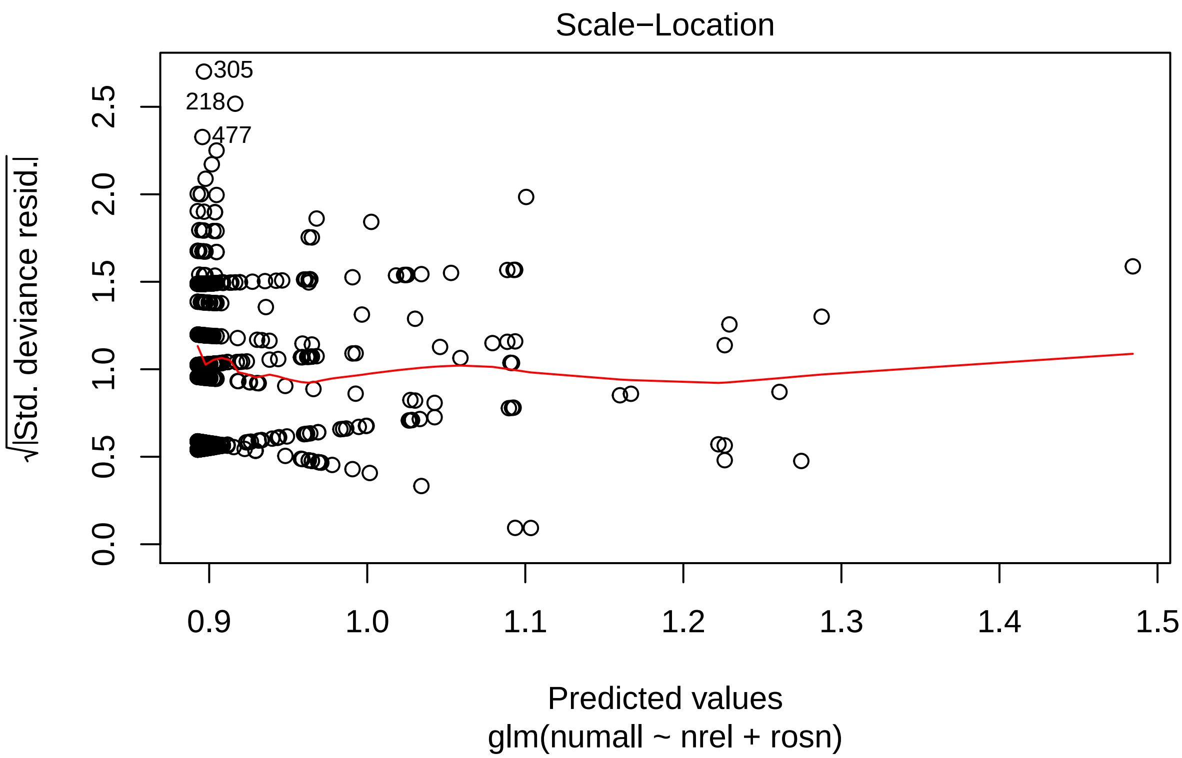
```r
# hypothesis tests for statistical significance
Anova(N.mod.0, test = "LR")
```

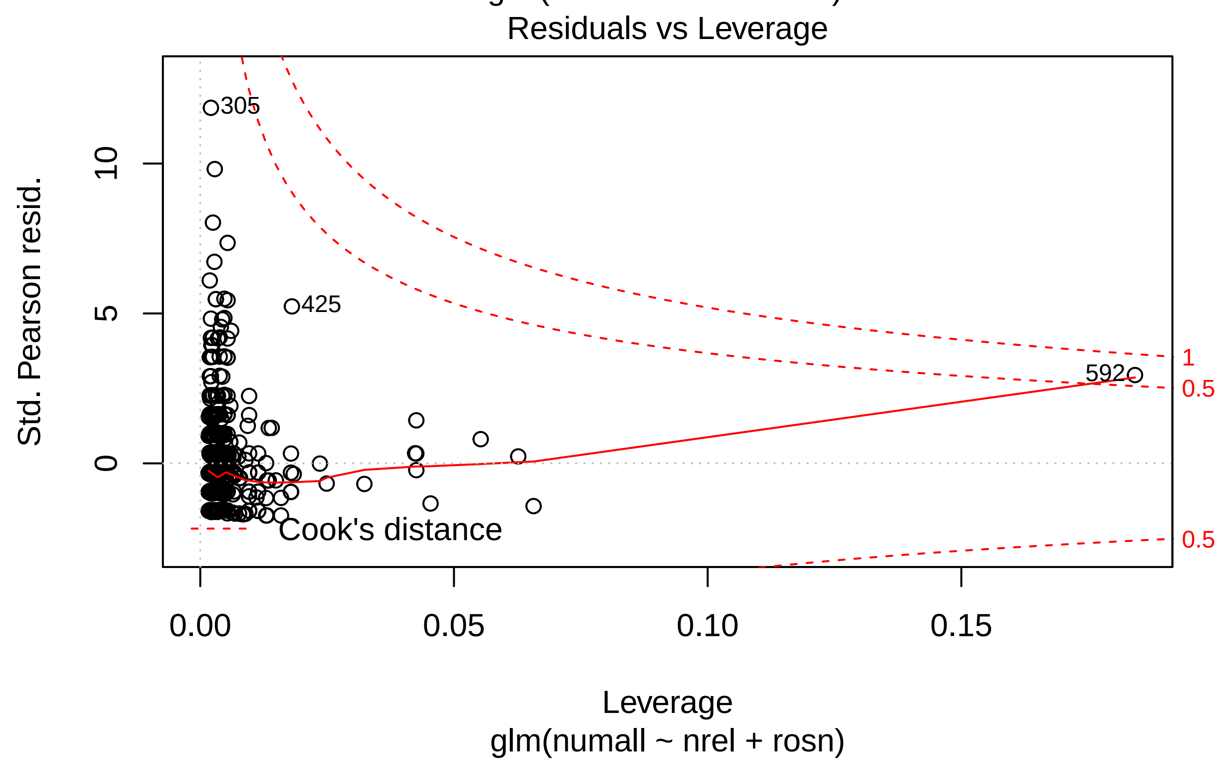
```
## Analysis of Deviance Table (Type II tests)
##
## Response: numall
##      LR Chisq Df Pr(>Chisq)
## nrel   6.9884  1   0.008204 **
## rosn   0.0270  1   0.869573
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# diagnostic plots
plot(N.mod.0)
```

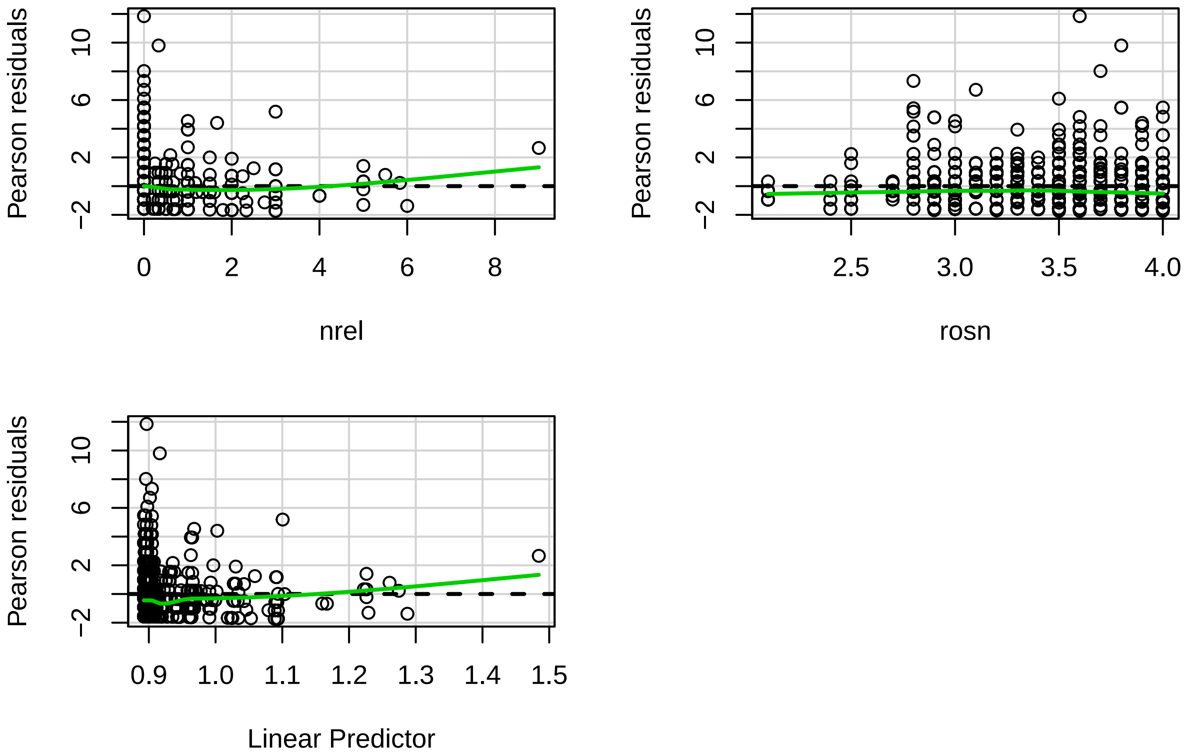

Residuals vs Fitted
Residuals
Predicted values
glm(numall ~ nrel + rosn)
305
218
477
8
6
4
2
0
−2
0.9
1.0
1.1
1.2
1.3
1.4
1.5


Normal Q−Q
Std. deviance resid.
Theoretical Quantiles
glm(numall ~ nrel + rosn)
305
218
477
8
6
4
2
0
−2
−3
−2
−1
0
1
2
3



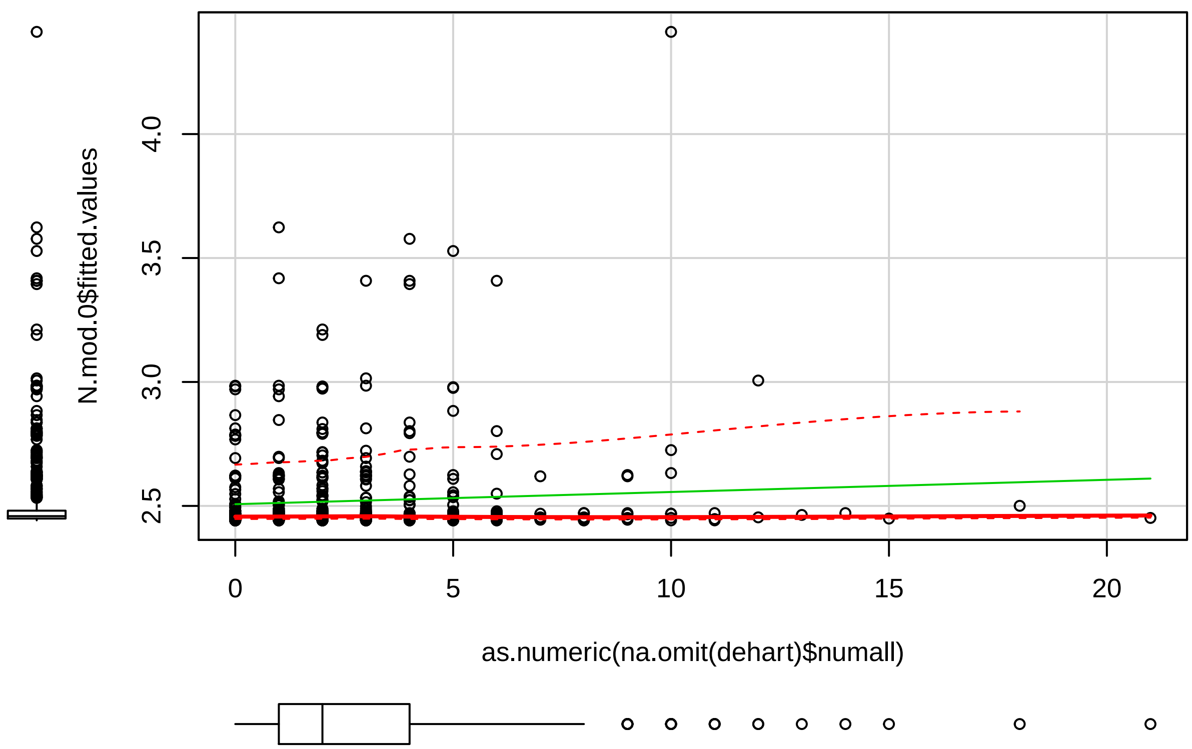



```
# Pearson residuals
residualPlots(N.mod.0)
```

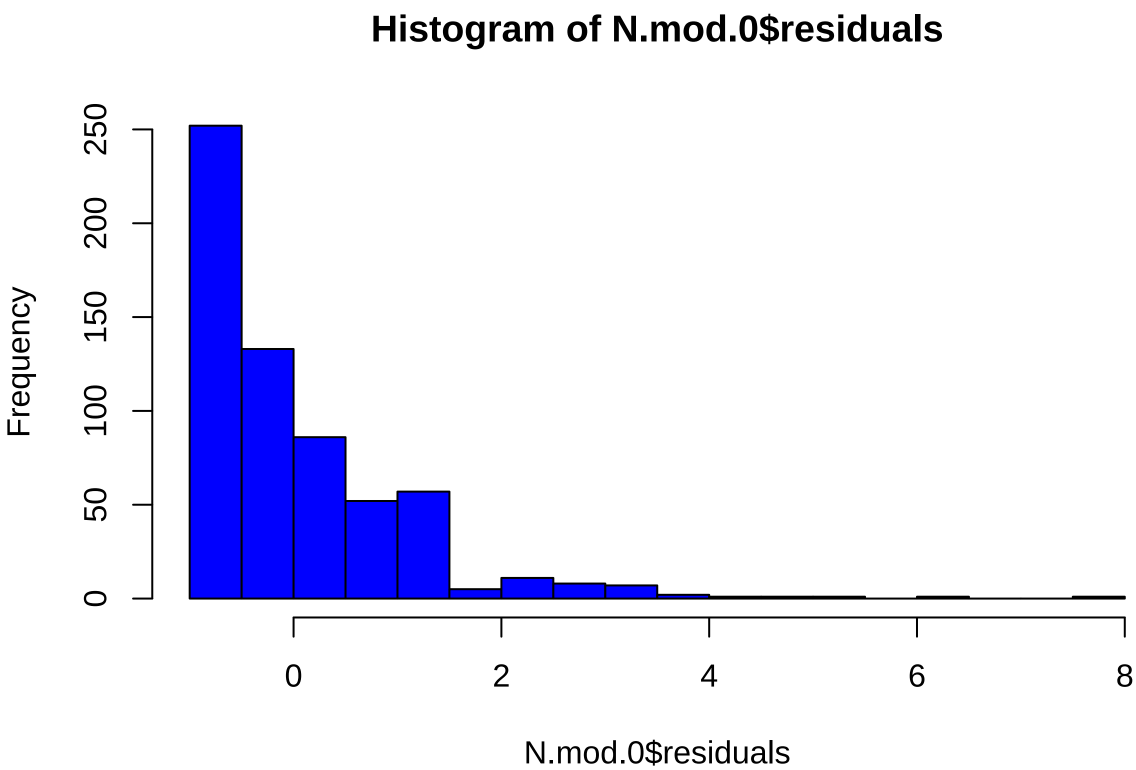



```
##       Test stat Pr(>|t|)
## nrel      7.028    0.008
## rosn      2.364    0.124
```

```
# scatterplot of actual and fitted values
scatterplot(as.numeric(na.omit(dehart)$numall), N.mod.0$fitted.values)
```



```r
# histogram of residuals
hist(N.mod.0$residuals, breaks = 20, col = "blue")
```




```r
N.mod.1 <- glm(formula = numall ~ nrel + rosn + nrel*rosn, family = poisson(link = "log"), data = na.omi
summary(N.mod.1)
```

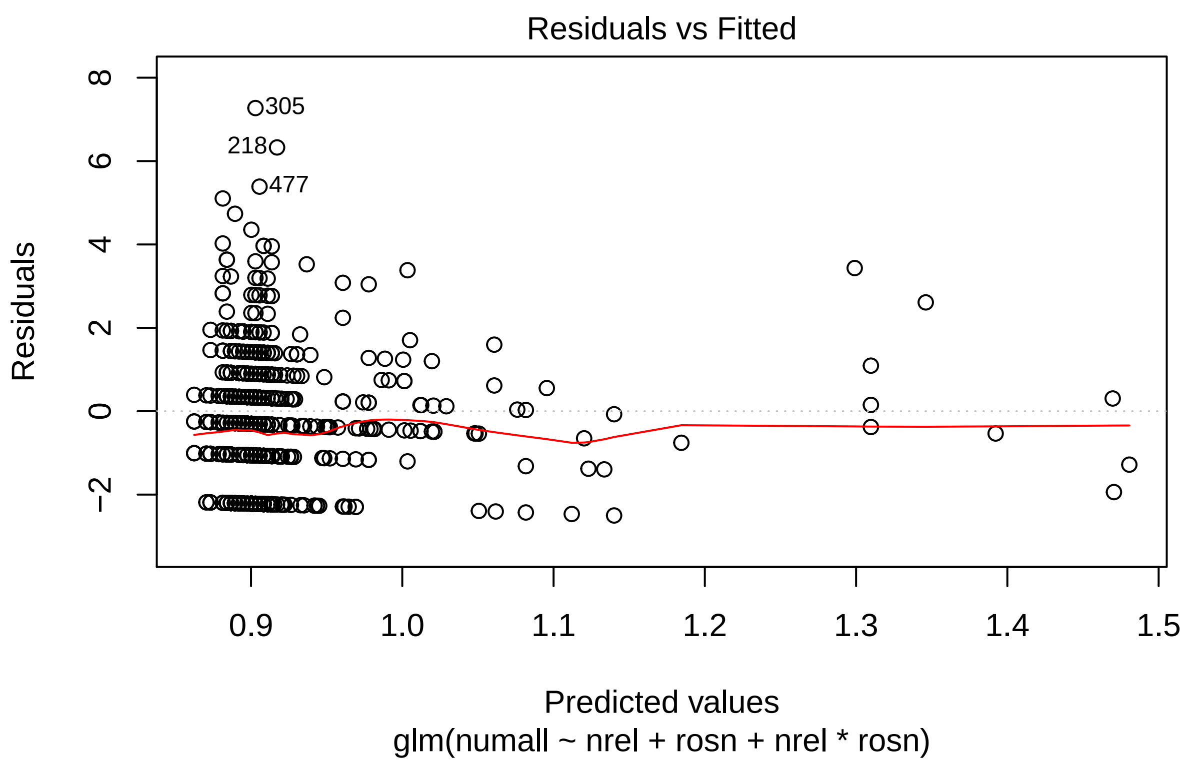
```
##
## Call:
## glm(formula = numall ~ nrel + rosn + nrel * rosn, family = poisson(link = "log"),
##     data = na.omit(dehart))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5007  -1.0966  -0.3156   0.3795   7.2719
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.80573    0.22392   3.598  0.00032 ***
## nrel         0.45430    0.23218   1.957  0.05039 .
## rosn         0.02700    0.06477   0.417  0.67682
## nrel:rosn   -0.11252    0.06722  -1.674  0.09416 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1583.5  on 617  degrees of freedom
## Residual deviance: 1573.8  on 614  degrees of freedom
## AIC: 2943.7
##
```

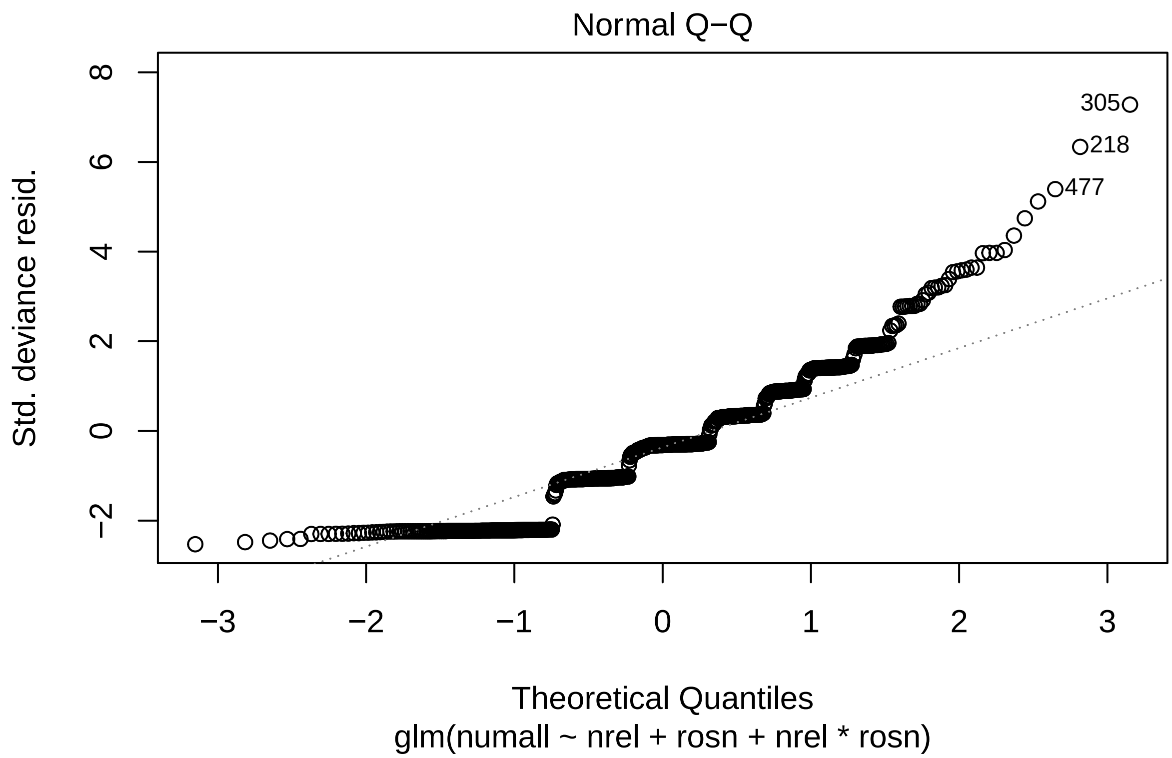
```
## Number of Fisher Scoring iterations: 5
```

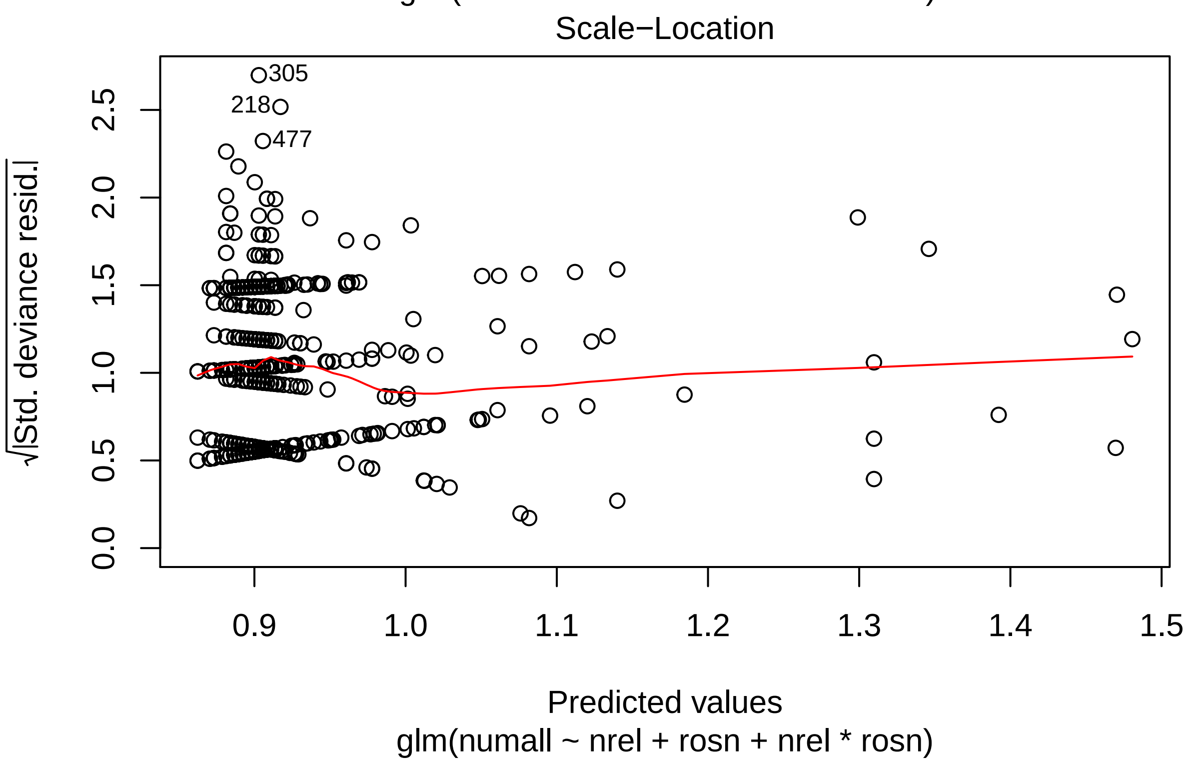
```r
# hypothesis tests for statistical significance
Anova(N.mod.1, test = "LR")
```

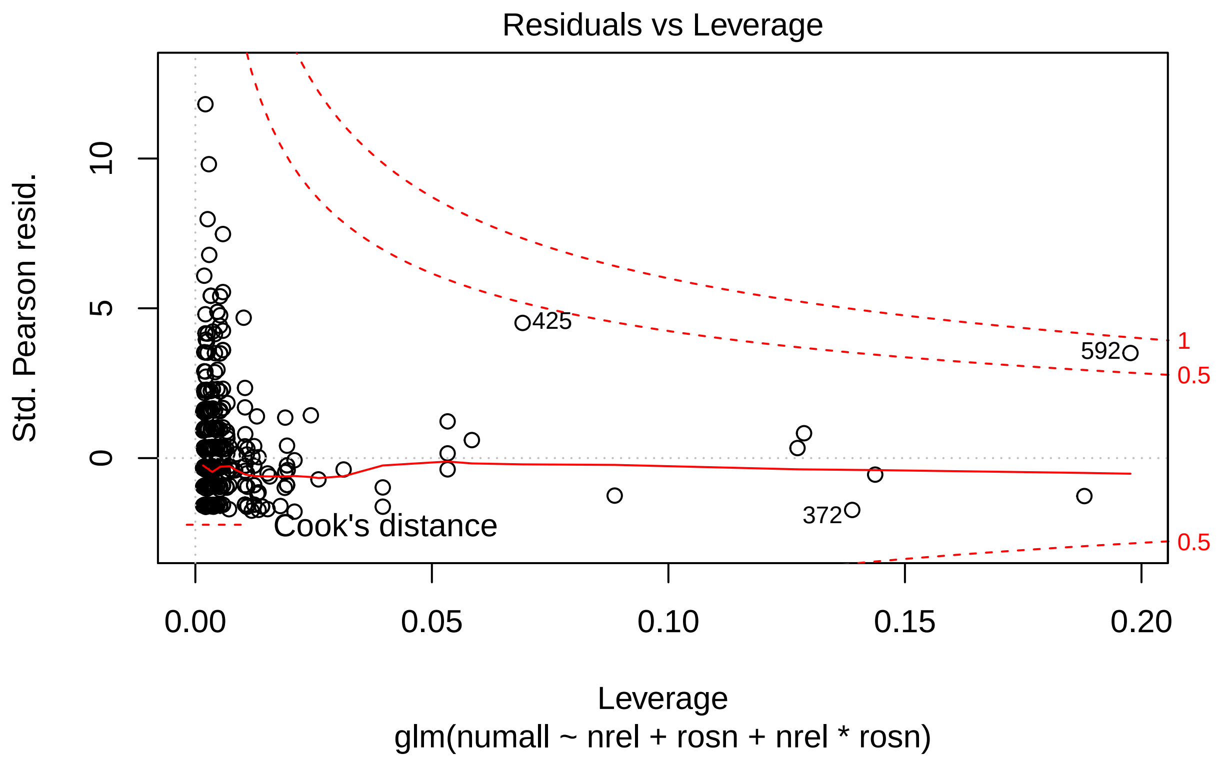
```
## Analysis of Deviance Table (Type II tests)
##
## Response: numall
##             LR Chisq Df Pr(>Chisq)
## nrel          6.9884  1   0.008204 **
## rosn          0.0270  1   0.869573
## nrel:rosn     2.7409  1   0.097809 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# diagnostic plots
plot(N.mod.1)
```

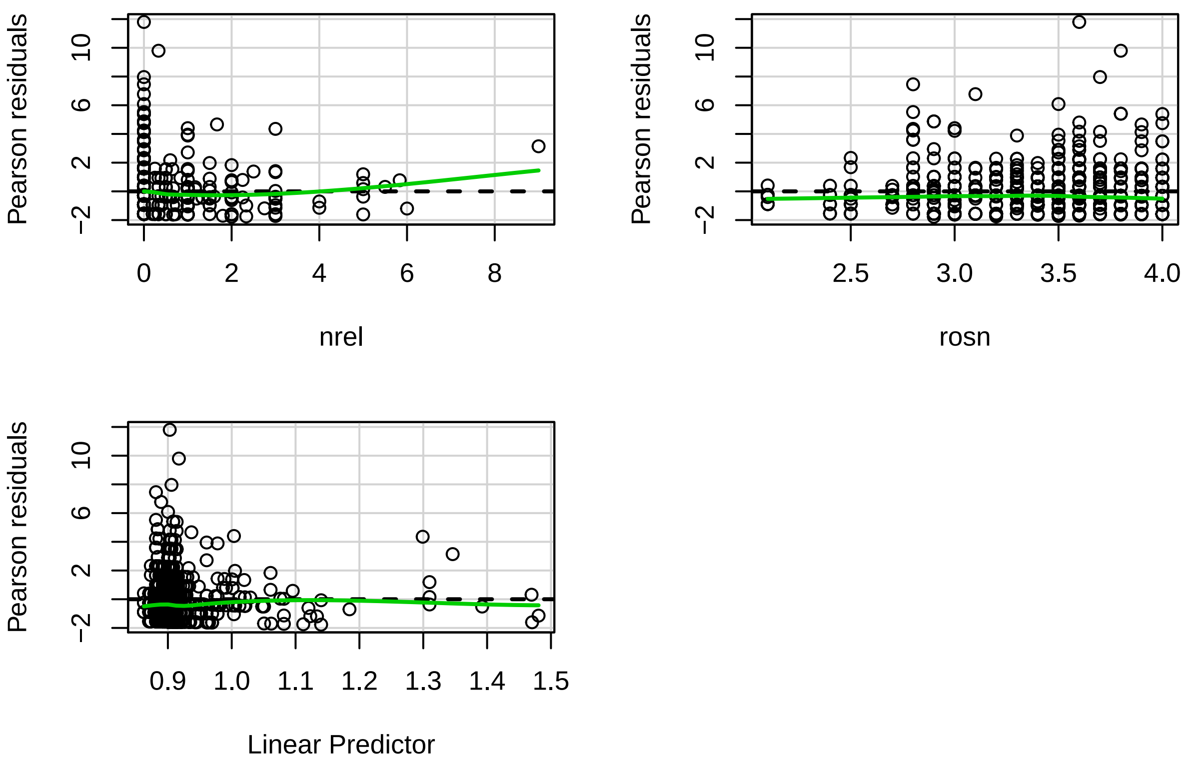



Normal Q−Q
Std. deviance resid.
305
218
477
Theoretical Quantiles
glm(numall ~ nrel + rosn + nrel * rosn)


Scale−Location
√|Std. deviance resid.|
305
218
477
Predicted values
glm(numall ~ nrel + rosn + nrel * rosn)

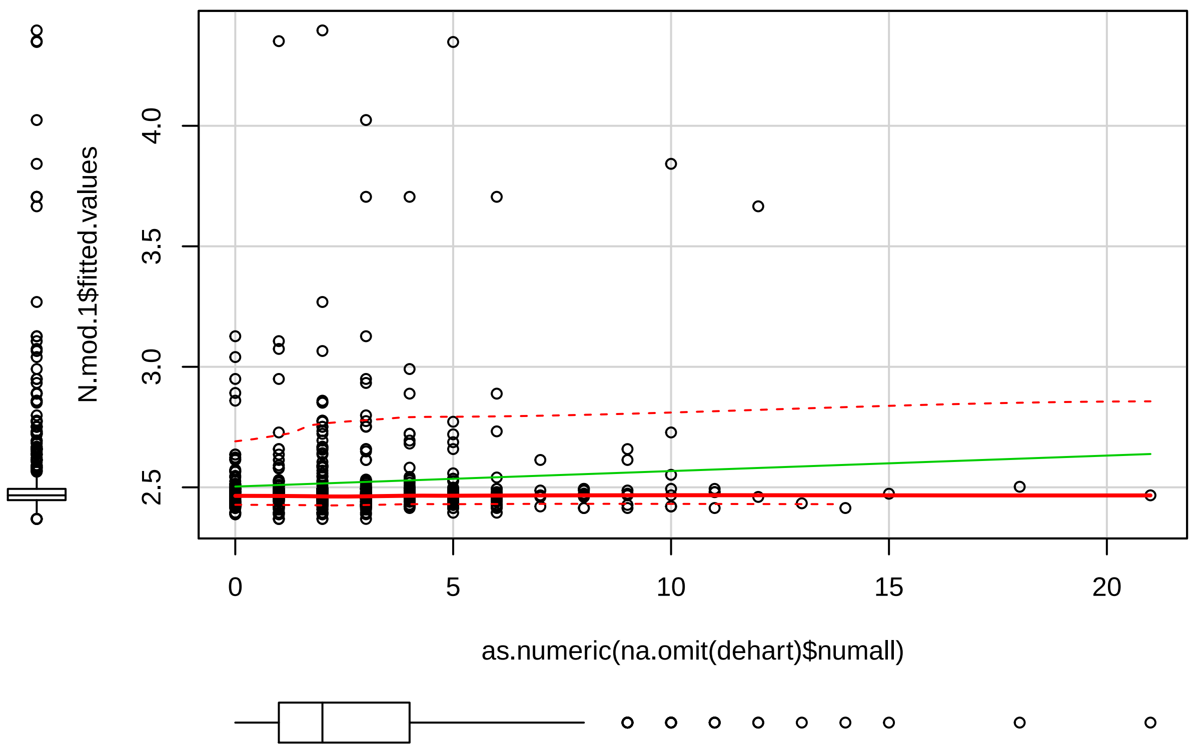



```r
# Pearson residuals
residualPlots(N.mod.1)
```

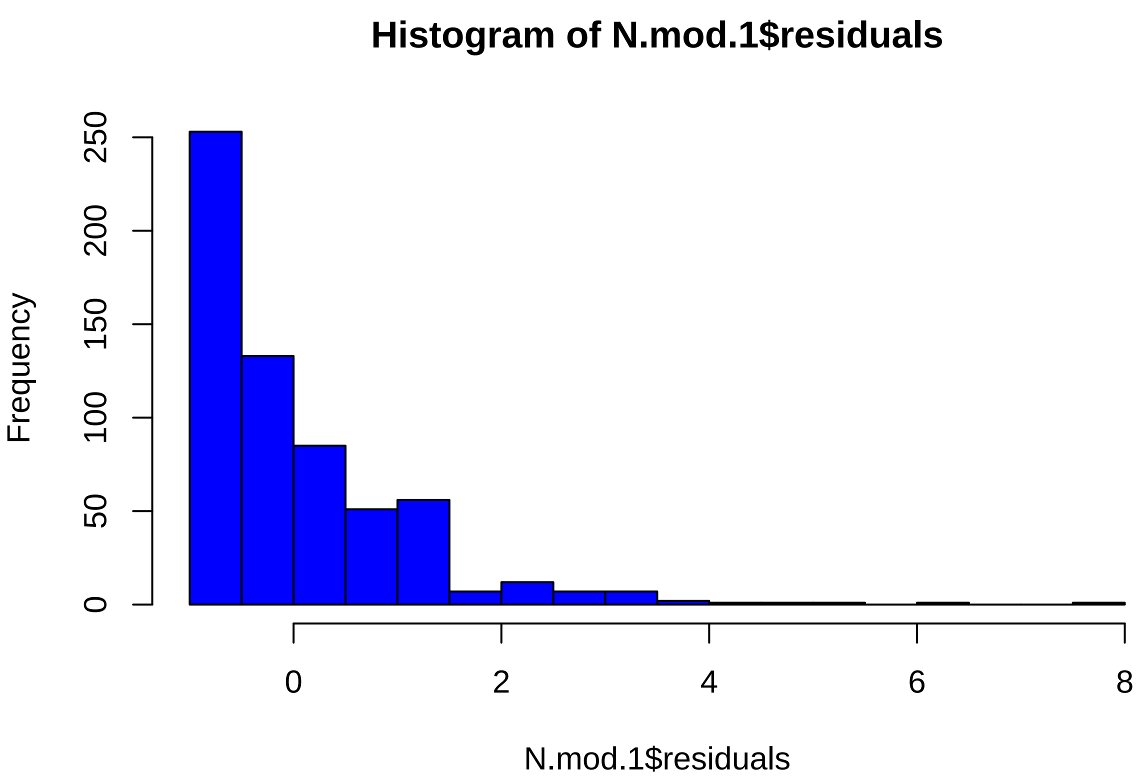



```
##        Test stat Pr(>|t|)
## nrel       8.826    0.003
## rosn       1.934    0.164
```

```r
# scatterplot of actual and fitted values
scatterplot(as.numeric(na.omit(dehart)$numall), N.mod.1$fitted.values)
```




```r
# histogram of residuals
hist(N.mod.1$residuals, breaks = 20, col = "blue")
```



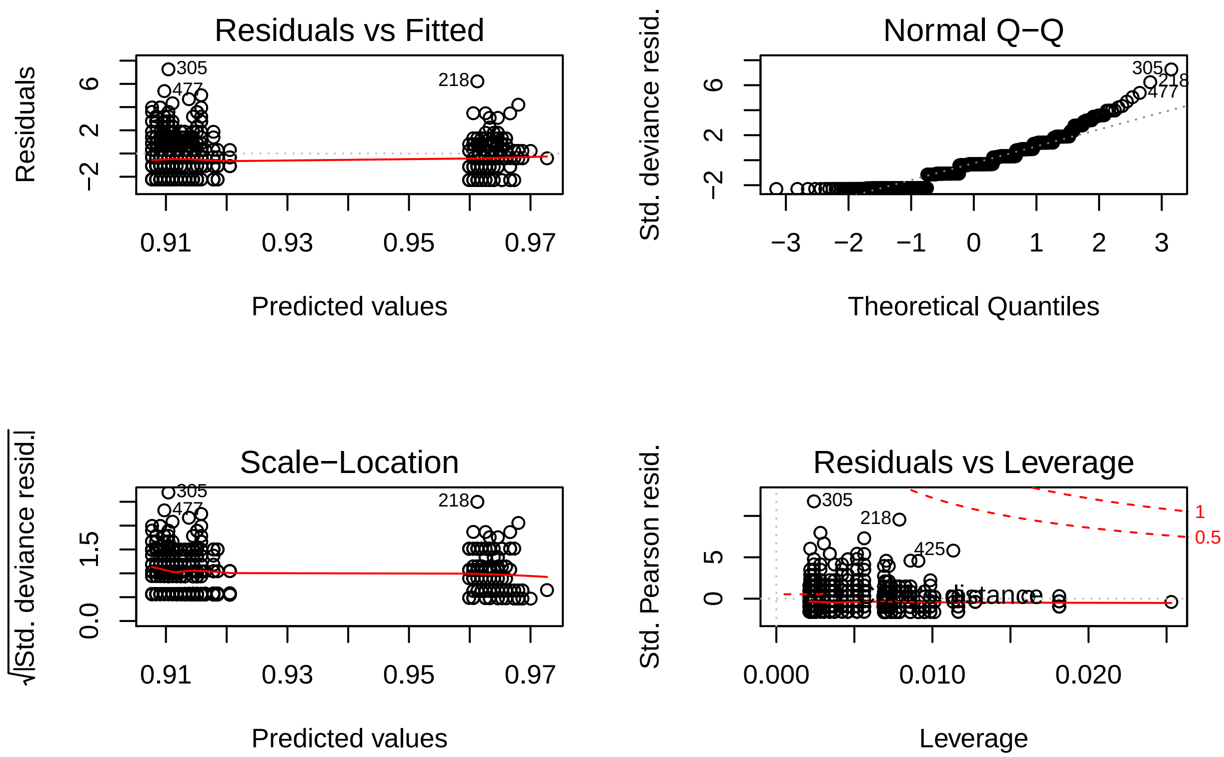
```r
# AGAIN WITH FACTOR FOR NREL
dehart$nrel_f <- dehart$nrel
dehart$nrel_f[dehart$nrel_f != 0] = 1
dehart$nrel_f <- factor(dehart$nrel_f)
levels(dehart$nrel_f) <- c("low","high")

N.mod.0a <- glm(formula = numall ~ nrel_f + rosn, family = poisson(link = "log"), data = na.omit(dehart
summary(N.mod.0a)
```

```
##
## Call:
## glm(formula = numall ~ nrel_f + rosn, family = poisson(link = "log"),
##     data = na.omit(dehart))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2939  -1.1442  -0.3229   0.6693   7.2528
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.934688   0.208906   4.474 7.67e-06 ***
## nrel_fhigh   0.052194   0.059103   0.883    0.377
## rosn        -0.006743   0.060503  -0.111    0.911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1583.5  on 617  degrees of freedom
## Residual deviance: 1582.8  on 615  degrees of freedom
## AIC: 2950.7
##
## Number of Fisher Scoring iterations: 5
```

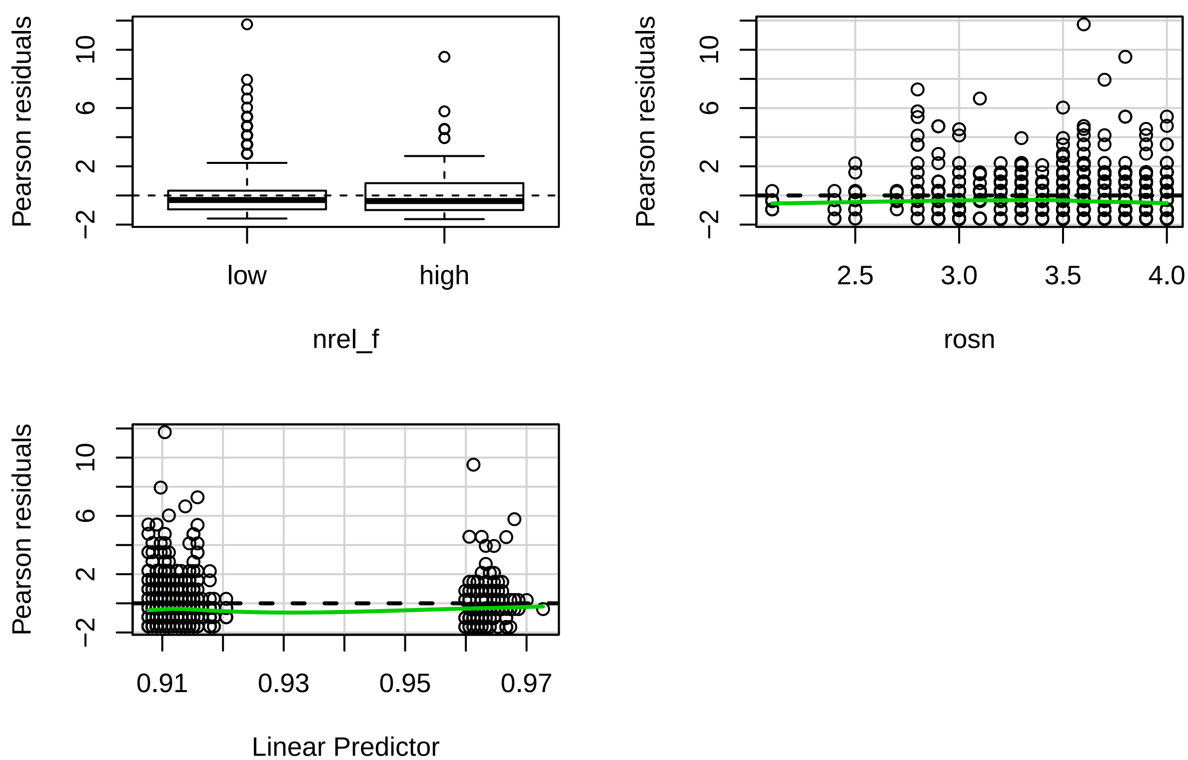
```r
# hypothesis tests for statistical significance
Anova(N.mod.0a, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: numall
##        LR Chisq Df Pr(>Chisq)
## nrel_f  0.77295  1     0.3793
## rosn    0.01241  1     0.9113
```

```r
# diagnostic plots
par(mfrow=c(2,2))
plot(N.mod.0a)
```

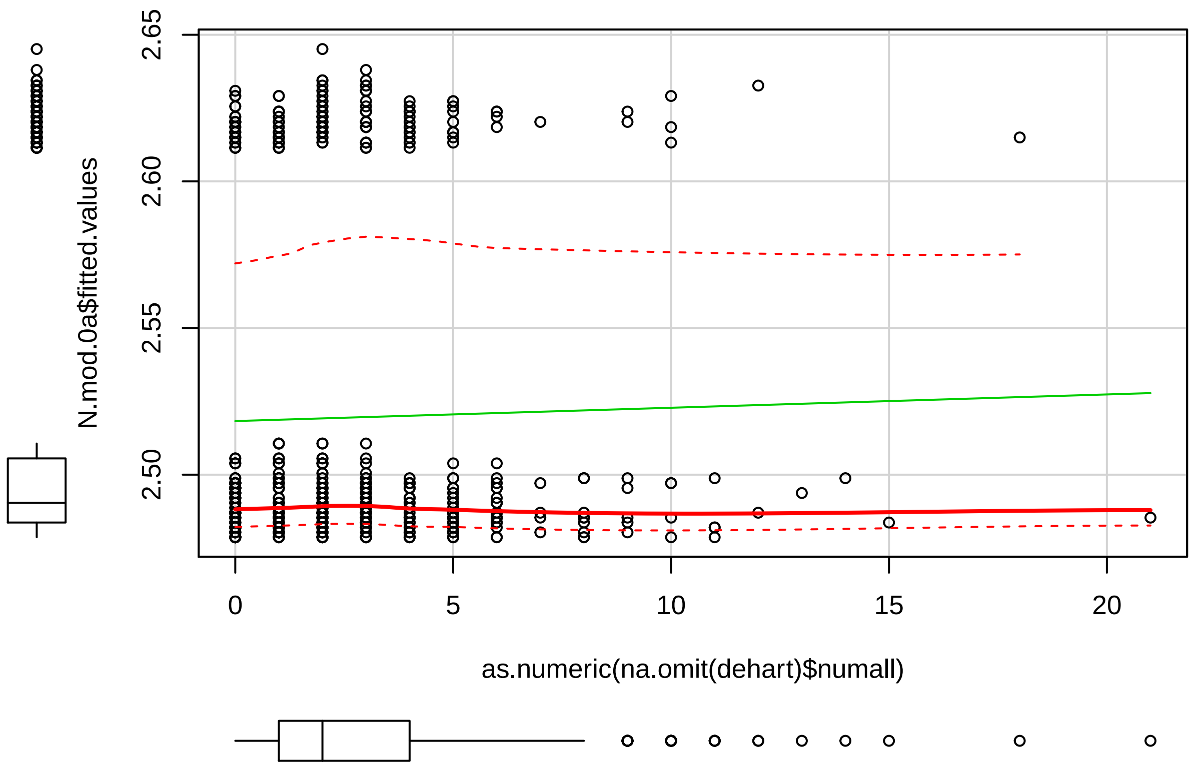



```r
# Pearson residuals
residualPlots(N.mod.0a)
```




```
##         Test stat Pr(>|t|)
## nrel_f         NA       NA
## rosn        2.591    0.107
```

```r
# scatterplot of actual and fitted values
scatterplot(as.numeric(na.omit(dehart)$numall), N.mod.0a$fitted.values)
```




```r
# histogram of residuals
hist(N.mod.0a$residuals, breaks = 20, col = "blue")


N.mod.1a <- glm(formula = numall ~ nrel_f + rosn + nrel_f*rosn, family = poisson(link = "log"), data =
summary(N.mod.1a)
```

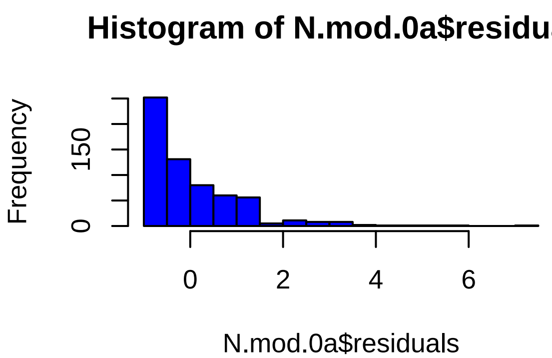
```
##
## Call:
## glm(formula = numall ~ nrel_f + rosn + nrel_f * rosn, family = poisson(link = "log"),
##     data = na.omit(dehart))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.4504  -1.1137  -0.3339   0.5868   7.2275
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)        0.74179    0.23394   3.171  0.00152 **
## nrel_fhigh         1.06068    0.52477   2.021  0.04326 *
## rosn               0.04959    0.06768   0.733  0.46375
## nrel_fhigh:rosn   -0.29204    0.15154  -1.927  0.05397 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
```

```
##     Null deviance: 1583.5  on 617  degrees of freedom
## Residual deviance: 1579.1  on 614  degrees of freedom
## AIC: 2949
##
## Number of Fisher Scoring iterations: 5
```

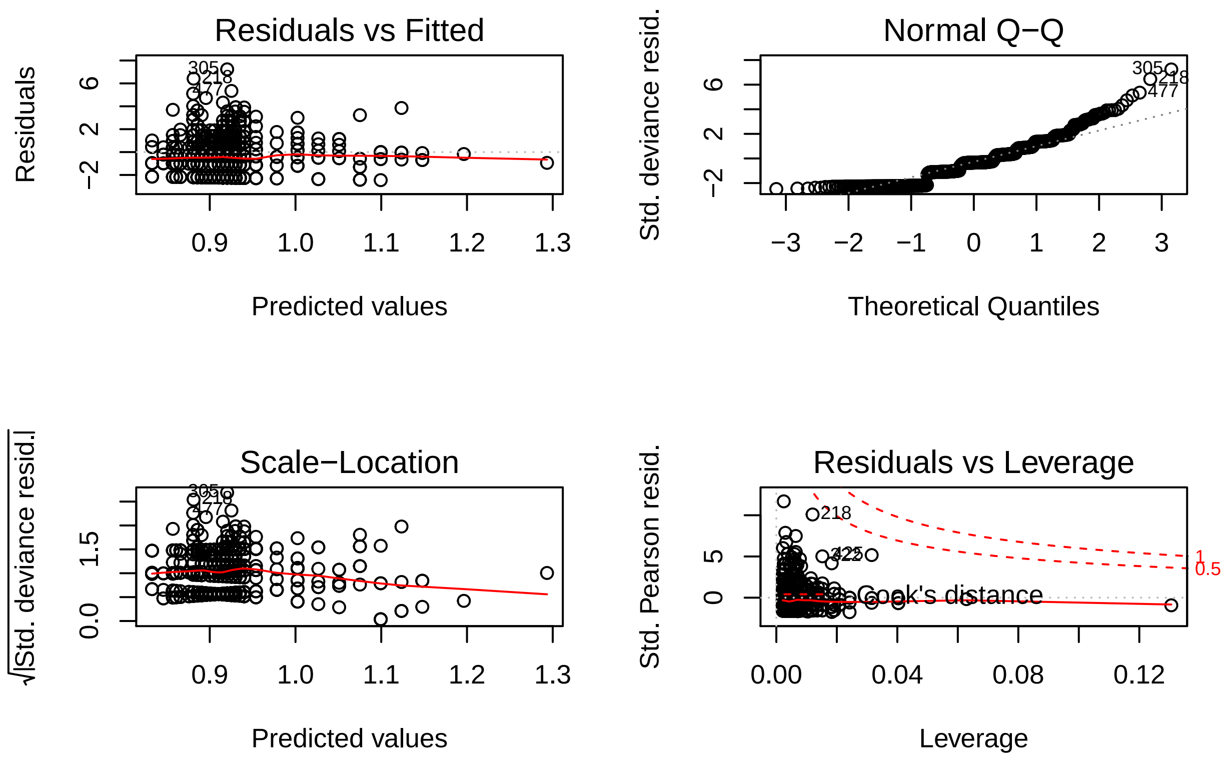
```r
# hypothesis tests for statistical significance
Anova(N.mod.1a, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: numall
##             LR Chisq Df Pr(>Chisq)
## nrel_f        0.7729  1    0.37931
## rosn          0.0124  1    0.91129
## nrel_f:rosn   3.6538  1    0.05594 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

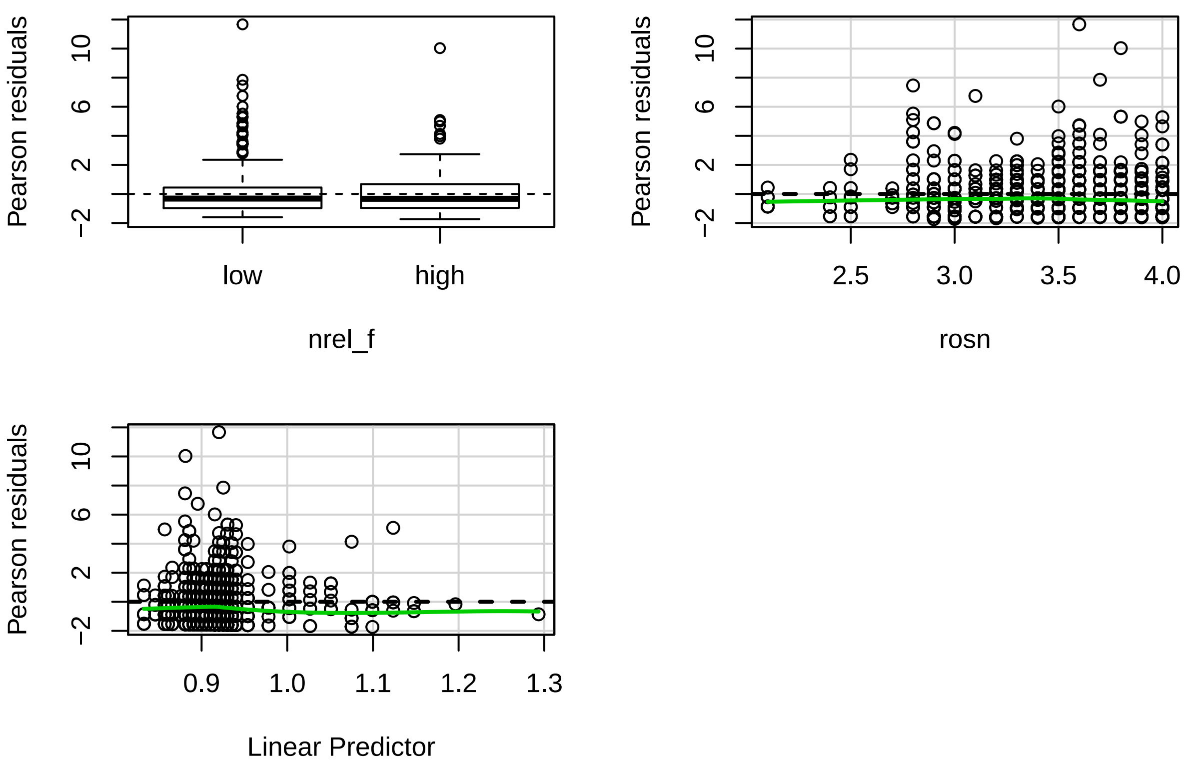
```r
# diagnostic plots
par(mfrow=c(2,2))
```




```r
plot(N.mod.1a)
```

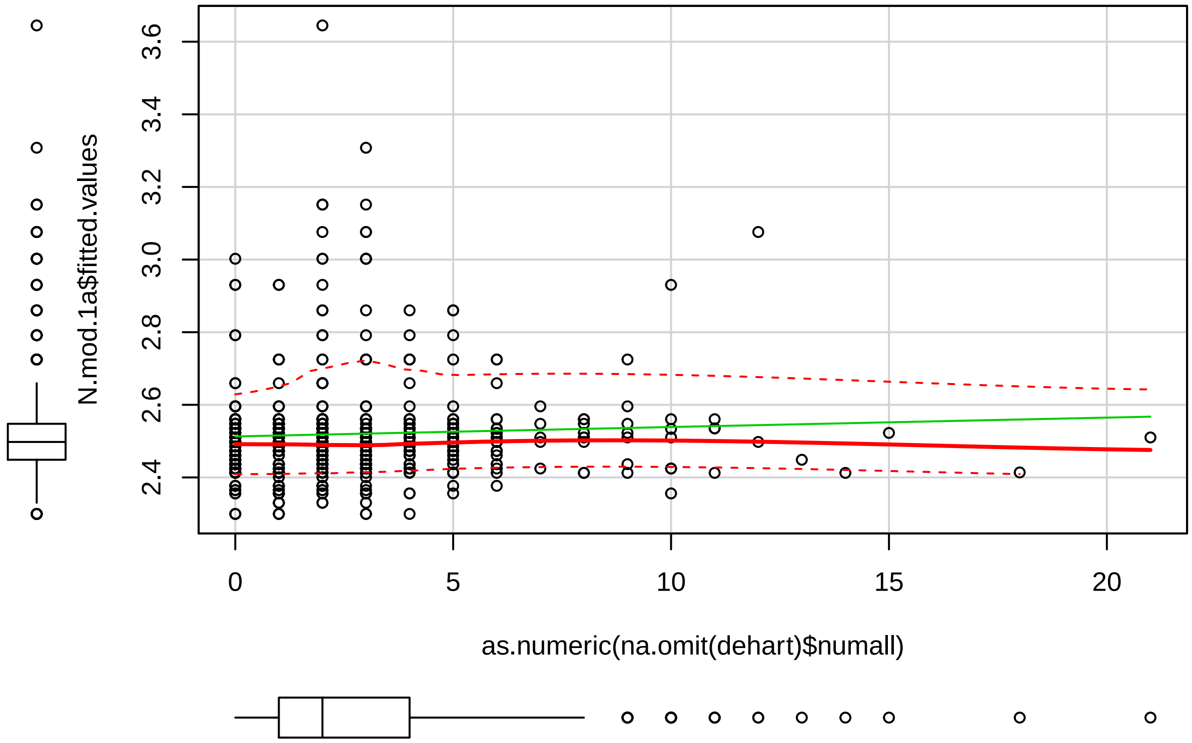



```r
# Pearson residuals
residualPlots(N.mod.1a)
```

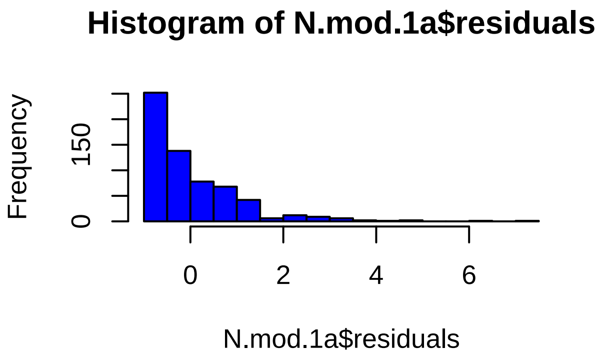



```
##          Test stat Pr(>|t|)
## nrel_f          NA       NA
## rosn         2.118    0.146
```

```r
# scatterplot of actual and fitted values
scatterplot(as.numeric(na.omit(dehart)$numall), N.mod.1a$fitted.values)
```




```r
# histogram of residuals
hist(N.mod.1a$residuals, breaks = 20, col = "blue")



# little significance for anything
```




```r
# INTERMEDIATE MODEL; DAYS OF THE WEEK

N.mod.2 <- glm(formula = numall ~ nrel + prel + rosn + nrel*rosn + prel*rosn + dayweek_f, family = pois
summary(N.mod.2)
```

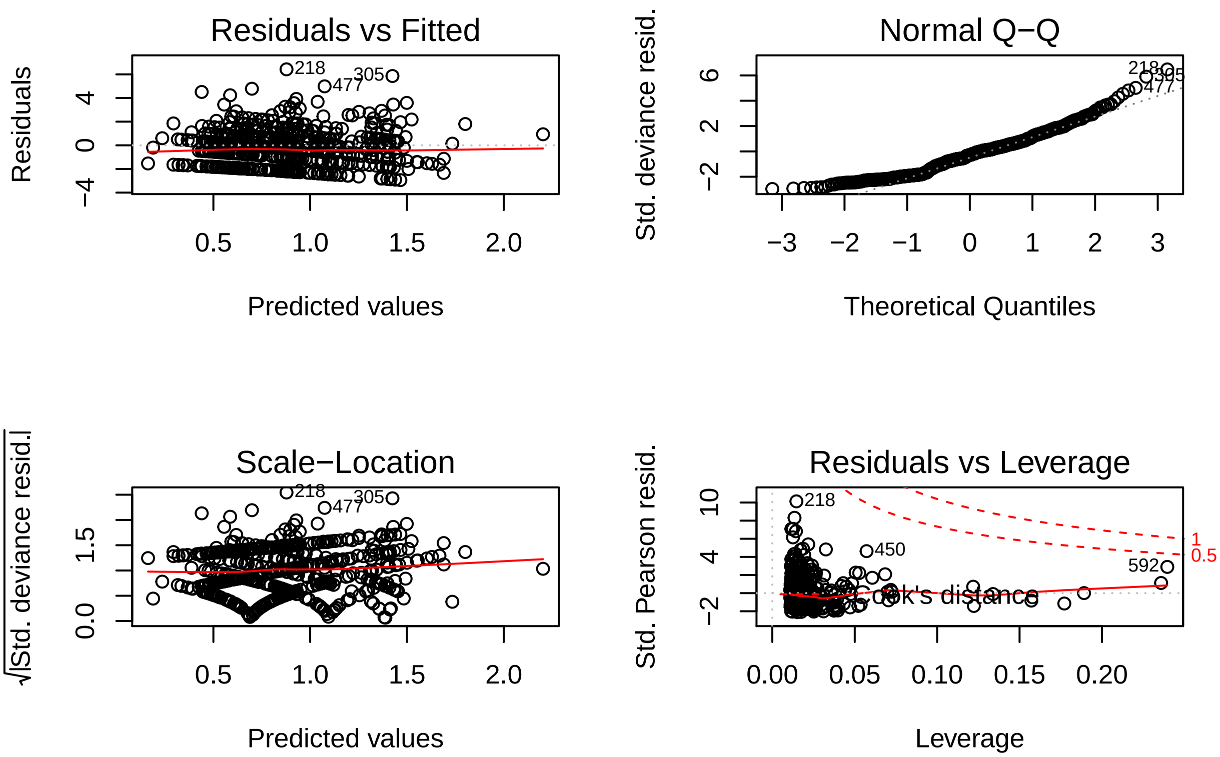
```
##
## Call:
## glm(formula = numall ~ nrel + prel + rosn + nrel * rosn + prel *
##     rosn + dayweek_f, family = poisson(link = "log"), data = na.omit(dehart))
```

```
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.9432  -1.5586  -0.2800   0.5938   6.4240
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.32074    0.35651  -0.900 0.368296
## nrel          0.75464    0.23851   3.164 0.001556 **
## prel          0.26556    0.07484   3.549 0.000387 ***
## rosn          0.26221    0.10174   2.577 0.009955 **
## dayweek_ftue -0.14614    0.11041  -1.324 0.185635
## dayweek_fwed -0.09892    0.10809  -0.915 0.360100
## dayweek_fthu  0.19492    0.10104   1.929 0.053711 .
## dayweek_ffri  0.36276    0.09749   3.721 0.000199 ***
## dayweek_fsat  0.66487    0.09225   7.207  5.7e-13 ***
## dayweek_fsun  0.18081    0.10180   1.776 0.075714 .
## nrel:rosn    -0.19302    0.06875  -2.808 0.004991 **
## prel:rosn    -0.06616    0.02216  -2.986 0.002830 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1583.5  on 617  degrees of freedom
## Residual deviance: 1416.1  on 606  degrees of freedom
## AIC: 2802
##
## Number of Fisher Scoring iterations: 5
```

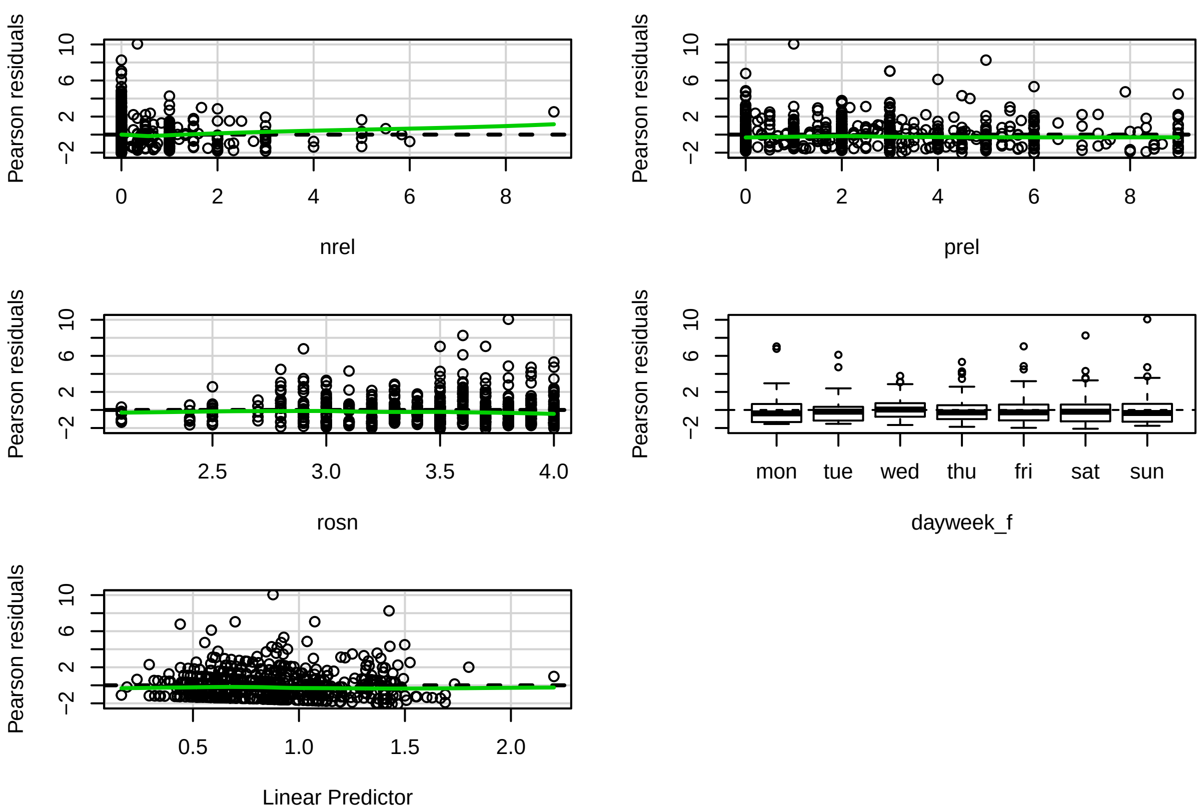
```r
# hypothesis tests for statistical significance
Anova(N.mod.2, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: numall
##           LR Chisq Df Pr(>Chisq)
## nrel        10.510  1   0.001187 **
## prel        16.789  1  4.178e-05 ***
## rosn         0.089  1   0.765921
## dayweek_f  116.305  6  < 2.2e-16 ***
## nrel:rosn    7.642  1   0.005702 **
## prel:rosn    8.807  1   0.003000 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# diagnostic plots
par(mfrow=c(2,2))
plot(N.mod.2)
```



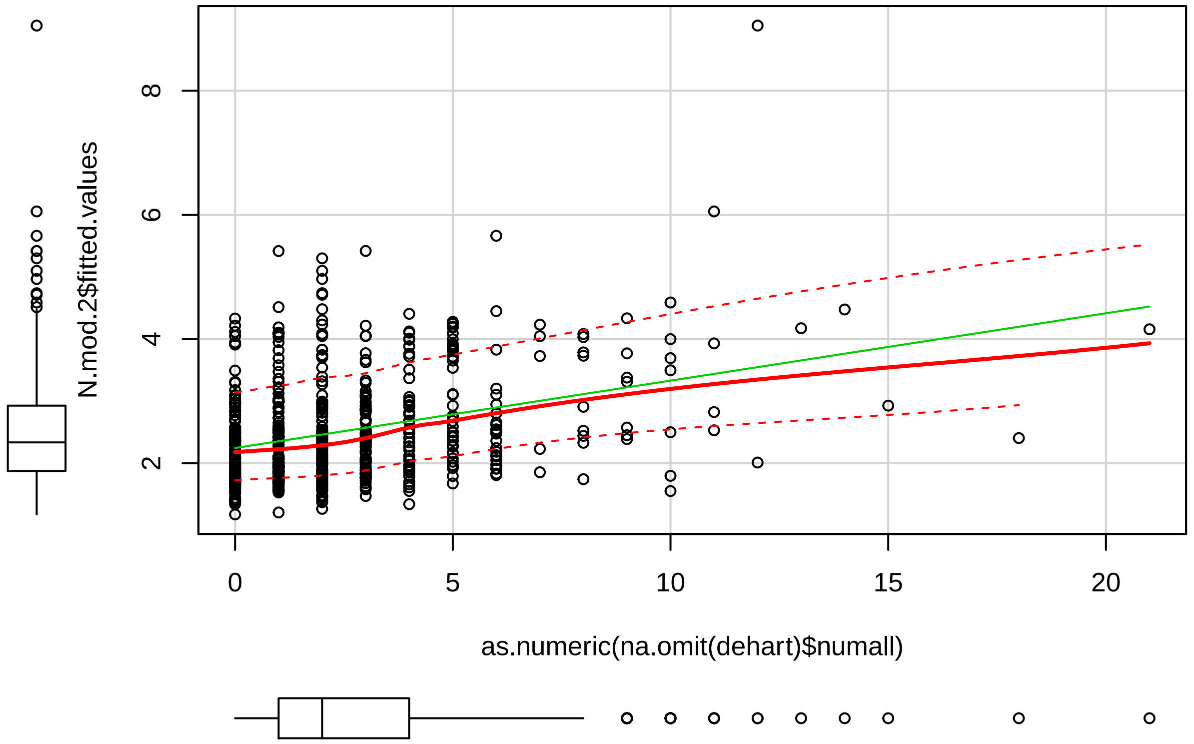

```r
# Pearson residuals
residualPlots(N.mod.2)
```




```
##           Test stat Pr(>|t|)
## nrel          6.696    0.010
```

```
## prel          0.000     0.996
## rosn          2.867     0.090
## dayweek_f        NA        NA
```

```r
# scatterplot of actual and fitted values
scatterplot(as.numeric(na.omit(dehart)$numall), N.mod.2$fitted.values)
```




```r
# histogram of residuals
hist(N.mod.2$residuals, breaks = 20, col = "blue")


# AGAIN WITH FACTOR FOR NREL
N.mod.2a <- glm(formula = numall ~ nrel + prel + rosn + nrel*rosn + prel*rosn + dayweek_f, family = poi
summary(N.mod.2a)
```

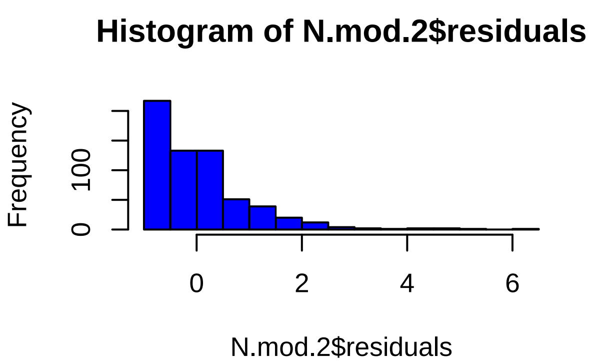
```
##
## Call:
## glm(formula = numall ~ nrel + prel + rosn + nrel * rosn + prel *
##     rosn + dayweek_f, family = poisson(link = "log"), data = na.omit(dehart))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.9432  -1.5586  -0.2800   0.5938   6.4240
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.32074    0.35651  -0.900 0.368296
## nrel          0.75464    0.23851   3.164 0.001556 **
## prel          0.26556    0.07484   3.549 0.000387 ***
## rosn          0.26221    0.10174   2.577 0.009955 **
## dayweek_ftue -0.14614    0.11041  -1.324 0.185635
```

```
## dayweek_fwed -0.09892    0.10809  -0.915 0.360100
## dayweek_fthu  0.19492    0.10104   1.929 0.053711 .
## dayweek_ffri  0.36276    0.09749   3.721 0.000199 ***
## dayweek_fsat  0.66487    0.09225   7.207  5.7e-13 ***
## dayweek_fsun  0.18081    0.10180   1.776 0.075714 .
## nrel:rosn    -0.19302    0.06875  -2.808 0.004991 **
## prel:rosn    -0.06616    0.02216  -2.986 0.002830 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1583.5  on 617  degrees of freedom
## Residual deviance: 1416.1  on 606  degrees of freedom
## AIC: 2802
##
## Number of Fisher Scoring iterations: 5
```

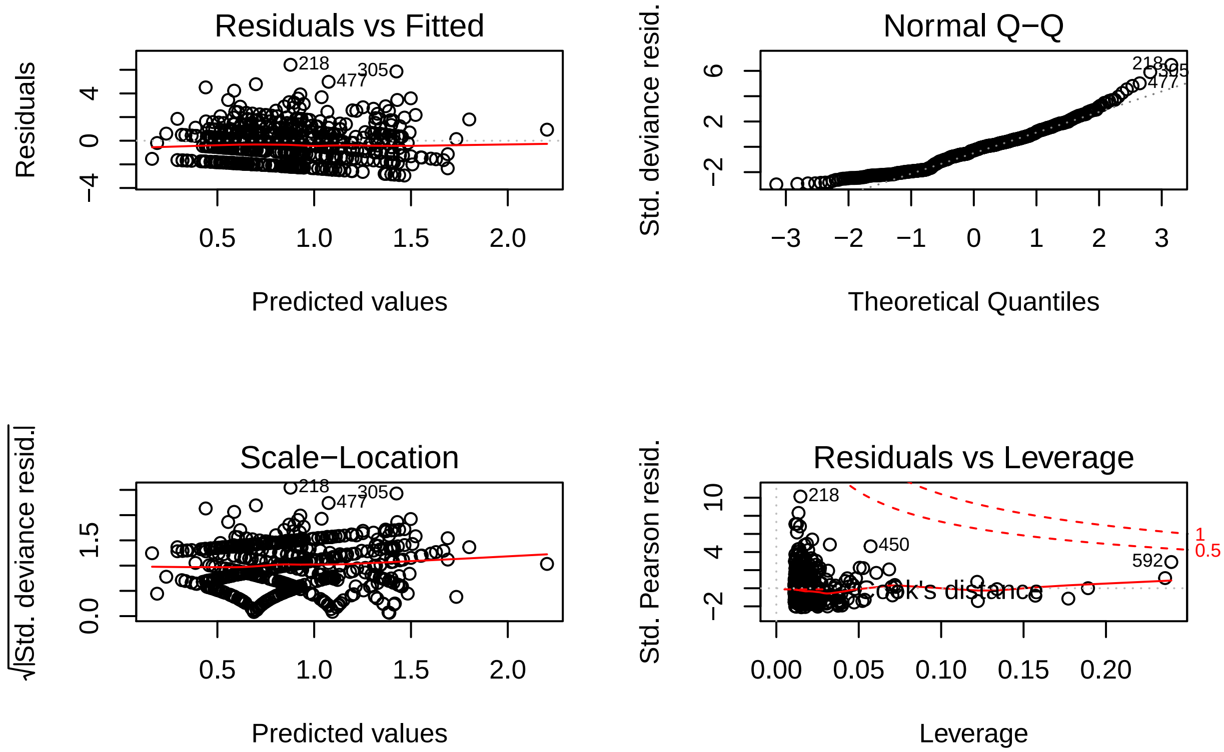
```
# hypothesis tests for statistical significance
Anova(N.mod.2a, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: numall
##           LR Chisq Df Pr(>Chisq)
## nrel        10.510  1   0.001187 **
## prel        16.789  1  4.178e-05 ***
## rosn         0.089  1   0.765921
## dayweek_f  116.305  6  < 2.2e-16 ***
## nrel:rosn    7.642  1   0.005702 **
## prel:rosn    8.807  1   0.003000 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

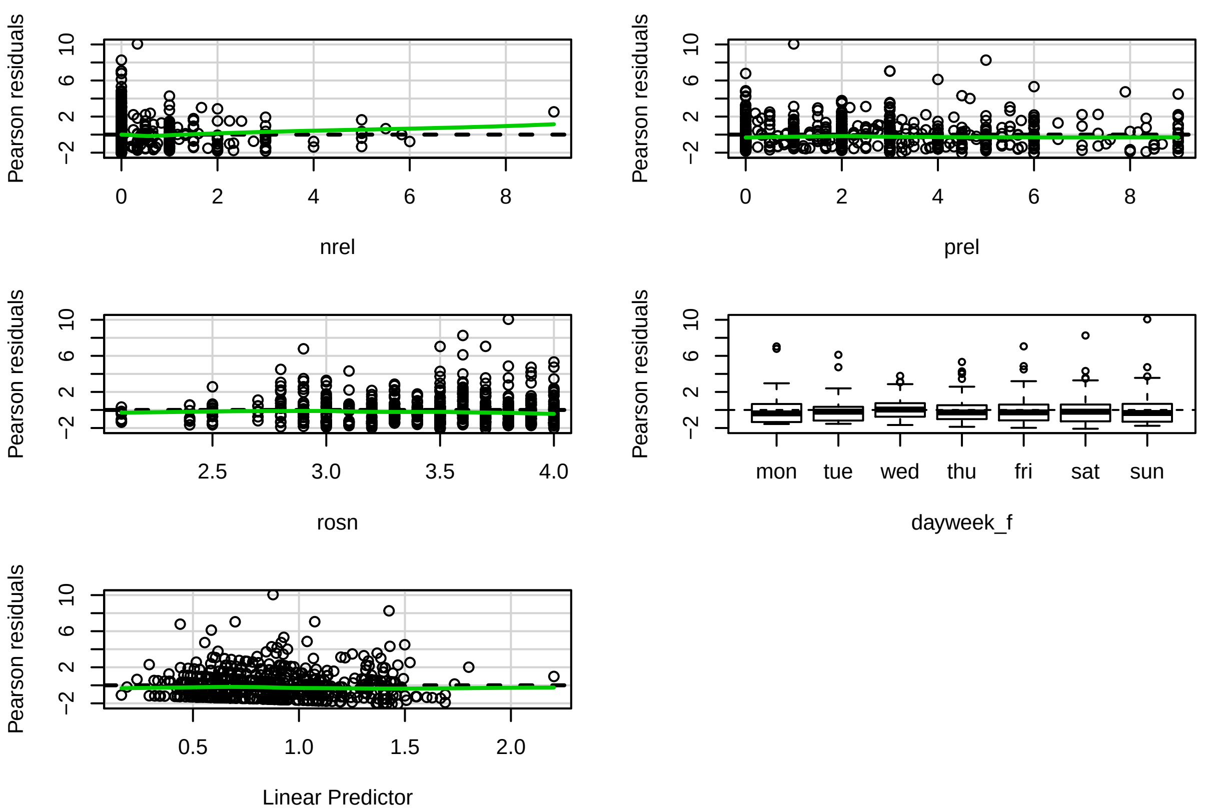
```
# diagnostic plots
par(mfrow=c(2,2))
```




```
plot(N.mod.2a)
```



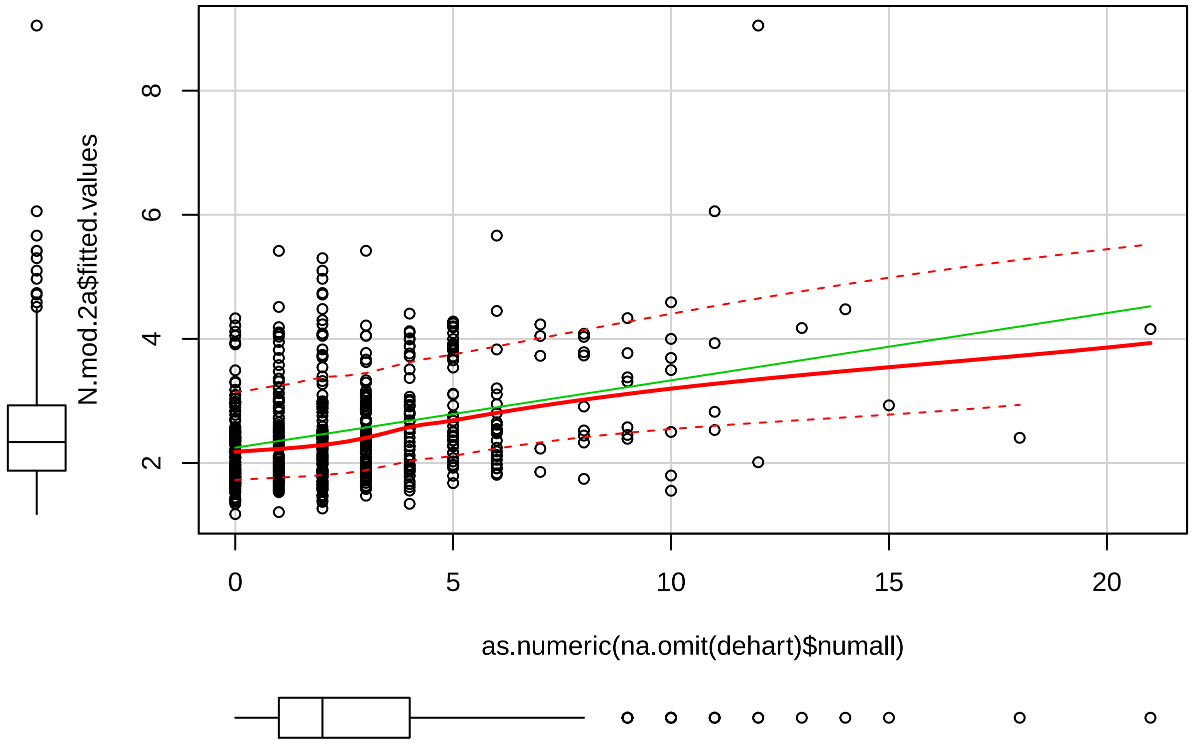

```r
# Pearson residuals
residualPlots(N.mod.2a)
```



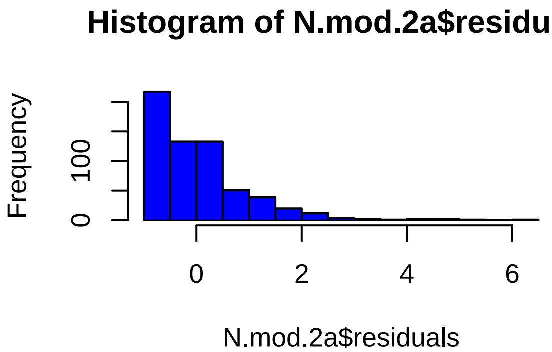

```
##          Test stat Pr(>|t|)
## nrel         6.696    0.010
```

```
## prel           0.000     0.996
## rosn           2.867     0.090
## dayweek_f         NA        NA
```

```
# scatterplot of actual and fitted values
scatterplot(as.numeric(na.omit(dehart)$numall), N.mod.2a$fitted.values)
```




```
# histogram of residuals
hist(N.mod.2a$residuals, breaks = 20, col = "blue")
```




```
# EVERYTHING MODEL

N.mod.3 <- glm(formula = numall ~ nrel + prel + negevent + posevent + rosn + nrel*rosn + prel*rosn + da
summary(N.mod.3)
```

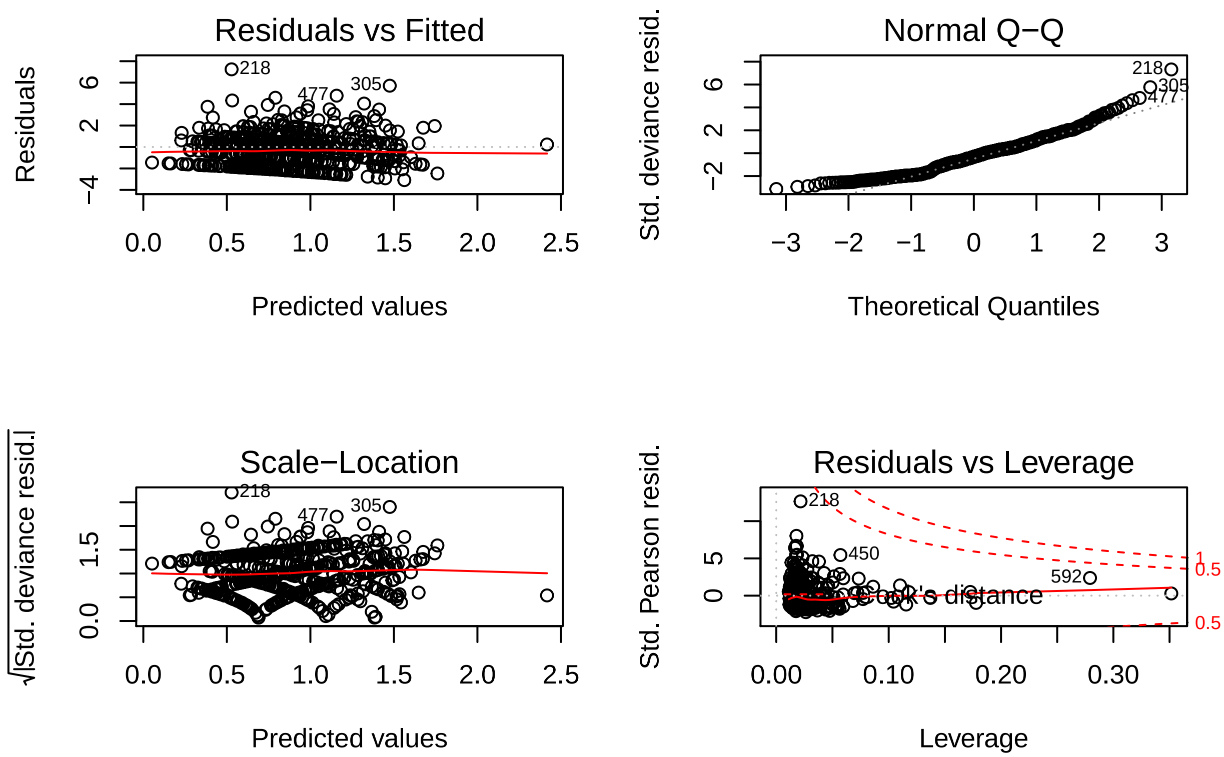
```
##
## Call:
## glm(formula = numall ~ nrel + prel + negevent + posevent + rosn +
##     nrel * rosn + prel * rosn + dayweek_f + gender_f, family = poisson(link = "log"),
##     data = na.omit(dehart))
```

```
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.0888  -1.5326  -0.3009   0.5579   7.2406
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.18977    0.35862  -0.529 0.596697
## nrel            0.67296    0.24626   2.733 0.006282 **
## prel            0.23071    0.07679   3.004 0.002660 **
## negevent       -0.28695    0.07873  -3.645 0.000268 ***
## posevent        0.05717    0.04553   1.256 0.209197
## rosn            0.25820    0.10247   2.520 0.011742 *
## dayweek_ftue   -0.14105    0.11041  -1.278 0.201401
## dayweek_fwed   -0.06779    0.10818  -0.627 0.530885
## dayweek_fthu    0.20608    0.10109   2.039 0.041491 *
## dayweek_ffri    0.35610    0.09750   3.652 0.000260 ***
## dayweek_fsat    0.65559    0.09222   7.109 1.17e-12 ***
## dayweek_fsun    0.15534    0.10198   1.523 0.127684
## gender_ffemale -0.12184    0.05287  -2.304 0.021200 *
## nrel:rosn      -0.15359    0.07125  -2.156 0.031108 *
## prel:rosn      -0.05709    0.02252  -2.535 0.011243 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1583.5  on 617  degrees of freedom
## Residual deviance: 1393.9  on 603  degrees of freedom
## AIC: 2785.9
##
## Number of Fisher Scoring iterations: 5
```

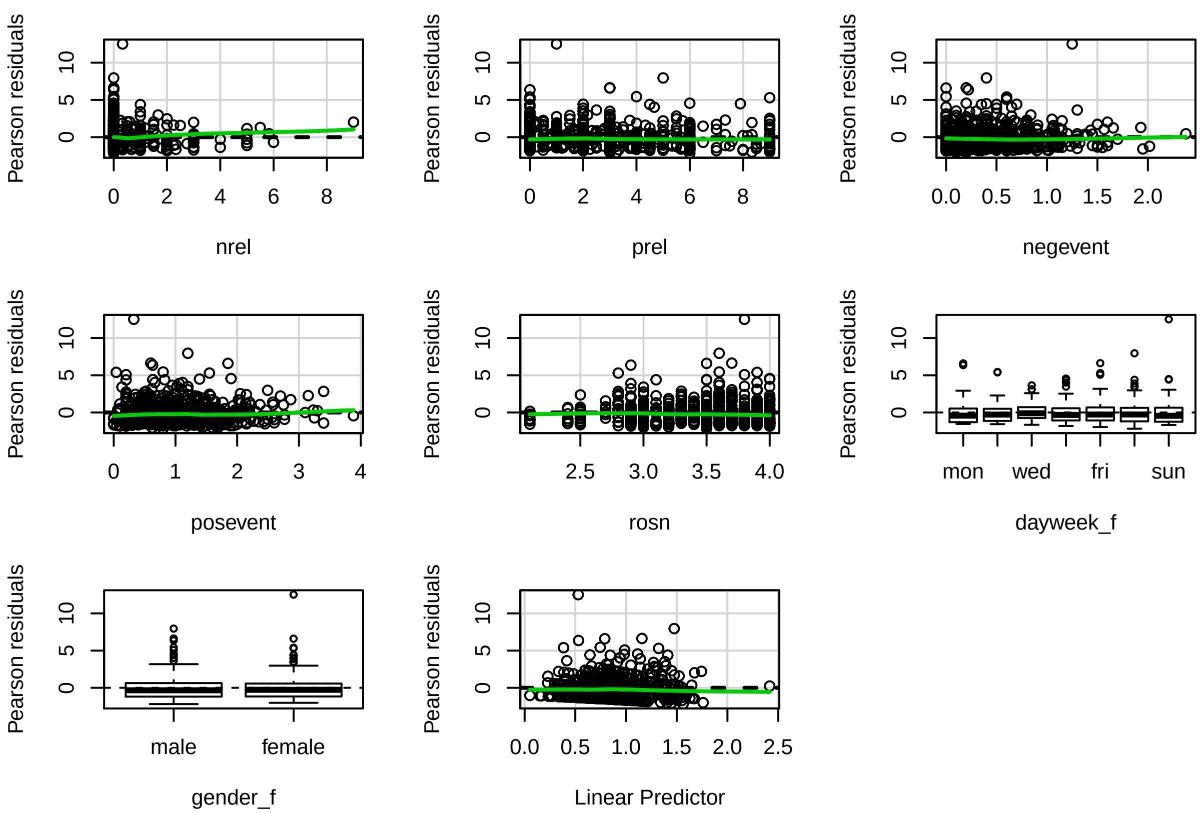
```r
# hypothesis tests for statistical significance
Anova(N.mod.3, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: numall
##           LR Chisq Df Pr(>Chisq)
## nrel        23.990  1  9.684e-07 ***
## prel         8.845  1   0.002938 **
## negevent    13.739  1   0.000210 ***
## posevent     1.558  1   0.211933
## rosn         0.152  1   0.696398
## dayweek_f  109.411  6  < 2.2e-16 ***
## gender_f     5.302  1   0.021296 *
## nrel:rosn    4.527  1   0.033371 *
## prel:rosn    6.355  1   0.011707 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# diagnostic plots
par(mfrow=c(2,2))
plot(N.mod.3)
```



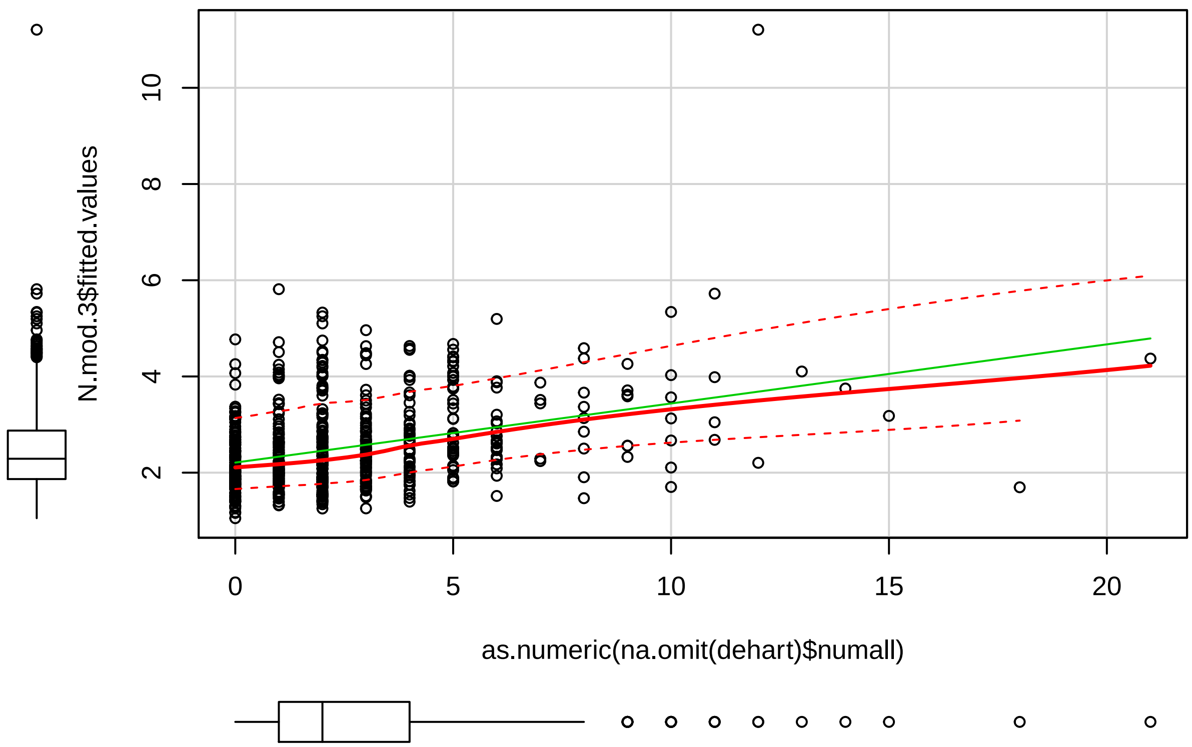

```r
# Pearson residuals
residualPlots(N.mod.3)
```




```
##          Test stat Pr(>|t|)
## nrel         4.860    0.027
```

```
## prel          0.086    0.769
## negevent      0.434    0.510
## posevent      1.369    0.242
## rosn          1.595    0.207
## dayweek_f        NA       NA
## gender_f         NA       NA
```

```r
# scatterplot of actual and fitted values
scatterplot(as.numeric(na.omit(dehart)$numall), N.mod.3$fitted.values)
```




```r
# histogram of residuals
hist(N.mod.3$residuals, breaks = 20, col = "blue")


# AGAIN WITH FACTOR FOR NREL
N.mod.3a <- glm(formula = numall ~ nrel_f + prel + negevent + posevent + rosn + nrel_f*rosn + prel*rosn
summary(N.mod.3a)
```

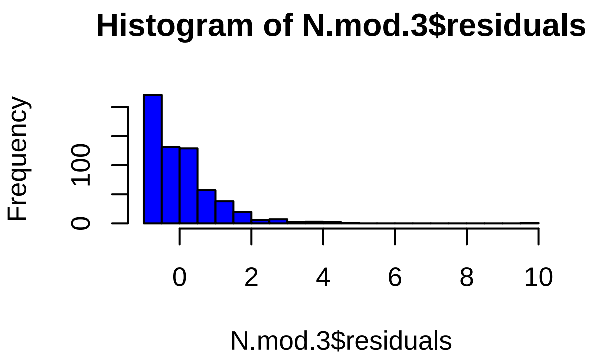
```
##
## Call:
## glm(formula = numall ~ nrel_f + prel + negevent + posevent +
##     rosn + nrel_f * rosn + prel * rosn + dayweek_f + gender_f,
##     family = poisson(link = "log"), data = na.omit(dehart))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.0351  -1.4924  -0.3336   0.5733   6.8959
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -0.29535    0.36850  -0.801  0.42285
```

```
## nrel_fhigh       1.27887    0.54581   2.343  0.01913 *
## prel             0.23940    0.07676   3.119  0.00181 **
## negevent        -0.18588    0.07493  -2.481  0.01311 *
## posevent         0.06225    0.04583   1.358  0.17441
## rosn             0.28017    0.10473   2.675  0.00747 **
## dayweek_ftue    -0.13216    0.11070  -1.194  0.23254
## dayweek_fwed    -0.05320    0.10803  -0.492  0.62240
## dayweek_fthu     0.21604    0.10143   2.130  0.03317 *
## dayweek_ffri     0.38033    0.09755   3.899 9.66e-05 ***
## dayweek_fsat     0.67713    0.09217   7.347 2.03e-13 ***
## dayweek_fsun     0.17687    0.10213   1.732  0.08329 .
## gender_ffemale  -0.12023    0.05294  -2.271  0.02313 *
## nrel_fhigh:rosn -0.32065    0.15819  -2.027  0.04266 *
## prel:rosn       -0.06174    0.02250  -2.744  0.00606 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1583.5  on 617  degrees of freedom
## Residual deviance: 1411.2  on 603  degrees of freedom
## AIC: 2803.1
##
## Number of Fisher Scoring iterations: 5
```

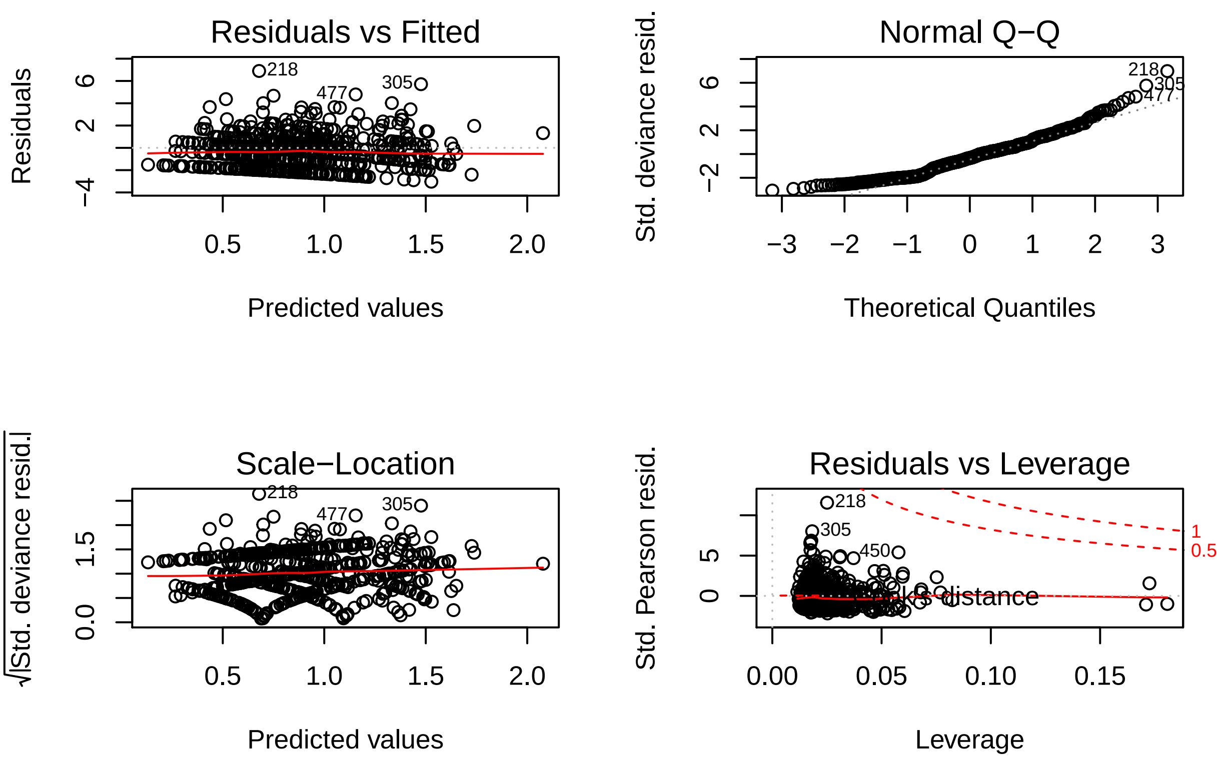
```
# hypothesis tests for statistical significance
Anova(N.mod.3a, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: numall
##               LR Chisq Df Pr(>Chisq)
## nrel_f           7.187  1   0.007344 **
## prel             5.871  1   0.015388 *
## negevent         6.308  1   0.012018 *
## posevent         1.820  1   0.177352
## rosn             0.101  1   0.751226
## dayweek_f      114.269  6  < 2.2e-16 ***
## gender_f         5.151  1   0.023234 *
## nrel_f:rosn      4.044  1   0.044319 *
## prel:rosn        7.444  1   0.006364 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

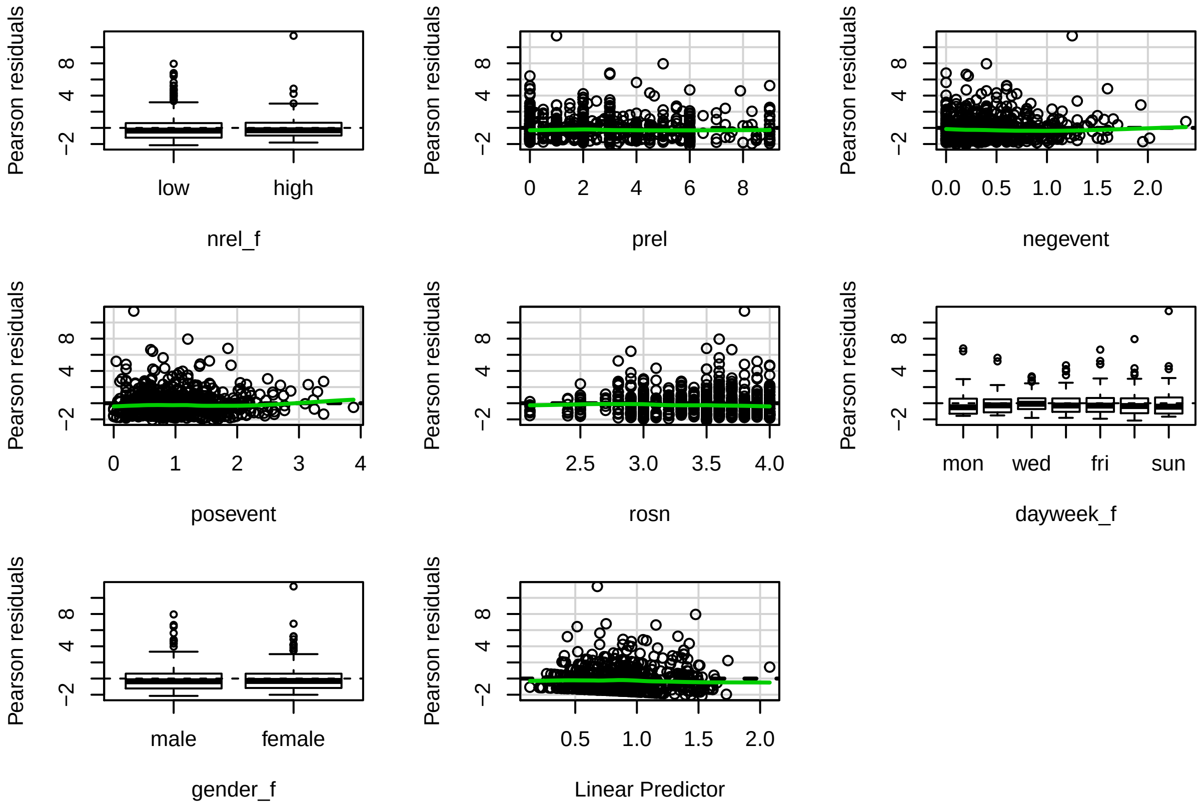
```
# diagnostic plots
par(mfrow=c(2,2))
```




```r
plot(N.mod.3a)
```

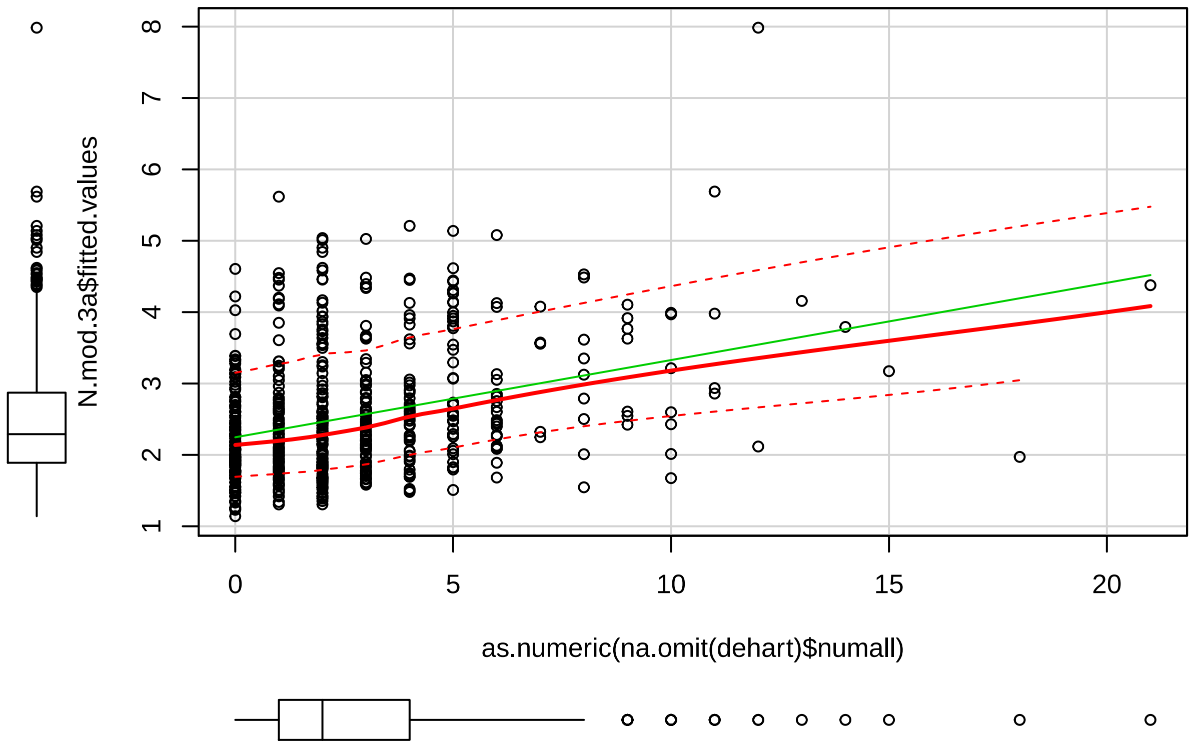



```r
# Pearson residuals
residualPlots(N.mod.3a)
```

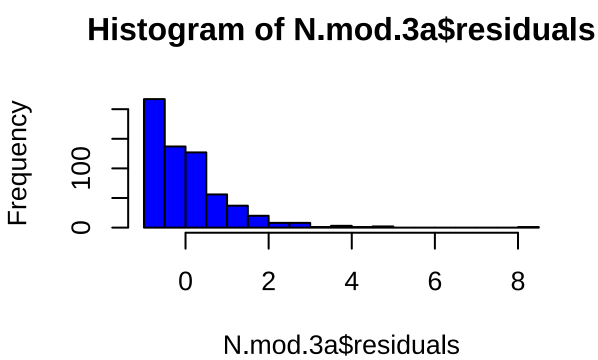



```
##           Test stat Pr(>|t|)
## nrel_f           NA       NA
## prel          0.668    0.414
## negevent      2.283    0.131
## posevent      3.083    0.079
## rosn          2.133    0.144
## dayweek_f        NA       NA
## gender_f         NA       NA
```

```
# scatterplot of actual and fitted values
scatterplot(as.numeric(na.omit(dehart)$numall), N.mod.3a$fitted.values)
```




```r
# histogram of residuals
hist(N.mod.3a$residuals, breaks = 20, col = "blue")
```



DESIRE TO DRINK

```r
# DESIRE TO DRINK; ORDINAL LOGISTIC MODEL
# No significant interaction effect

# number of 'bins' :
N <- 3

dehart$desired_o <- cut(dehart$desired,N)
levels(dehart$desired_o) <- c("low","med","high")

library(package = MASS)
```

```
## Warning: package 'MASS' was built under R version 3.4.3

##
## Attaching package: 'MASS'

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

```
##
##     select
```

```r
# BASIC MODEL

D.mod.0 <- polr(formula = desired_o ~ nrel + rosn, method = "logistic", data = na.omit(dehart))
summary(D.mod.0)
```

```
##
## Re-fitting to get Hessian
## Call:
## polr(formula = desired_o ~ nrel + rosn, data = na.omit(dehart),
##     method = "logistic")
##
## Coefficients:
##        Value Std. Error t value
## nrel  0.2327     0.0842   2.763
## rosn -0.6614     0.1820  -3.633
##
## Intercepts:
##          Value   Std. Error t value
## low|med  -3.2555  0.6391    -5.0940
## med|high -1.2734  0.6258    -2.0347
##
## Residual Deviance: 1301.456
## AIC: 1309.456
```

```r
# hypothesis tests for statistical significance
Anova(D.mod.0, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: desired_o
##      LR Chisq Df Pr(>Chisq)
## nrel   7.9511  1  0.0048058 **
## rosn  13.4152  1  0.0002496 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
D.mod.1 <- polr(formula = desired_o ~ nrel + rosn + nrel*rosn, method = "logistic", data = na.omit(deha
summary(D.mod.1)
```

```
##
## Re-fitting to get Hessian
## Call:
## polr(formula = desired_o ~ nrel + rosn + nrel * rosn, data = na.omit(dehart),
##     method = "logistic")
##
## Coefficients:
##              Value Std. Error t value
## nrel        0.7295     0.8564  0.8518
## rosn       -0.6242     0.1925 -3.2418
## nrel:rosn -0.1429     0.2448 -0.5838
##
## Intercepts:
```

```
##          Value   Std. Error t value
## low|med  -3.1279  0.6741    -4.6401
## med|high -1.1457  0.6619    -1.7311
##
## Residual Deviance: 1301.11
## AIC: 1311.11
```

```r
# hypothesis tests for statistical significance
Anova(D.mod.1, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: desired_o
##           LR Chisq Df Pr(>Chisq)
## nrel        7.9511  1  0.0048058 **
## rosn       13.4152  1  0.0002496 ***
## nrel:rosn   0.3467  1  0.5559720
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# factors for nrel
D.mod.0a <- polr(formula = desired_o ~ nrel_f + rosn, method = "logistic", data = na.omit(dehart))
summary(D.mod.0a)
```

```
##
## Re-fitting to get Hessian
```

```
## Call:
## polr(formula = desired_o ~ nrel_f + rosn, data = na.omit(dehart),
##     method = "logistic")
##
## Coefficients:
##               Value Std. Error t value
## nrel_fhigh  0.2042     0.1776   1.149
## rosn       -0.6536     0.1822  -3.588
##
## Intercepts:
##          Value   Std. Error t value
## low|med  -3.2532  0.6387    -5.0933
## med|high -1.2863  0.6256    -2.0562
##
## Residual Deviance: 1308.084
## AIC: 1316.084
```

```r
# hypothesis tests for statistical significance
Anova(D.mod.0a, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: desired_o
##        LR Chisq Df Pr(>Chisq)
## nrel_f   1.3231  1  0.2500292
## rosn    13.0747  1  0.0002993 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
D.mod.1a <- polr(formula = desired_o ~ nrel_f + rosn + nrel_f*rosn, method = "logistic", data = na.omit
summary(D.mod.1a)
```

```
##
## Re-fitting to get Hessian
## Call:
## polr(formula = desired_o ~ nrel_f + rosn + nrel_f * rosn, data = na.omit(dehart),
##     method = "logistic")
##
## Coefficients:
##                    Value Std. Error t value
## nrel_fhigh       1.9635     1.6524   1.188
## rosn            -0.5624     0.2006  -2.804
## nrel_fhigh:rosn -0.5057     0.4720  -1.071
##
## Intercepts:
##          Value   Std. Error t value
## low|med  -2.9419  0.6995    -4.2058
## med|high -0.9726  0.6889    -1.4118
##
## Residual Deviance: 1306.926
## AIC: 1316.926
```

```r
# hypothesis tests for statistical significance
Anova(D.mod.1a, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: desired_o
##             LR Chisq Df Pr(>Chisq)
## nrel_f        1.3231  1  0.2500292
## rosn         13.0747  1  0.0002993 ***
## nrel_f:rosn   1.1583  1  0.2818095
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# INTERMEDIATE MODEL; DAYS OF THE WEEK

D.mod.2 <- polr(formula = desired_o ~ nrel + prel + rosn + nrel*rosn + prel*rosn + dayweek_f, method =
summary(D.mod.2)
```

```
##
## Re-fitting to get Hessian
## Call:
## polr(formula = desired_o ~ nrel + prel + rosn + nrel * rosn +
##     prel * rosn + dayweek_f, data = na.omit(dehart), method = "logistic")
##
## Coefficients:
##                  Value Std. Error t value
## nrel           1.04024    0.88741  1.1722
## prel           0.40201    0.25436  1.5805
## rosn          -0.41787    0.28431 -1.4698
## dayweek_ftue  0.17013    0.28610  0.5947
## dayweek_fwed  0.24405    0.28909  0.8442
```

```
## dayweek_fthu  0.43213    0.29240  1.4779
## dayweek_ffri  0.58517    0.28812  2.0310
## dayweek_fsat  0.72353    0.29169  2.4804
## dayweek_fsun -0.36303    0.29268 -1.2404
## nrel:rosn    -0.21341    0.25329 -0.8425
## prel:rosn    -0.08328    0.07412 -1.1237
##
## Intercepts:
##          Value   Std. Error t value
## low|med  -1.8847  0.9971    -1.8901
## med|high  0.1882  0.9927     0.1896
##
## Residual Deviance: 1265.545
## AIC: 1291.545
```

```r
# hypothesis tests for statistical significance
Anova(D.mod.2, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: desired_o
##            LR Chisq Df Pr(>Chisq)
## nrel        11.8245  1  0.0005846 ***
## prel        12.9135  1  0.0003262 ***
## rosn        14.5145  1  0.0001391 ***
## dayweek_f   19.7301  6  0.0030929 **
## nrel:rosn    0.7236  1  0.3949748
## prel:rosn    1.2785  1  0.2581742
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# AGAIN WITH FACTOR FOR NREL
D.mod.2a <- polr(formula = desired_o ~ nrel + prel + rosn + nrel*rosn + prel*rosn + dayweek_f, method =
summary(D.mod.2a)
```

```
##
## Re-fitting to get Hessian
## Call:
## polr(formula = desired_o ~ nrel + prel + rosn + nrel * rosn +
##     prel * rosn + dayweek_f, data = na.omit(dehart), method = "logistic")
##
## Coefficients:
##                  Value Std. Error t value
## nrel           1.04024    0.88741  1.1722
## prel           0.40201    0.25436  1.5805
## rosn          -0.41787    0.28431 -1.4698
## dayweek_ftue  0.17013    0.28610  0.5947
## dayweek_fwed  0.24405    0.28909  0.8442
## dayweek_fthu  0.43213    0.29240  1.4779
## dayweek_ffri  0.58517    0.28812  2.0310
## dayweek_fsat  0.72353    0.29169  2.4804
## dayweek_fsun -0.36303    0.29268 -1.2404
## nrel:rosn    -0.21341    0.25329 -0.8425
## prel:rosn    -0.08328    0.07412 -1.1237
##
```

```
## Intercepts:
##          Value   Std. Error t value
## low|med  -1.8847  0.9971    -1.8901
## med|high  0.1882  0.9927     0.1896
##
## Residual Deviance: 1265.545
## AIC: 1291.545
```

```
# hypothesis tests for statistical significance
Anova(D.mod.2a, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: desired_o
##            LR Chisq Df Pr(>Chisq)
## nrel        11.8245  1  0.0005846 ***
## prel        12.9135  1  0.0003262 ***
## rosn        14.5145  1  0.0001391 ***
## dayweek_f   19.7301  6  0.0030929 **
## nrel:rosn    0.7236  1  0.3949748
## prel:rosn    1.2785  1  0.2581742
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# EVERYTHING MODEL

D.mod.3 <- polr(formula = desired_o ~ nrel + prel + negevent + posevent + rosn + nrel*rosn + prel*rosn +
summary(D.mod.3)
```

```
##
## Re-fitting to get Hessian
## Call:
## polr(formula = desired_o ~ nrel + prel + negevent + posevent +
##     rosn + nrel * rosn + prel * rosn + dayweek_f + gender_f,
##     data = na.omit(dehart), method = "logistic")
##
## Coefficients:
##                   Value Std. Error t value
## nrel             1.06457    0.89038  1.1956
## prel             0.36682    0.25715  1.4265
## negevent         0.07835    0.23000  0.3406
## posevent         0.29010    0.14827  1.9565
## rosn            -0.38060    0.28738 -1.3244
## dayweek_ftue     0.17462    0.28700  0.6085
## dayweek_fwed     0.23381    0.29036  0.8052
## dayweek_fthu     0.44950    0.29310  1.5336
## dayweek_ffri     0.57864    0.28920  2.0008
## dayweek_fsat     0.73233    0.29259  2.5029
## dayweek_fsun    -0.36900    0.29467 -1.2522
## gender_ffemale -0.37536    0.16101 -2.3313
## nrel:rosn       -0.21804    0.25446 -0.8569
## prel:rosn       -0.08630    0.07465 -1.1561
##
## Intercepts:
##          Value   Std. Error t value
```

```
## low|med  -1.7482  1.0062    -1.7375
## med|high  0.3456  1.0024     0.3448
##
## Residual Deviance: 1257.573
## AIC: 1289.573
```

```
# hypothesis tests for statistical significance
Anova(D.mod.3, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: desired_o
##           LR Chisq Df Pr(>Chisq)
## nrel        9.8935  1  0.0016586 **
## prel        3.3432  1  0.0674818 .
## negevent    0.1161  1  0.7332954
## posevent    3.8755  1  0.0489964 *
## rosn       12.8532  1  0.0003369 ***
## dayweek_f  19.9499  6  0.0028268 **
## gender_f    5.4631  1  0.0194223 *
## nrel:rosn   0.7440  1  0.3883749
## prel:rosn   1.3545  1  0.2444937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# AGAIN WITH FACTOR FOR NREL
D.mod.3a <- polr(formula = desired_o ~ nrel_f + prel + negevent + posevent + rosn + nrel_f*rosn + prel*r
summary(D.mod.3a)
```

```
##
## Re-fitting to get Hessian
```

```
## Call:
## polr(formula = desired_o ~ nrel_f + prel + negevent + posevent +
##     rosn + nrel_f * rosn + prel * rosn + dayweek_f + gender_f,
##     data = na.omit(dehart), method = "logistic")
##
## Coefficients:
##                      Value Std. Error t value
## nrel_fhigh       2.37635     1.7146  1.3860
## prel             0.37847     0.2555  1.4811
## negevent         0.31361     0.2197  1.4276
## posevent         0.27271     0.1475  1.8492
## rosn            -0.31480     0.2917 -1.0791
## dayweek_ftue     0.19285     0.2878  0.6702
## dayweek_fwed     0.26519     0.2904  0.9132
## dayweek_fthu     0.45382     0.2936  1.5457
## dayweek_ffri     0.61483     0.2894  2.1248
## dayweek_fsat     0.75681     0.2923  2.5889
## dayweek_fsun    -0.33017     0.2954 -1.1178
## gender_ffemale  -0.36240     0.1610 -2.2515
## nrel_fhigh:rosn -0.61124     0.4895 -1.2486
## prel:rosn       -0.09397     0.0742 -1.2665
##
## Intercepts:
##          Value   Std. Error t value
```

```
## low|med  -1.4866  1.0214    -1.4555
## med|high  0.5897  1.0189     0.5788
##
## Residual Deviance: 1265.037
## AIC: 1297.037
```

```
# hypothesis tests for statistical significance
Anova(D.mod.3a, test = "LR")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: desired_o
##              LR Chisq Df Pr(>Chisq)
## nrel_f         1.6008  1  0.2057878
## prel           2.1860  1  0.1392746
## negevent       2.0408  1  0.1531332
## posevent       3.4588  1  0.0629154 .
## rosn          13.1261  1  0.0002912 ***
## dayweek_f     19.9268  6  0.0028537 **
## gender_f       5.0941  1  0.0240073 *
## nrel_f:rosn    1.5731  1  0.2097587
## prel:rosn      1.6256  1  0.2023108
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```