Analysis of the paper Nascimento et al.

Diogo B. Provete

2017-03-16

Contents

Loading packages

```r
R.Version()$version.string; R.Version()$platform
```

```
## [1] "R version 3.3.2 (2016-10-31)"
```

```
## [1] "x86_64-apple-darwin13.4.0"
```

```r
library(visdat)
packageVersion("visdat")
```

```
## [1] '0.0.5.9000'
```

```r
library(tidyverse)
```

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
```

```
## Loading tidyverse: dplyr
## Conflicts with tidy packages ----------------------------------------------
## filter(): dplyr, stats
## lag():    dplyr, stats
```

```r
packageVersion("tidyverse")
```

```
## [1] '1.1.1'
```

```r
library(sjPlot)
packageVersion("sjPlot")
```

```
## [1] '2.3.1'
```

```r
library(MuMIn)
packageVersion("MuMIn")
```

```
## [1] '1.15.6'
```

```r
library(bbmle)
```

```
## Loading required package: stats4
```

```
##
## Attaching package: 'bbmle'
```

```
## The following object is masked from 'package:MuMIn':
##
##     AICc
```

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

```r
packageVersion("bbmle")
```

```
## [1] '1.0.18'
```

Data input

```r
lizdata<-read.table("camila_cauda_lagartos.txt", h=TRUE)
head(lizdata);str(lizdata)
```

```
##   Numero    Sex   SVL Intact_tail_length Autotomized_tail_length
## 1      2   Male 20.70                 NA                   12.88
## 2      3   Male 21.10                 NA                   13.07
## 3      6 Female 23.72                 NA                   17.56
## 4      9   Male 18.84              17.38                      NA
## 5     21   Male 22.20                 NA                   16.50
## 6     22   <NA> 20.59                 NA                   12.46
##   Tail_state
## 1          0
## 2          0
## 3          0
## 4          1
## 5          0
## 6          0
```

```
## 'data.frame':    223 obs. of  6 variables:
##  $ Numero              : int  2 3 6 9 21 22 24 25 26 27 ...
##  $ Sex                 : Factor w/ 2 levels "Female","Male": 2 2 1 2 2 NA NA NA NA NA ...
##  $ SVL                 : num  20.7 21.1 23.7 18.8 22.2 ...
##  $ Intact_tail_length  : num  NA NA NA 17.4 NA ...
##  $ Autotomized_tail_length: num  12.9 13.1 17.6 NA 16.5 ...
##  $ Tail_state          : int  0 0 0 1 0 0 1 1 1 1 ...
```

Questions

1. The probability of tail autotomy in the lizard *Coleodactylus meridionalis* increases with body size (SVL)?

- Hypothesis: THe bigger the lizard, the higher the probability of tail autotomy.

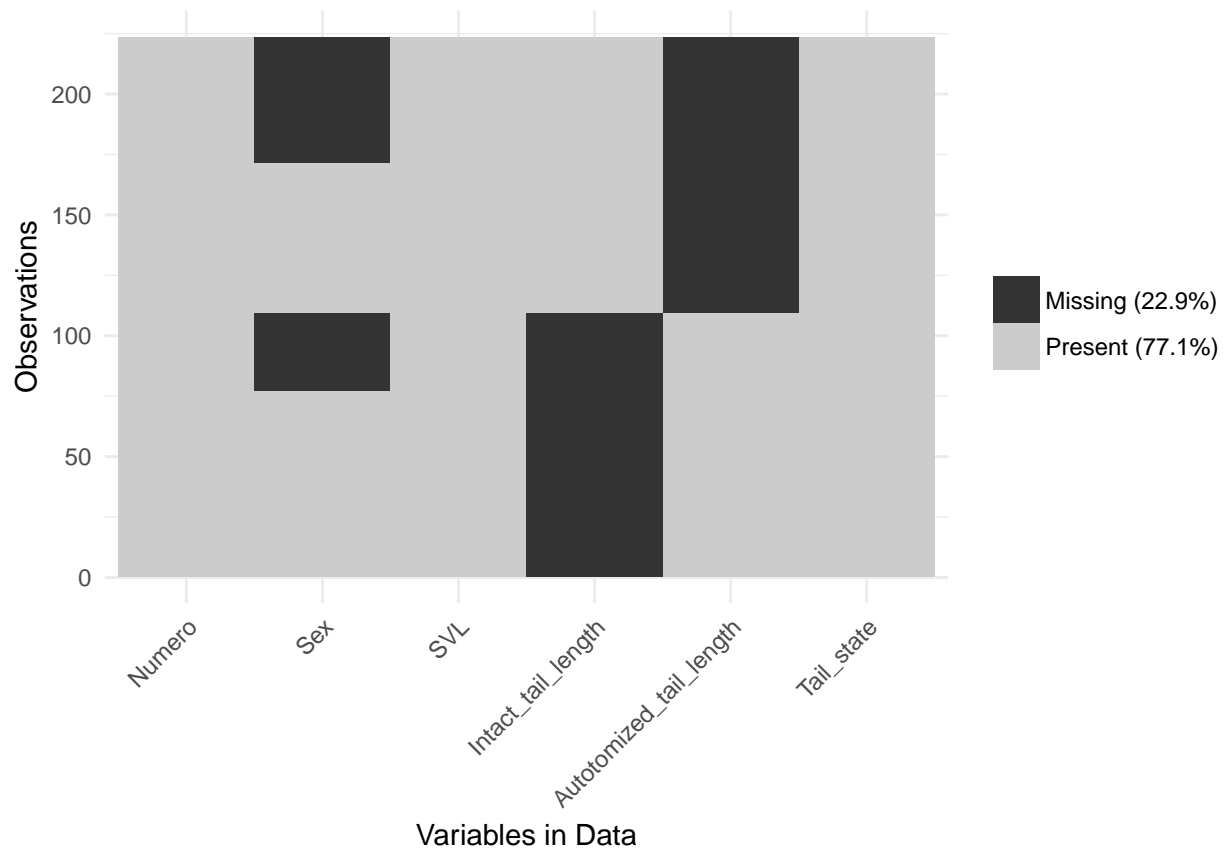2. The size of the autotomized tail is correlated with SVL?

The variable `Tail_state` was coded as 1= intact tail; 0= autotomized tail.

Vizualizing missing data

```
vis_dat(lizdata)
```
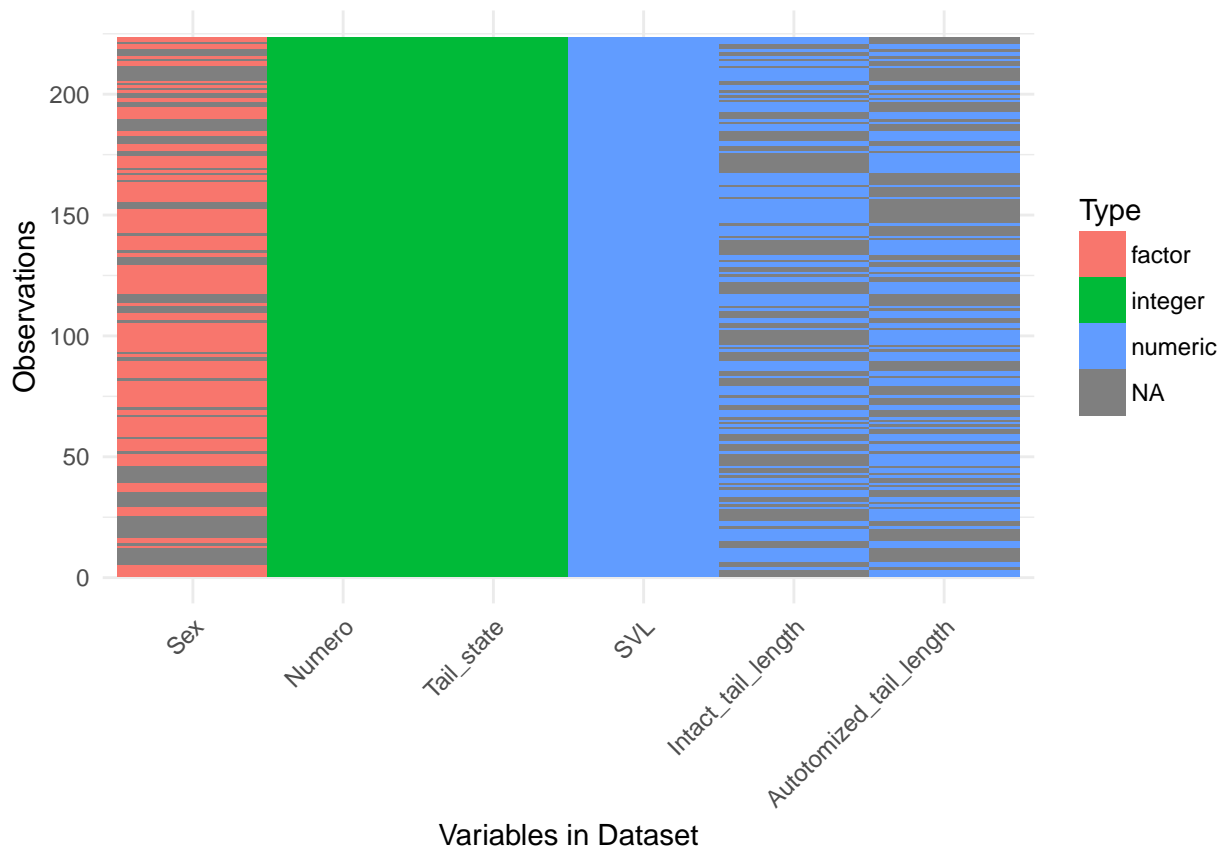


```
vis_miss(lizdata,cluster = TRUE)#22.9% of missing data
```
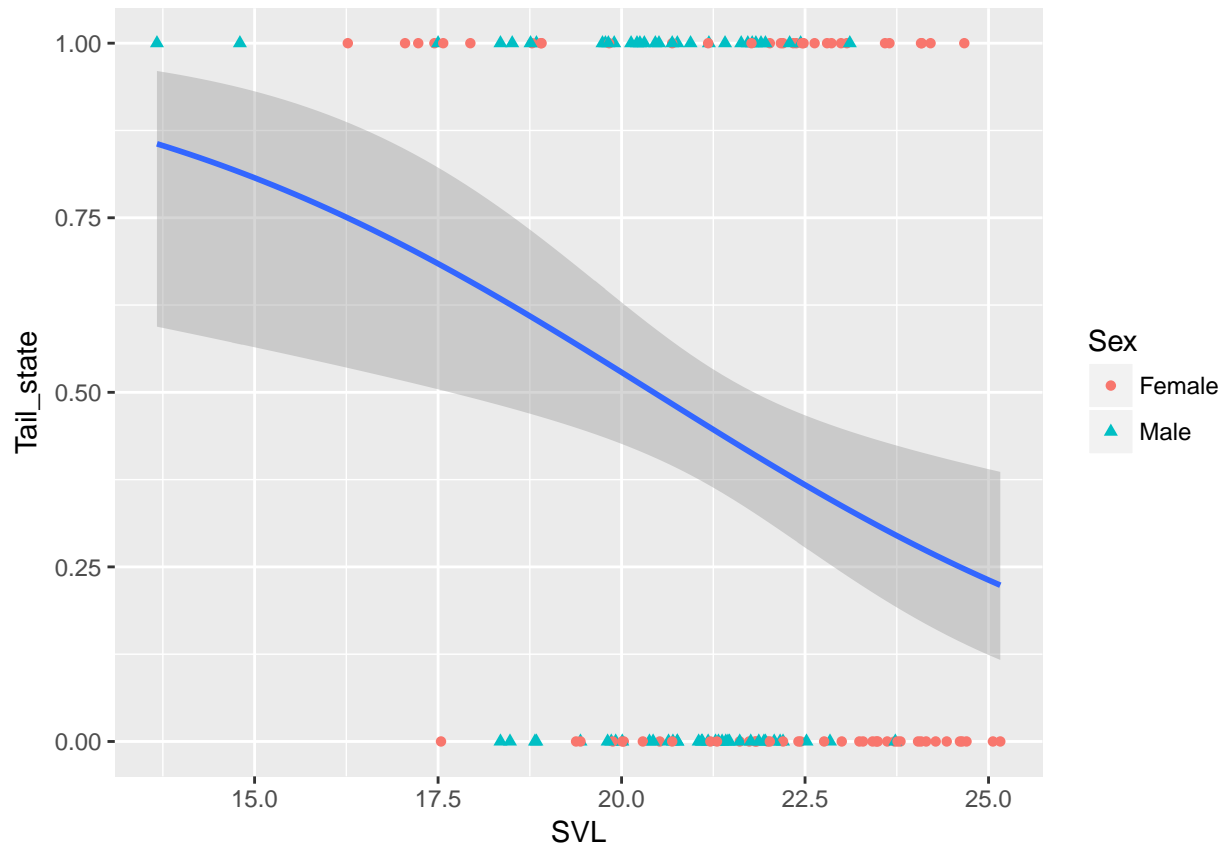
```
dados<-remove_missing(lizdata, vars = "Sex")#excluding rows with missing data for the variable Sex
dim(dados)#check the dim of table after excluding the rows
```

```
## [1] 139    6
```

Vizualizing the data

```
ggplot(dados, aes(SVL, Tail_state))+
    geom_point(aes(shape=Sex, color=Sex))+
    geom_smooth(method = "glm", method.args=list(family="binomial"))
```

It looks like there's a slight influence of SVL. The probability of the tail being intact decreases with increasing SVL. Let's see if the effect is significant.
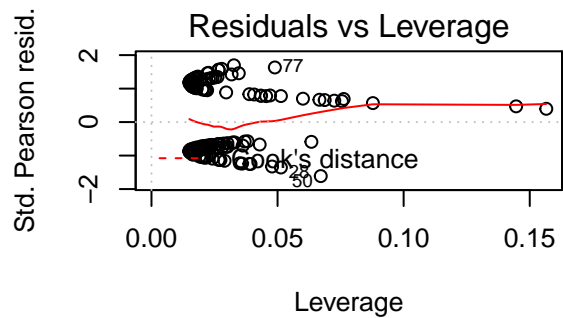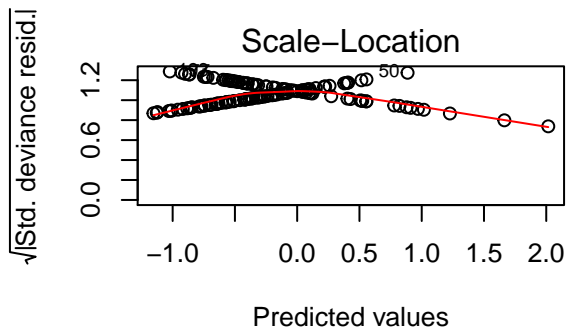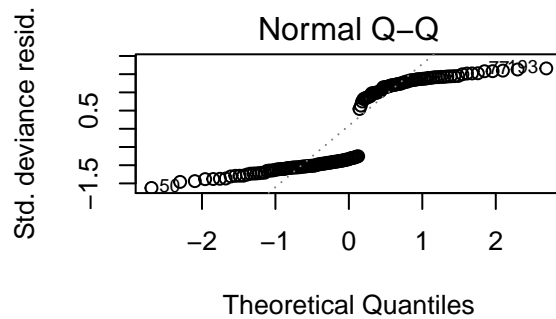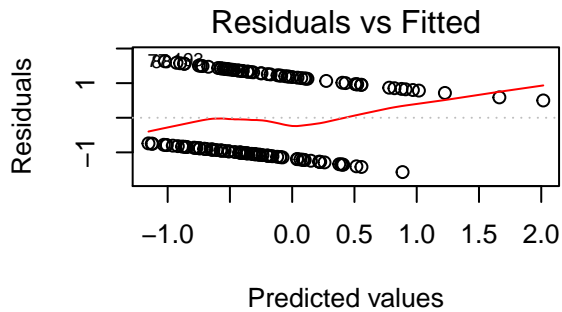
Model for question 1

```
mod_log<-glm(Tail_state~SVL*Sex, data=dados, family = binomial(link="logit"))
mod_pro<-glm(Tail_state~SVL*Sex, data=dados, family = binomial(link="probit"))
AICctab(mod_log, mod_pro, nobs=139)#slight difference between probit and logit. Let's only interpret lo
```

```
##         dAICc df
## mod_pro 0.0   4
## mod_log 0.1   4
```
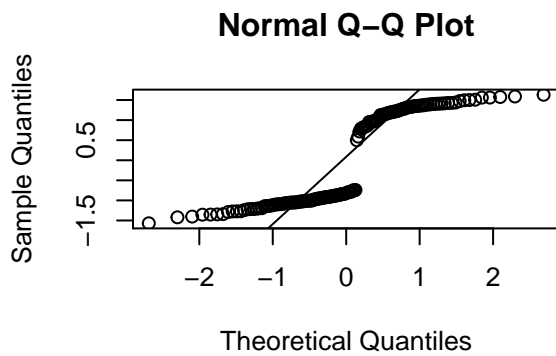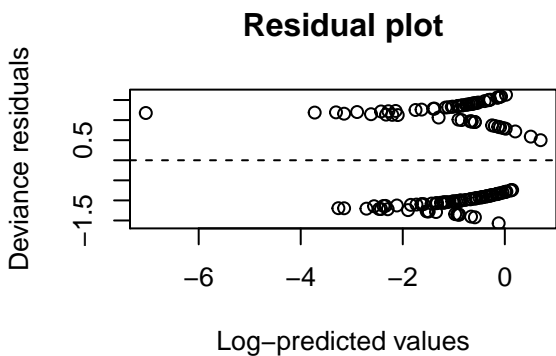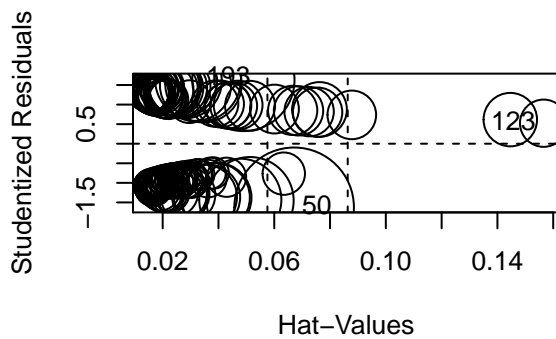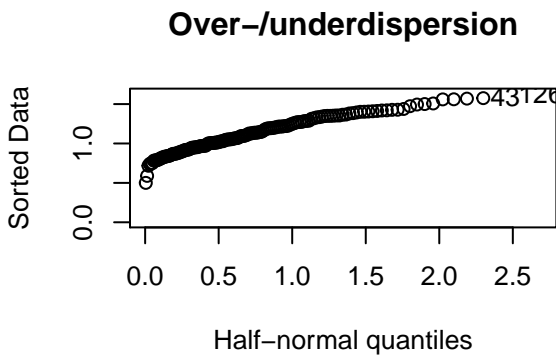
Model diagnosis

```
par(mfrow=c(2,2))
plot(mod_log)
```

Residuals vs Fitted

Normal Q–Q

Scale–Location

Residuals vs Leverage

```
sjp.glm(mod_log, type = "ma")
```

```
## No outliers detected.
```

```
## NULL
```

Over–/underdispersion

Residual plot

Normal Q–Q Plot

Residual plot
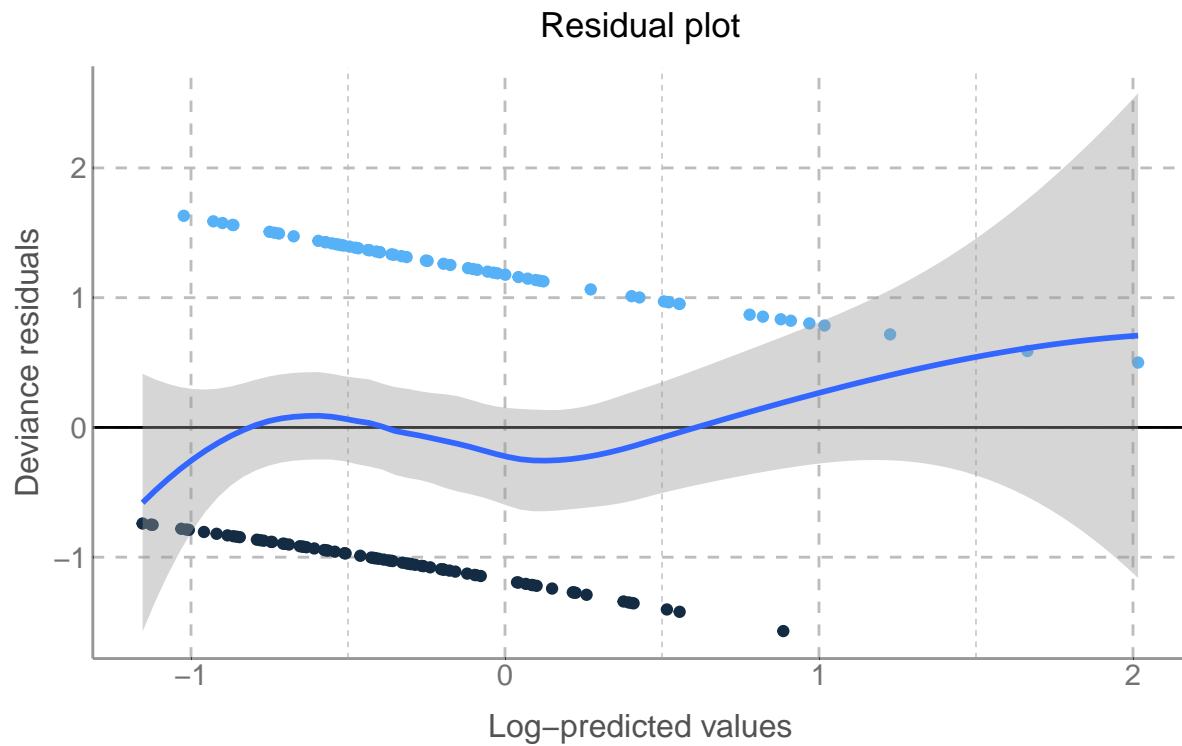
Check significance of terms when they entered the model...

Anova:
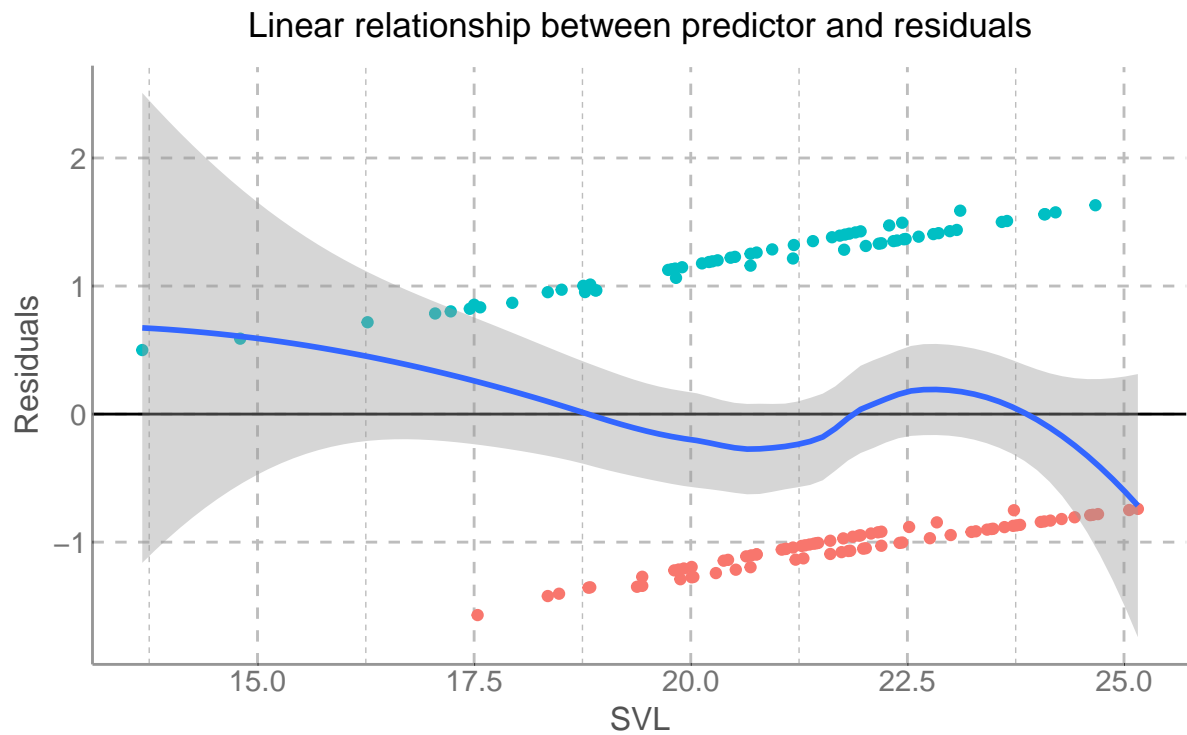
Linear relationship between predictor and residuals

Analysis of Deviance Table
##
Model: binomial, link: logit

```
## 
## Response: Tail_state
## 
## Terms added sequentially (first to last)
## 
## 
##          Df Deviance Resid. Df Resid. Dev Pr(>Chi)
## NULL                      138     191.07
## SVL      1   9.2563       137     181.82 0.002347 **
## Sex      1   0.3920       136     181.43 0.531262
## SVL:Sex  1   0.0454       135     181.38 0.831292
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Inference

```r
sjp.glm(mod_log, type = "pred", vars = "SVL")
```
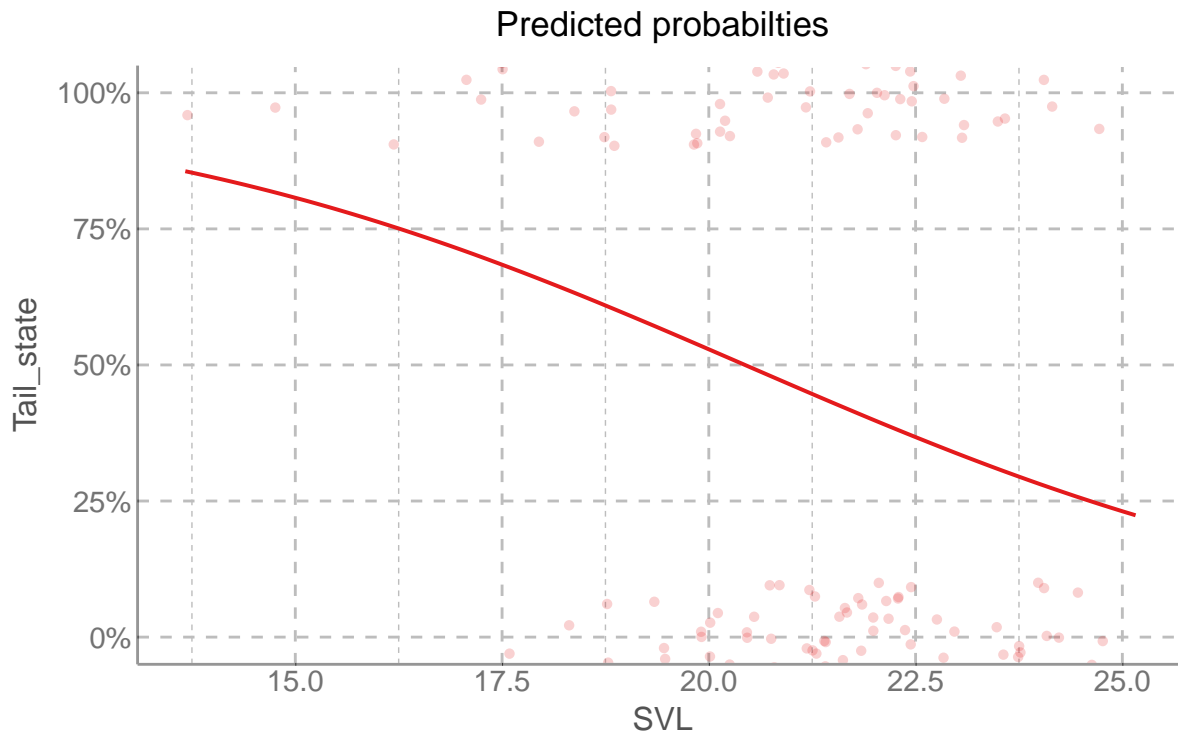

Predicted probabilties

Table with results

```r
sjt.glm(mod_log, show.family = TRUE, string.est="Estimate", show.se = TRUE, show.ci = FALSE, show.dev=TI
```

Tail_state

	Estimate	std. Error	p
(Intercept)	265.98	689.14	0.031
SVL	0.77	0.09	0.023
SexMale	2.01	8.85	0.874
SVL:SexMale	0.96	0.20	0.832
Observations	139		
Deviance	181.379		
X2deviance	p=0.021		
Family	binomial (logit)		

chi-square = A well-fitting model with predictors should significantly differ from the null-model (without predictors), thus, a p-value less than 0.05 indicates a good model-fit.

```r
anova(mod_log, test="Chisq" )
```

```
## Analysis of Deviance Table
##
## Model: binomial, link: logit
##
## Response: Tail_state
##
## Terms added sequentially (first to last)
##
##
##          Df Deviance Resid. Df Resid. Dev Pr(>Chi)
## NULL                      138     191.07
## SVL      1   9.2563       137     181.82 0.002347 **
## Sex      1   0.3920       136     181.43 0.531262
## SVL:Sex  1   0.0454       135     181.38 0.831292
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#For models with known dispersion (e.g., binomial and Poisson fits) the chi-squared test is most approp
round(r.squaredGLMM(mod_log)[[1]], 3)#R2 of the model = 0.081
```

```
## [1] 0.081
```

```r
std.coef(mod_log, partial.sd = TRUE)#standardized coefficients
```

```
##              Estimate* Std. Error*  df
## (Intercept)  0.000000    0.000000 135
## SVL         -0.438858    0.193015 135
## SexMale      0.028284    0.178585 135
## SVL:SexMale -0.037792    0.178271 135
```

Conclusion

The SVL negatively affects (-0.4388583) the probability of the tail being intact, ie., with increasing SVL, the probability of the tail being intact decreases. The interaction wasn't significant (0.8312917), so this effect is **independent** of the lizard sex.

Model for question 2

```r
cauda<-lm(log(Autotomized_tail_length)~SVL, data=dados)
```

Diagnosis

```r
sjp.lm(cauda, type="ma")
```

```
## 4 outliers removed in updated model.
```
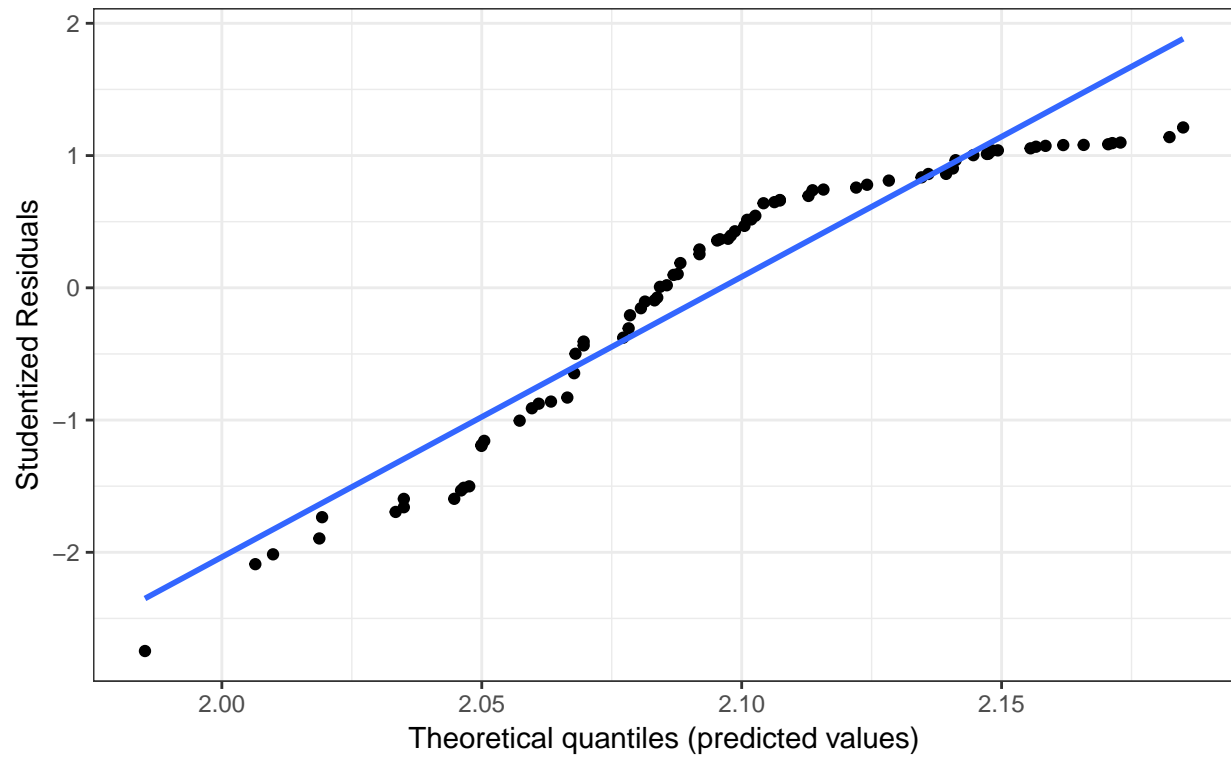
```
## # A tibble: 2 × 3
##     models  adjusted.r2      aic
##      <chr>        <dbl>    <dbl>
```

```
## 1 original -0.008546912 159.6691
## 2  updated -0.008546912 159.6691
```
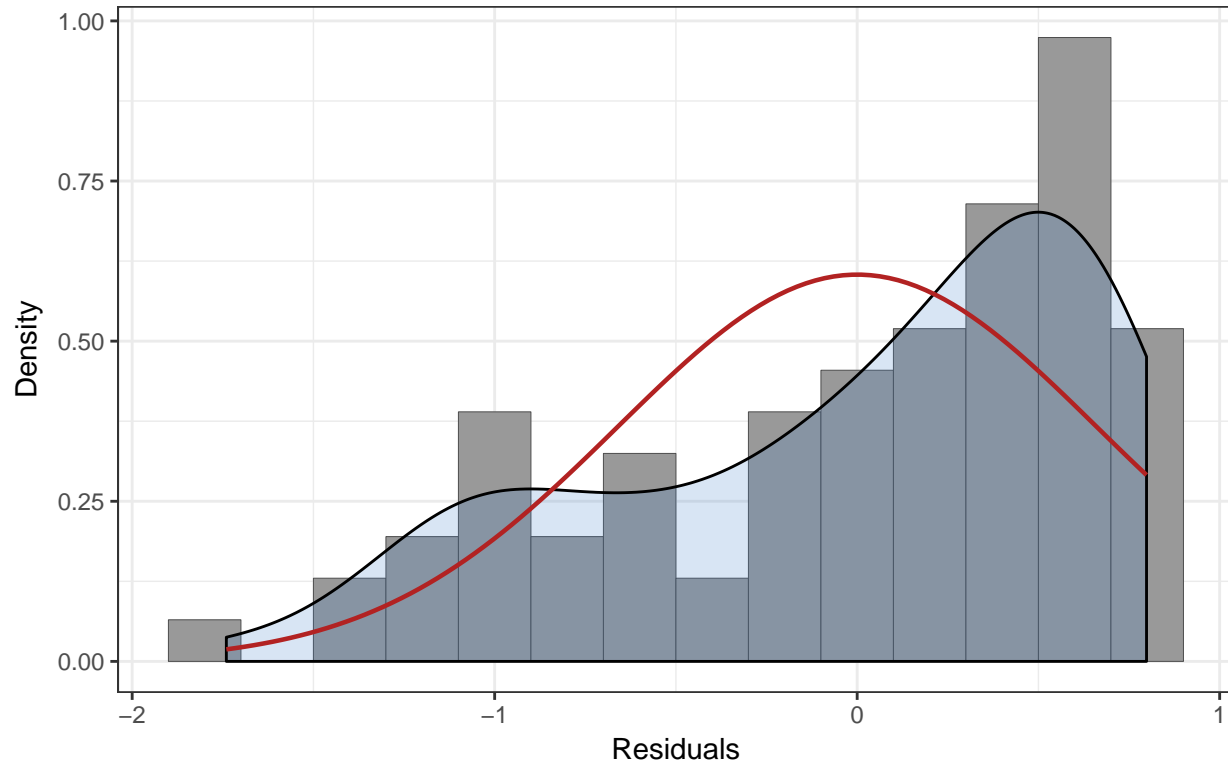
Non−normality of residuals and outliers
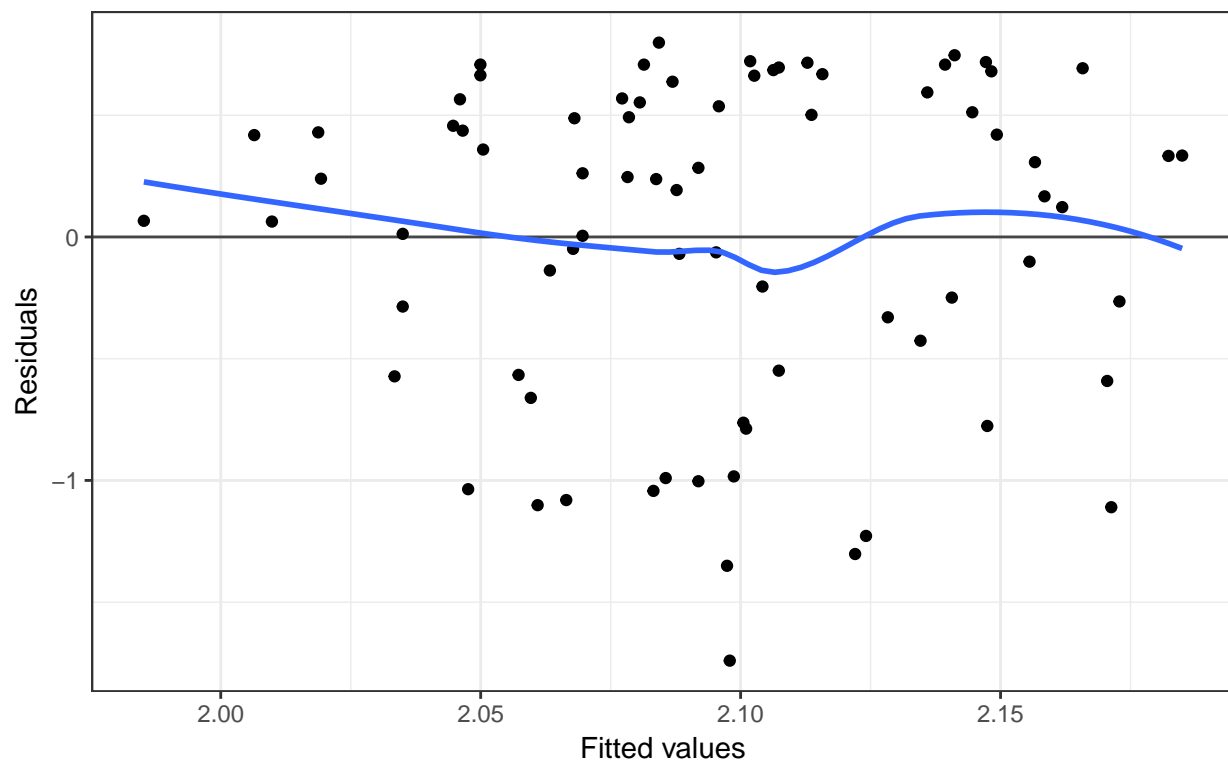
Dots should be plotted along the line

Non−normality of residuals

Distribution should look like normal curve

Homoscedasticity (constant variance of residuals)

Amount and distance of points scattered above/below line is equal or randomly spread

Inference

```
summary(cauda)
```

```
##
## Call:
## lm(formula = log(Autotomized_tail_length) ~ SVL, data = dados)
##
## Residuals:
##      Min      1Q  Median      3Q     Max
## -1.7402 -0.5492  0.2377  0.5527  0.7983
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.52549    0.95903   1.591    0.116
## SVL          0.02621    0.04393   0.597    0.553
##
## Residual standard error: 0.665 on 75 degrees of freedom
##   (62 observations deleted due to missingness)
## Multiple R-squared:  0.004723,   Adjusted R-squared:  -0.008547
## F-statistic: 0.3559 on 1 and 75 DF,  p-value: 0.5526
```

```
anova(cauda)
```

```
## Analysis of Variance Table
##
## Response: log(Autotomized_tail_length)
##           Df Sum Sq Mean Sq F value Pr(>F)
## SVL        1  0.157 0.15742  0.3559 0.5526
## Residuals 75 33.169 0.44226
```

```
sjt.lm(cauda, show.ci = FALSE, show.se = TRUE)
```

log(Autotomized_tail_length)

B

std. Error

p

(Intercept)

1.53

0.96

.116

SVL

0.03

0.04

.553

Observations

77

R2 / adj. R2

.005 / -.009

Conclusion

There's no relationship bewteen SVL and the size of the autotomized tail.