

Universidade Estadual de Campinas
Instituto de Matemática, Estatística e Computação Científica
Departamento de Estatística

Relatório - Parte 2

Exercício 1

Guilherme Pazian RA:160323
Henrique Capatto RA:146406
Hugo Calegari RA:155738
Leonardo Uchoa Pedreira RA:156231

Professor: Caio Lucidius Naberezny Azevedo

Campinas-SP, 20 de Junho de 2017

Parte descritiva

De maneira geral, pode-se observar, pelas tabelas 1, 2, 3 e 4, as seguintes características:

- A quantidade de indivíduos que tem pelo menos uma infecção, para o grupo cujo hábito de nadar é frequente, é menor do que para o grupo que nada ocasionalmente;
- A quantidade de indivíduos que tem pelo menos uma infecção, para o grupo cujo local onde costuma nadar é a praia, é menor do que para o grupo que não costuma nadar na praia;
- A quantidade de indivíduos que tem pelo menos uma infecção, para a faixa etária 15-19, é maior do que para os grupos das demais faixas; e o número de indivíduos que tem pelo menos uma infecção na faixa etária 20-24 é maior do que a faixa etária 25-29;
- A quantidade de indivíduos que tem pelo menos uma infecção, para o sexo masculino, é maior do que para o para o sexo feminino;

Avaliar o comportamento do número de infecções para as diferentes covariáveis e suas possíveis interações, é de fundamental importância para a futura modelagem. A seguir, são apresentados os gráficos de perfis para verificar as possíveis interações entre fatores (em geral a combinação de dois fatores). Por estes gráficos, é aceitável considerar as interações entre a faixa etária e as demais covariáveis: hábito de nadar, local onde nada e o sexo; interações entre sexo e as covariáveis: local onde nada e o hábito de nadar e a interação entre o local onde costuma nadar e o hábito de nadar.

Hábito de nadar	Contagem do número de infecções
Frequente	140
Ocasional	258

Tabela 1: Tabela com as quantidades do número de infecções para os diferentes níveis de hábito de nadar.

Local	Contagem do número de infecções
Praia	155
Piscina	243

Tabela 2: Tabela com as quantidades do número de infecções para os diferentes níveis do lugar onde se costuma nadar.

Faixa etária	Contagem do número de infecções
15-19	223
20-24	92
25-29	83

Tabela 3: Tabela com as quantidades do número de infecções para os diferentes níveis de faixa etária.

Sexo	Contagem do número de infecções
Feminino	131
Masculino	267

Tabela 4: Tabela com as quantidades do número de infecções para os sexos.

Hábito de nadar	Local	Contagem do número de infecções
Frequente	Praia	59
Frequente	Piscina	81
Ocasional	Praia	96
Ocasional	Piscina	162

Tabela 5: Tabela com as quantidades do número de infecções para as combinações dos diferentes hábitos de nadar e local onde se costuma nadar.

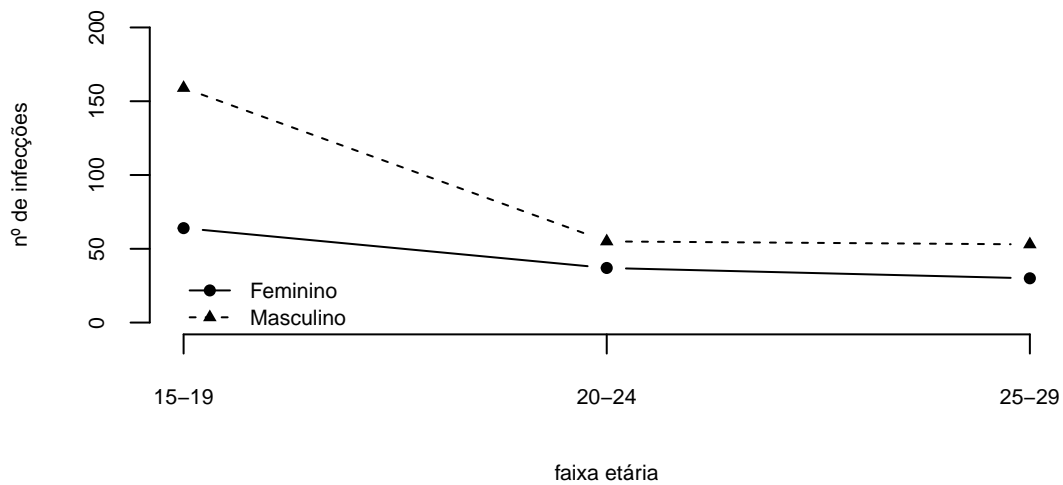


Figura 1: Gráfico de perfil que representa o número de infecções de ouvido para cada sexo. Nota-se que ao mudar de faixa etária há diminuição do número de infecções para ambos os sexos. Observa-se que a redução é maior ao mudar da faixa etária de 15-19 para 20-24 para o sexo masculino. Pelo gráfico, pode-se pensar que existe interação entre as covariáveis faixa etária e sexo.

Hábito de nadar	Sexo	Contagem do número de infecções
Frequente	Feminino	47
Frequente	Masculino	93
Ocasional	Feminino	84
Ocasional	Masculino	174

Tabela 6: Tabela com as quantidades do número de infecções para as combinações dos diferentes hábitos de nadar e os sexos.

Hábito de nadar	Faixa etária	Contagem do número de infecções
Frequente	15-19	87
Frequente	20-24	33
Frequente	25-29	20
Ocasional	15-19	136
Ocasional	20-24	59
Ocasional	25-29	63

Tabela 7: Tabela com as quantidades do número de infecções para as combinações dos diferentes hábitos de nadar e as diferentes faixas etárias.

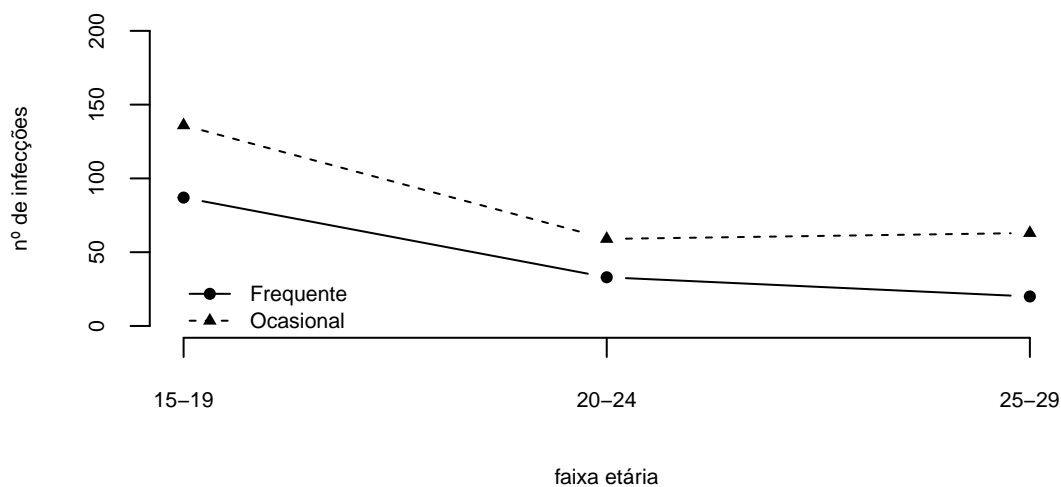


Figura 2: Gráfico de perfil que representa o número de infecções para os grupos que têm hábitos (frequente e ocasional) de nadar diferentes. Para estes diferentes grupos de hábitos de nadar, nota-se que o número de infecções diminui. No entanto, para aqueles cujo hábito é ocasional, a quantidade de infecções é superior para os diferentes níveis de faixa etária. Cogita-se que possa existir interação entre o hábito de nadar e a faixa etária.

Local	Faixa etária	Contagem do número de infecções
Praia	15-19	90
Praia	20-24	20
Praia	25-29	45
Piscina	15-19	133
Piscina	20-24	72
Piscina	25-29	38

Tabela 8: Tabela com as quantidades do número de infecções para as combinações dos diferentes locais onde se costuma nadar e as diferentes faixas etárias.

Local	Sexo	Contagem do número de infecções
Praia	Feminino	81
Praia	Masculino	74
Piscina	Feminino	50
Piscina	Masculino	193

Tabela 9: Tabela com as quantidades do número de infecções para as combinações dos diferentes locais onde se costuma nadar e os sexos.

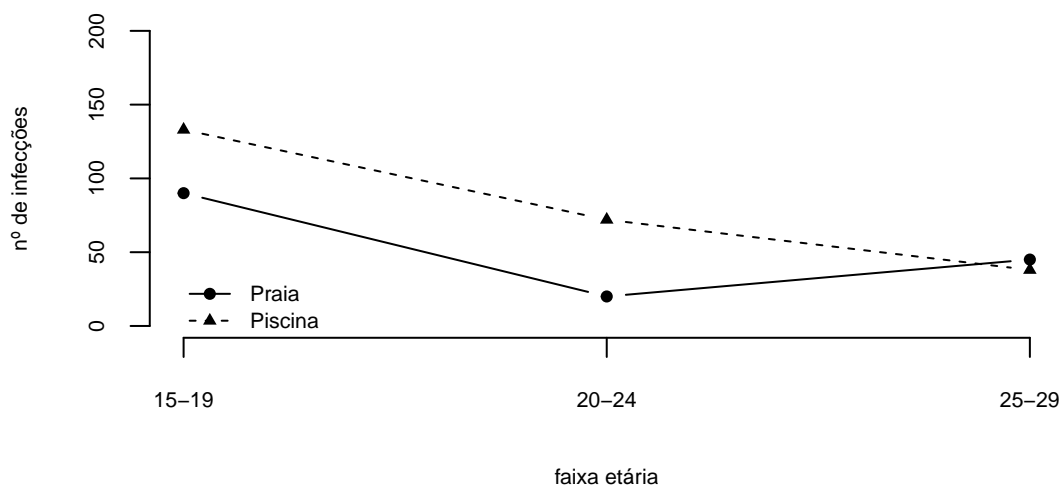


Figura 3: Gráfico de perfil que representa o número de infecções para os indivíduos que costumam nadar na praia ou na piscina. Com exceção da última faixa etária (25-29), o número de infecções para indivíduos que costumam nadar em piscinas é maior para as diferentes faixas etárias, de 15-19 e 20-24. Observa-se que para os diferentes lugares onde se costuma nadar, nas faixas etárias 15-19 e 20-24, as retas possuem um pequeno desnível que as tornam não paralelas. No entanto, para a última faixa etária, nota-se que o comportamento é diferente para os diferentes locais onde se costuma nadar. Assim, é possível pensar em interação entre a faixa etária e o local onde se costuma nadar.

Sexo	Faixa etária	Contagem do número de infecções
Feminino	15-19	64
Feminino	20-24	37
Feminino	25-29	30
Masculino	15-19	159
Masculino	20-24	55
Masculino	25-29	53

Tabela 10: Tabela com as quantidades do número de infecções para as combinações dos sexos as diferentes faixas etárias.

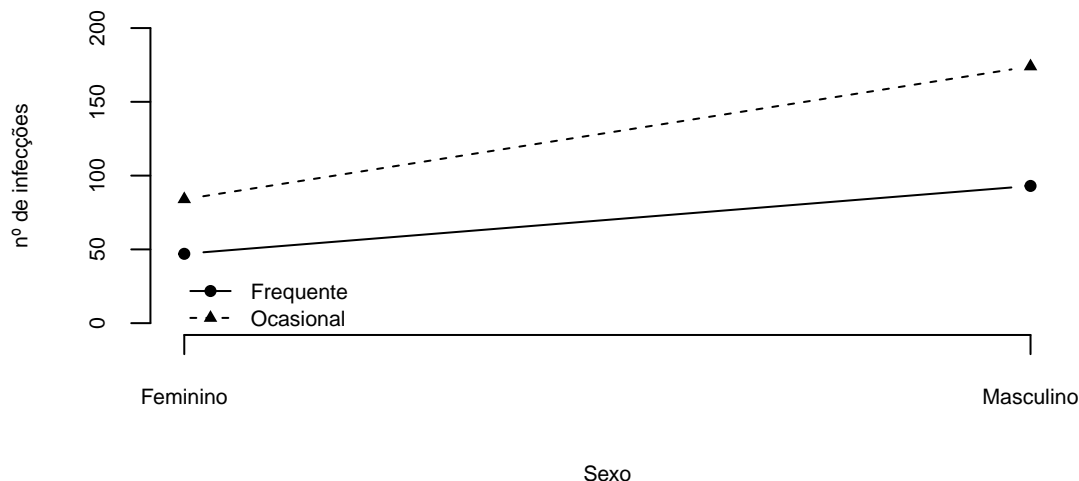


Figura 4: Gráfico de perfil que representa o número de infecções para os indivíduos que tem um determinado hábito de nadar, para os diferentes sexos. Observa-se que as retas não são paralelas, com isso pode-se pensar que existe interação entre o sexo e o hábito de nadar.

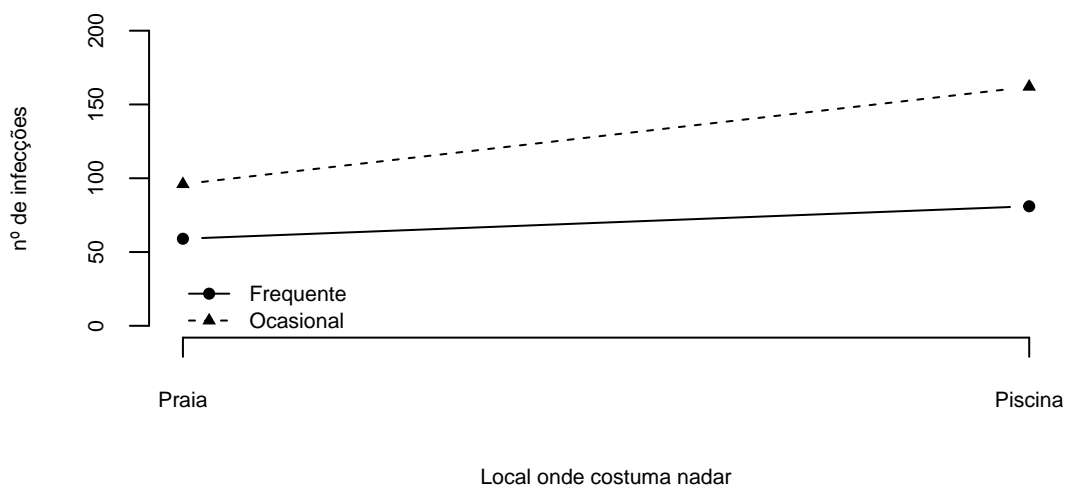


Figura 5: Gráfico de perfil que representa o número de infecções para os indivíduos que tem determinado hábito de nadar e o local que costumam nadar. O comportamento é muito similar em relação à figura 4 e cogita-se na possível interação entre local onde costuma nadar e hábito de nadar.

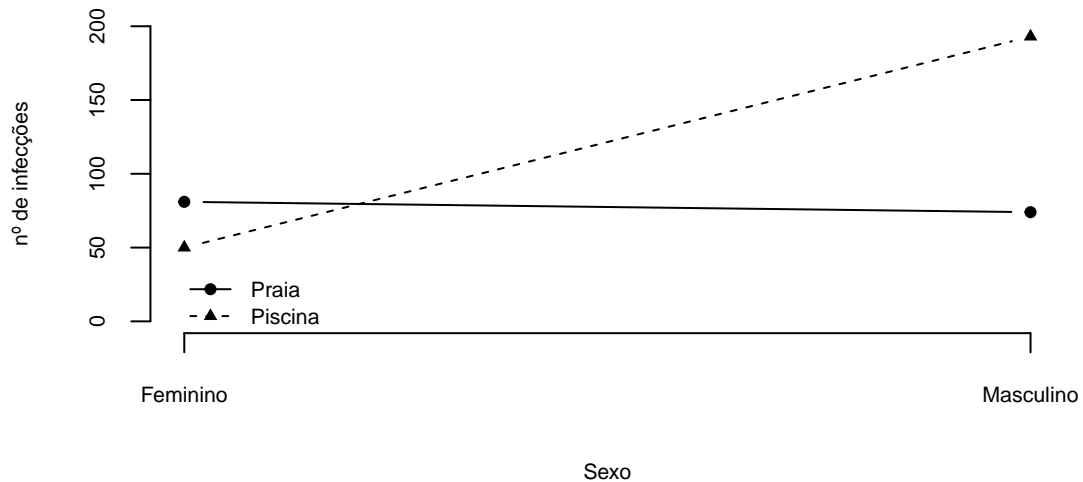


Figura 6: Gráfico de perfil que representa o número de infecções para os sexos, para indivíduos que tem costume de nadar em diferentes lugares. Nota-se que é plausível a interação entre sexo e o lugar onde costuma nadar.

Análise inferencial e ajuste de modelos

O modelo “saturado” considerado, foi aquele no qual os fatores principais e todas as possíveis interações foram consideradas, isto é, a interação de terceira ordem foi considerada (combinação de todas as covariáveis em estudo).

Pela análise descritiva, as interações de primeira ordem (combinação de dois fatores) serão consideradas para elaboração do primeiro modelo.

```
##
## Call:
## glm(formula = dados$`Nº de infecções` ~ dados$`Hab. de nadar` *
##      dados$Local * dados$`Faixa etária` * dados$Sexo, family = poisson(link = "log"),
##      data = dados)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2,6458  -1,6115  -0,8165   0,5154   5,9978
##
## Coefficients:
##                                     Estimate
## (Intercept)                        0,35667
## dados$`Hab. de nadar`Occas          0,11333
## dados$LocalNonBeach                 0,04879
## dados$`Faixa etária`20-24          -1,04982
## dados$`Faixa etária`25-29          -0,35667
## dados$SexoMale                     -0,57982
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach 0,73397
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`20-24 0,32850
```

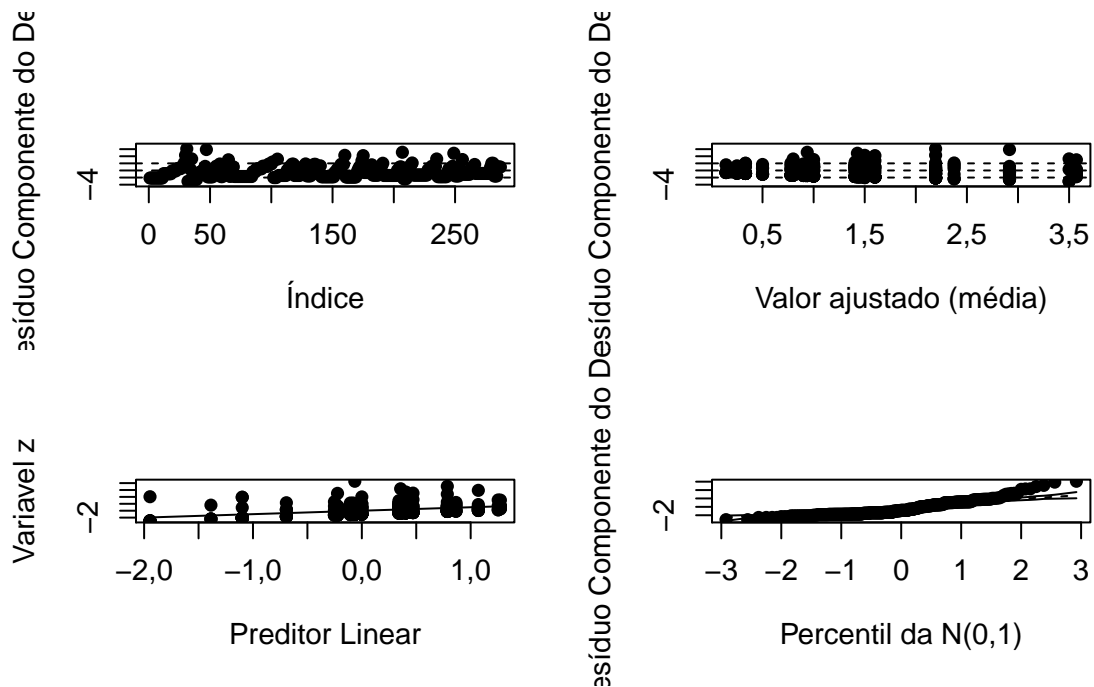
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`25-29	0,75167
## dados\$LocalNonBeach:dados\$`Faixa etária`20-24	0,53900
## dados\$LocalNonBeach:dados\$`Faixa etária`25-29	-1,43508
## dados\$`Hab. de nadar`Occas:dados\$SexoMale	0,51528
## dados\$LocalNonBeach:dados\$SexoMale	0,51528
## dados\$`Faixa etária`20-24:dados\$SexoMale	1,27297
## dados\$`Faixa etária`25-29:dados\$SexoMale	-0,51879
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`20-24	-0,69575
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`25-29	-0,21267
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$SexoMale	-0,91798
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`20-24:dados\$SexoMale	-2,90302
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`25-29:dados\$SexoMale	-0,34620
## dados\$LocalNonBeach:dados\$`Faixa etária`20-24:dados\$SexoMale	-1,18311
## dados\$LocalNonBeach:dados\$`Faixa etária`25-29:dados\$SexoMale	1,96963
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`20-24:dados\$SexoMale	3,97616
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`25-29:dados\$SexoMale	0,63558
##	Std. Error
## (Intercept)	0,22361
## dados\$`Hab. de nadar`Occas	0,30276
## dados\$LocalNonBeach	0,46547
## dados\$`Faixa etária`20-24	0,50000
## dados\$`Faixa etária`25-29	0,46547
## dados\$SexoMale	0,33541
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach	0,57425
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`20-24	0,65918
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`25-29	0,55764
## dados\$LocalNonBeach:dados\$`Faixa etária`20-24	0,72648
## dados\$LocalNonBeach:dados\$`Faixa etária`25-29	1,17615
## dados\$`Hab. de nadar`Occas:dados\$SexoMale	0,43301
## dados\$LocalNonBeach:dados\$SexoMale	0,54899
## dados\$`Faixa etária`20-24:dados\$SexoMale	0,67480
## dados\$`Faixa etária`25-29:dados\$SexoMale	0,69222
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`20-24	0,91991
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`25-29	1,34129
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$SexoMale	0,68009
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`20-24:dados\$SexoMale	1,29353
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`25-29:dados\$SexoMale	0,82065
## dados\$LocalNonBeach:dados\$`Faixa etária`20-24:dados\$SexoMale	0,91580
## dados\$LocalNonBeach:dados\$`Faixa etária`25-29:dados\$SexoMale	1,33905
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`20-24:dados\$SexoMale	1,49457
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`25-29:dados\$SexoMale	1,53736
##	z value
## (Intercept)	1,595

## dados\$`Hab. de nadar`Occas	0,374
## dados\$LocalNonBeach	0,105
## dados\$`Faixa etária`20-24	-2,100
## dados\$`Faixa etária`25-29	-0,766
## dados\$SexoMale	-1,729
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach	1,278
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`20-24	0,498
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`25-29	1,348
## dados\$LocalNonBeach:dados\$`Faixa etária`20-24	0,742
## dados\$LocalNonBeach:dados\$`Faixa etária`25-29	-1,220
## dados\$`Hab. de nadar`Occas:dados\$SexoMale	1,190
## dados\$LocalNonBeach:dados\$SexoMale	0,939
## dados\$`Faixa etária`20-24:dados\$SexoMale	1,886
## dados\$`Faixa etária`25-29:dados\$SexoMale	-0,749
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`20-24	-0,756
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`25-29	-0,159
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$SexoMale	-1,350
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`20-24:dados\$SexoMale	-2,244
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`25-29:dados\$SexoMale	-0,422
## dados\$LocalNonBeach:dados\$`Faixa etária`20-24:dados\$SexoMale	-1,292
## dados\$LocalNonBeach:dados\$`Faixa etária`25-29:dados\$SexoMale	1,471
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`20-24:dados\$SexoMale	2,660
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`25-29:dados\$SexoMale	0,413
##	Pr(> z)
## (Intercept)	0,1107
## dados\$`Hab. de nadar`Occas	0,7082
## dados\$LocalNonBeach	0,9165
## dados\$`Faixa etária`20-24	0,0358
## dados\$`Faixa etária`25-29	0,4435
## dados\$SexoMale	0,0839
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach	0,2012
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`20-24	0,6182
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`25-29	0,1777
## dados\$LocalNonBeach:dados\$`Faixa etária`20-24	0,4581
## dados\$LocalNonBeach:dados\$`Faixa etária`25-29	0,2224
## dados\$`Hab. de nadar`Occas:dados\$SexoMale	0,2341
## dados\$LocalNonBeach:dados\$SexoMale	0,3479
## dados\$`Faixa etária`20-24:dados\$SexoMale	0,0592
## dados\$`Faixa etária`25-29:dados\$SexoMale	0,4536
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`20-24	0,4495
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$`Faixa etária`25-29	0,8740
## dados\$`Hab. de nadar`Occas:dados\$LocalNonBeach:dados\$SexoMale	0,1771
## dados\$`Hab. de nadar`Occas:dados\$`Faixa etária`20-24:dados\$SexoMale	0,0248

```

## dados$`Hab. de nadar`Occas:dados$`Faixa etária`25-29:dados$SexoMale 0,6731
## dados$LocalNonBeach:dados$`Faixa etária`20-24:dados$SexoMale 0,1964
## dados$LocalNonBeach:dados$`Faixa etária`25-29:dados$SexoMale 0,1413
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach:dados$`Faixa etária`20-24:dados$SexoMale 0,0078
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach:dados$`Faixa etária`25-29:dados$SexoMale 0,6793
##
## (Intercept)
## dados$`Hab. de nadar`Occas
## dados$LocalNonBeach
## dados$`Faixa etária`20-24 *
## dados$`Faixa etária`25-29
## dados$SexoMale .
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`20-24
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`25-29
## dados$LocalNonBeach:dados$`Faixa etária`20-24
## dados$LocalNonBeach:dados$`Faixa etária`25-29
## dados$`Hab. de nadar`Occas:dados$SexoMale
## dados$LocalNonBeach:dados$SexoMale
## dados$`Faixa etária`20-24:dados$SexoMale .
## dados$`Faixa etária`25-29:dados$SexoMale
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach:dados$`Faixa etária`20-24
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach:dados$`Faixa etária`25-29
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach:dados$SexoMale
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`20-24:dados$SexoMale *
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`25-29:dados$SexoMale
## dados$LocalNonBeach:dados$`Faixa etária`20-24:dados$SexoMale
## dados$LocalNonBeach:dados$`Faixa etária`25-29:dados$SexoMale
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach:dados$`Faixa etária`20-24:dados$SexoMale **
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach:dados$`Faixa etária`25-29:dados$SexoMale
## ---
## Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 824,51 on 286 degrees of freedom
## Residual deviance: 703,72 on 263 degrees of freedom
## AIC: 1124,1
##
## Number of Fisher Scoring iterations: 6

```



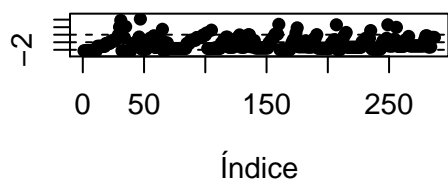
```
##
## Call:
## glm(formula = dados$`Nº de infecções` ~ dados$`Hab. de nadar` *
##   dados$Local + dados$`Hab. de nadar` * dados$`Faixa etária` +
##   dados$`Hab. de nadar` * dados$Sexo + dados$Local * dados$`Faixa etária` +
##   dados$Local * dados$Sexo + dados$`Faixa etária` * dados$Sexo,
##   family = poisson(link = "log"), data = dados)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2,2669  -1,5607  -1,0478   0,5761   5,9484
##
## Coefficients:
##
##                                Estimate Std. Error
## (Intercept)                   0,43933    0,18580
## dados$`Hab. de nadar`Occas      0,20266    0,22423
## dados$LocalNonBeach            -0,21031    0,28149
## dados$`Faixa etária`20-24      -1,03905    0,31257
## dados$`Faixa etária`25-29      -0,77279    0,32064
## dados$SexoMale                 -0,58197    0,23660
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach  0,35032    0,23809
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`20-24 0,10498    0,26459
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`25-29 0,70910    0,29297
## dados$`Hab. de nadar`Occas:dados$SexoMale      0,05472    0,24353
## dados$LocalNonBeach:dados$`Faixa etária`20-24  0,57879    0,31209
## dados$LocalNonBeach:dados$`Faixa etária`25-29  0,08302    0,27542
```

```

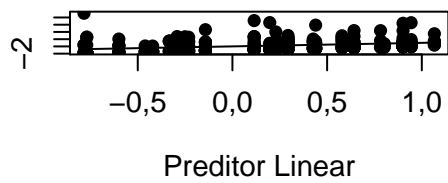
## dados$LocalNonBeach:dados$SexoMale          0,64622    0,24875
## dados$`Faixa etária`20-24:dados$SexoMale      0,39797    0,29546
## dados$`Faixa etária`25-29:dados$SexoMale      0,14612    0,28955
##                                               z value Pr(>|z|)
## (Intercept)                                2,365 0,018054 *
## dados$`Hab. de nadar`Occas                    0,904 0,366110
## dados$LocalNonBeach                         -0,747 0,454981
## dados$`Faixa etária`20-24                   -3,324 0,000887 ***
## dados$`Faixa etária`25-29                   -2,410 0,015947 *
## dados$SexoMale                              -2,460 0,013905 *
## dados$`Hab. de nadar`Occas:dados$LocalNonBeach 1,471 0,141184
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`20-24 0,397 0,691554
## dados$`Hab. de nadar`Occas:dados$`Faixa etária`25-29 2,420 0,015503 *
## dados$`Hab. de nadar`Occas:dados$SexoMale      0,225 0,822220
## dados$LocalNonBeach:dados$`Faixa etária`20-24 1,855 0,063659 .
## dados$LocalNonBeach:dados$`Faixa etária`25-29 0,301 0,763100
## dados$LocalNonBeach:dados$SexoMale            2,598 0,009379 **
## dados$`Faixa etária`20-24:dados$SexoMale      1,347 0,178003
## dados$`Faixa etária`25-29:dados$SexoMale      0,505 0,613810
## ---
## Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 824,51 on 286 degrees of freedom
## Residual deviance: 730,31 on 272 degrees of freedom
## AIC: 1132,7
##
## Number of Fisher Scoring iterations: 6

```

Resíduo Componente do De



Variável z



Resíduo Componente do Desídeo Componente do De

