

# Relatório 2

Luca Klein

2022-12-15

## 1. Modelos de Resposta Binária

Quando tratamos de regressões lineares tendemos a incluir variáveis cuja distribuição é contínua, contudo, existe a modelagem linear para variáveis discretas. Tal abordagem para variáveis numeráveis pode apresentar o comportamento binário, em que sua interpretação se pauta em probabilidades. No entanto, temos que a interpretação dos coeficientes e dos testes de inferência não se dão da mesma maneira, bem como a ocorrência de problemas de heterocedasticidade e probabilidades fora do intervalo entre 0 e 1. Nesse sentido, uma alternativa para modelagem desse tipo de informação se dá via modelos logito e probito. A estimação de tais métodos resulta em respostas não lineares nos parâmetros, sendo assim um dos caminhos que contornam tal característica é por meio de processos de máxima verossimilhança. Ademais, dado a dificuldade de interpretação dos coeficientes estimados nessas classes de modelos, muitas vezes cabe interpretar o sinal do parâmetro para que seja possível inferir se aquele evento analisado tem maior ou menor chance de ocorrer, dado uma variável independente.

### 1.1 Logit

A classe de modelos logísticos, como já comentado, se pauta em investigar a relação probabilística de uma variável ordinal ou binária. Nessa linha, a partir de uma variável categórica binária, nosso vetor dependente ( $Y$ ), temos que a combinação de regressores que maximiza a probabilidade do evento  $Y$  acontecer, dado pela máxima verossimilhança, irá resultar na curva logística. Posteriormente, como os coeficientes estarão no formato logarítmico basta exponencializarmos os estimadores para auferir a razão de chances e, assim, chegamos nas probabilidades ligadas a ocorrência do evento  $Y$  dado o comportamento dos regressores.

### 1.2 Probit

Os modelos probito tem uma aplicação muito similar aos de logito, em que a diferença entre eles se pauta na suposição da distribuição dos erros e na interpretação dos resultados. Então, o Probit é utilizado quando os erros têm dispersão gaussiana e o Logit quando a dispersão é logística. Além disso, as estimações dos coeficientes se dão da mesma maneira, ou seja, a partir da maximização da função da máxima verossimilhança, contudo, dado a forma funcional dessa classe de modelos, a interpretação dos parâmetros é feita via efeito marginal ou sinal do estimador.

### 1.3 Aplicação

Para o presente trabalho foi selecionado dois períodos e duas geografias a partir dos microdados da PNAD. A escolha foi pautada para o 2T2012 e 2T2022 para as Unidades Federativas (UFs) de maior e menor taxa de desemprego no segundo recorte temporal. A ideia por trás do trabalho é avaliar a evolução da desocupação mediante algumas categorias que possam sugerir maiores probabilidades do evento desemprego acontecer. Nessa linha, iremos avaliar Santa Catarina e a Bahia, cujas taxas de desocupação no 2T2022 foram 3,9% e 15,5%, respectivamente, em conjunto a características de raça, sexo e localização urbana ou rural. Dito isso, os modelos irão contemplar como variável explicativa um vetor binário de desemprego, sendo 1 pessoas ocupadas e 0 pessoas desocupadas, ao passo que as variáveis explicativas selecionadas para raça considerou

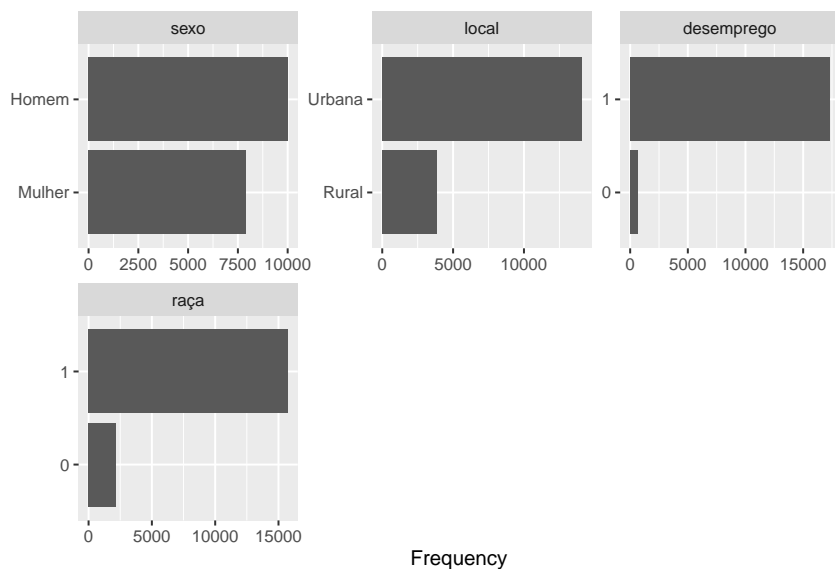
brancos sendo 1 e não brancos 0 e para as demais mantivemos as categorias masculino ou feminino e urbano ou rural.

### 1.31 Análise Descritiva

Nesta seção será apresentado as informações descritivas e o comportamento das variáveis para as UFs e períodos.

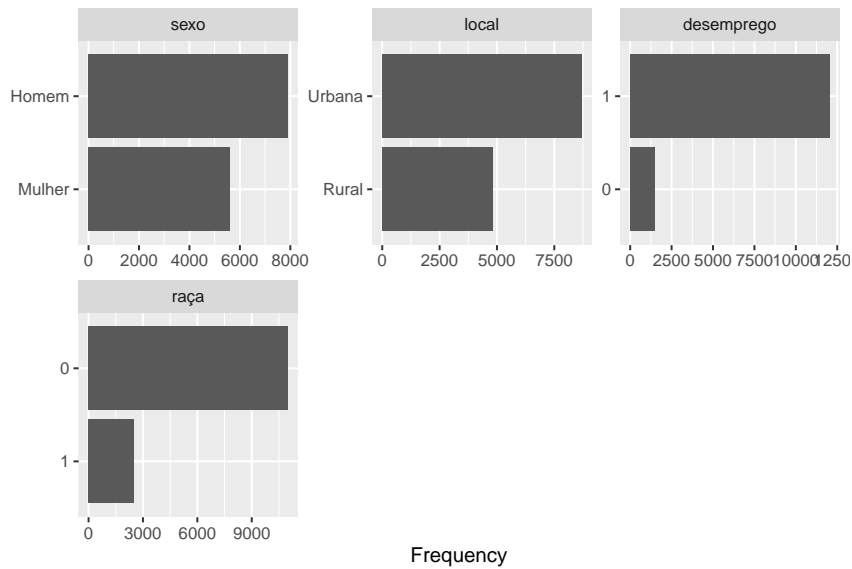
#### Santa Catarina 2T2012

```
##      desemprego      sexo      raça      local
##  Min.      :0.0000  Homem :10007  Min.      :0.0000  Urbana:14058
##  1st Qu.:1.0000  Mulher: 7907  1st Qu.:1.0000  Rural : 3856
##  Median :1.0000
##  Mean   :0.9644
##  3rd Qu.:1.0000
##  Max.   :1.0000
##              Mean   :0.8784
##              3rd Qu.:1.0000
##              Max.   :1.0000
```



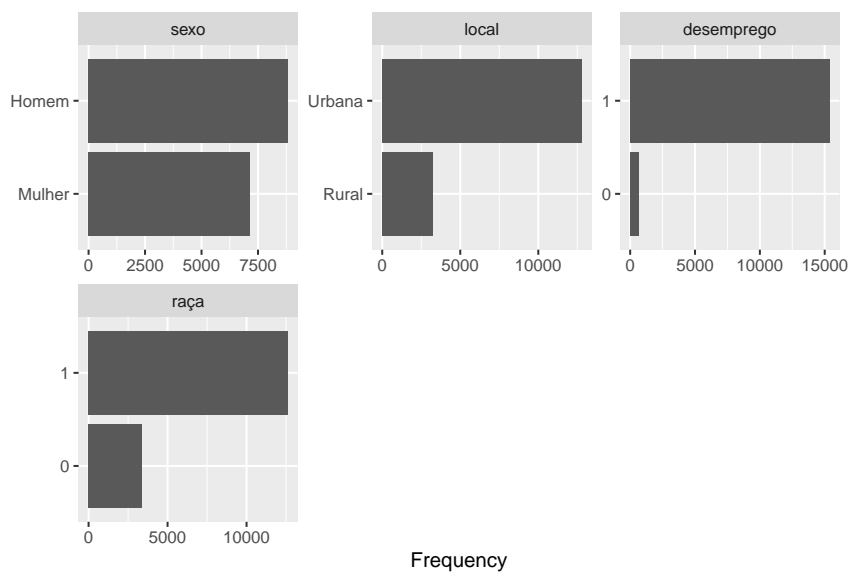
#### Bahia 2T2012

```
##      desemprego      sexo      raça      local
##  Min.      :0.0000  Homem :7900  Min.      :0.0000  Urbana:8684
##  1st Qu.:1.0000  Mulher:5615  1st Qu.:0.0000  Rural :4831
##  Median :1.0000
##  Mean   :0.8899
##  3rd Qu.:1.0000
##  Max.   :1.0000
##              Mean   :0.1869
##              3rd Qu.:0.0000
##              Max.   :1.0000
```



### Santa Catarina 2T2022

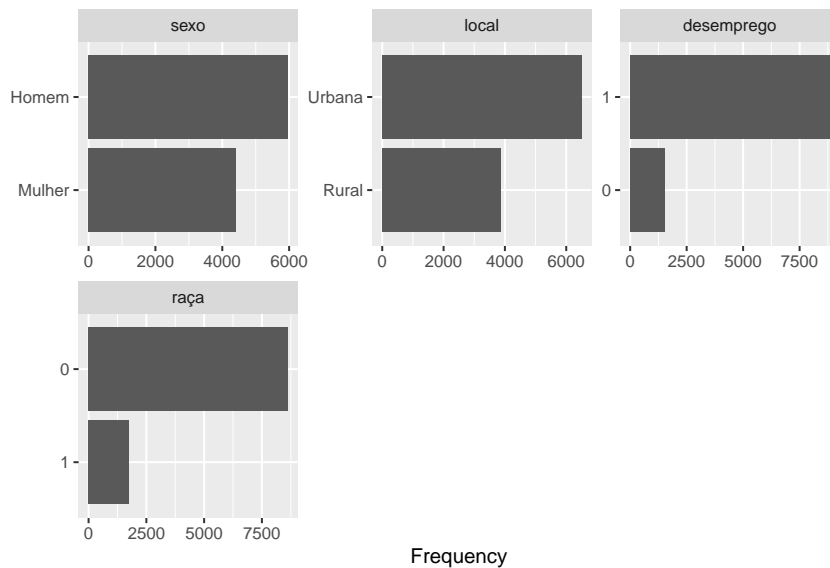
##	desemprego	sexo	raça	local
##	Min. :0.0000	Homem :8818	Min. :0.0000	Urbana:12768
##	1st Qu.:1.0000	Mulher:7162	1st Qu.:1.0000	Rural : 3212
##	Median :1.0000		Median :1.0000	
##	Mean :0.9614		Mean :0.7892	
##	3rd Qu.:1.0000		3rd Qu.:1.0000	
##	Max. :1.0000		Max. :1.0000	



### Bahia 2T2022

##	desemprego	sexo	raça	local
##	Min. :0.0000	Homem :5960	Min. :0.0000	Urbana:6497
##	1st Qu.:1.0000	Mulher:4400	1st Qu.:0.0000	Rural :3863

```
## Median :1.0000      Median :0.000
## Mean   :0.8513      Mean    :0.168
## 3rd Qu.:1.0000      3rd Qu.:0.000
## Max.   :1.0000      Max.    :1.000
```



No total, as amostras apresentaram uma queda no número de desempregados na ordem de 21 mil pessoas para SC e para BA uma alta de 53 mil pessoas. Quando olhamos para o sexo, em SC, temos que as mulheres registram uma queda de 27 mil pessoas desempregadas e homens uma alta de 6 mil pessoas, enquanto na BA as mulheres tiveram um incremento de 79 mil pessoas desocupadas e os homens uma queda de 26 mil. Pelo lado da localização, os desocupados da área rural de SC tiveram queda de 11 mil postos e na BA uma diminuição de 3 mil posto, ao mesmo tempo que na área urbana de SC a desocupação caiu em 1280 mil posto e na BA houve alta de 56 mil pessoas caindo no desemprego.

### 1.32 Modelos Logit

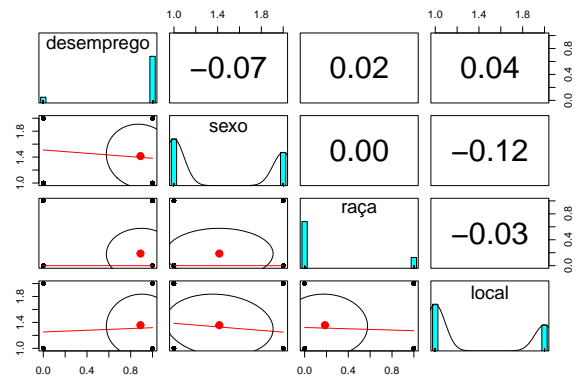
Para esta seção tomou-se o desemprego como variável explicativa e as demais como dependentes. Na sequência, como tais estimações podem sofrer com a influência de *outliers*, utilizou-se o método *Cook's Distance* para avaliar tal hipótese. Ao final, concluiu-se que os modelos não tiveram observações influentes, dado que nenhuma observação ultrapassou a linha de significância de 5%. Ademais, testou-se a possibilidade de multicolinearidade por meio do coeficiente de correlação de Pearson que, basicamente, pode ser entendido como um parâmetro que varia entre 1 e -1, sendo que, idealmente, ele deve estar abaixo de  $|0,8|$ . Para aplicação foi confirmado a ausência de multicolinearidade.

#### Santa Catarina 2T2012

```
##
## Call:
## glm(formula = desemprego ~ raça + sexo + local, family = binomial(link = "logit"),
##      data = sc_1)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.9161   0.2349   0.2349   0.3108   0.4215
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
```



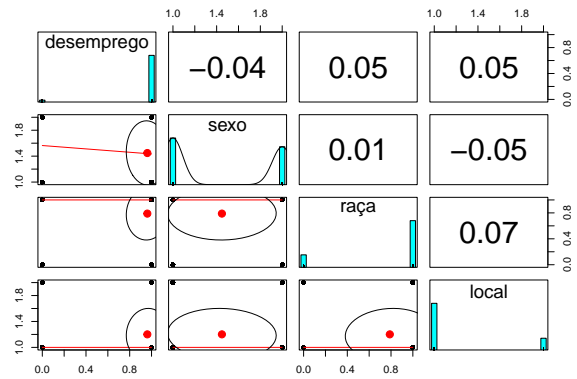
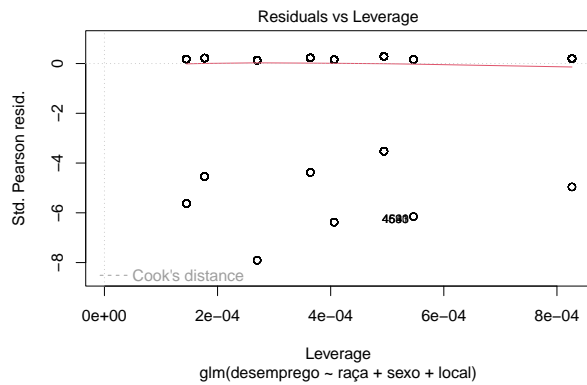
```
##
## Number of Fisher Scoring iterations: 5
```



Para o segundo período, estimou-se os mesmo modelos e testes acima. No que se refere aos teste, eles apontam ausência de *outliers* e de multicolinearidade.

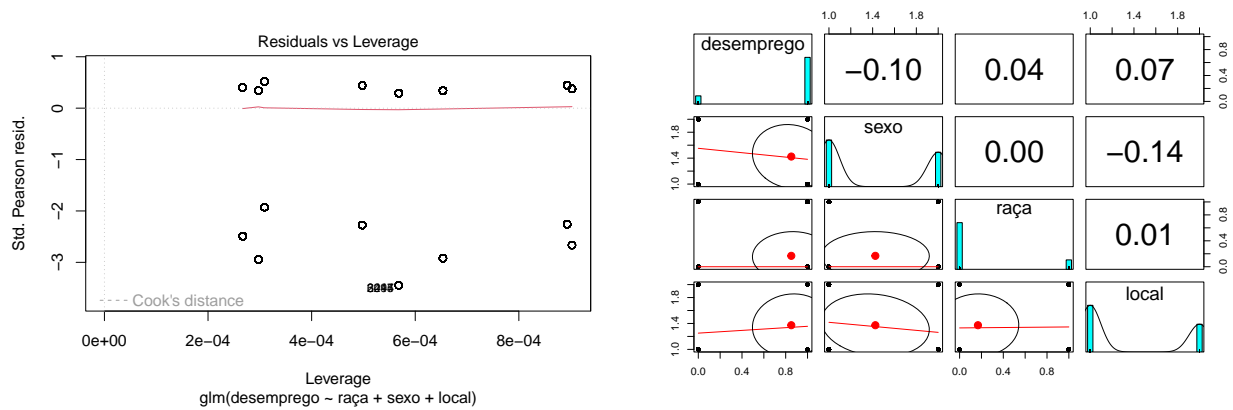
### Santa Catarina 2T2022

```
##
## Call:
## glm(formula = desemprego ~ raça + sexo + local, family = binomial(link = "logit"),
##      data = sc_2)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8815   0.2495   0.2495   0.3081   0.3932
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.95138    0.08781  33.610 < 2e-16 ***
## raça         0.50301    0.08984   5.599 2.16e-08 ***
## sexoMulher  -0.43055    0.08285  -5.197 2.03e-07 ***
## localRural   0.68141    0.13141   5.185 2.16e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 5225.6  on 15979  degrees of freedom
## Residual deviance: 5130.7  on 15976  degrees of freedom
## AIC: 5138.7
##
## Number of Fisher Scoring iterations: 6
```



## Bahia 2T2022

```
##
## Call:
## glm(formula = desemprego ~ raça + sexo + local, family = binomial(link = "logit"),
##      data = ba_2)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2606   0.4677   0.5464   0.5944   0.6905
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.82647    0.04727  38.641 < 2e-16 ***
## raça         0.31632    0.08055   3.927 8.60e-05 ***
## sexoMulher  -0.51407    0.05611  -9.163 < 2e-16 ***
## localRural   0.33154    0.06069   5.463 4.68e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 8713.3  on 10359  degrees of freedom
## Residual deviance: 8565.7  on 10356  degrees of freedom
## AIC: 8573.7
##
## Number of Fisher Scoring iterations: 4
```



## Resultados

Por fim, vamos calculou-se a razão de chances para os modelos e analisar os resultados.

Coeficientes	SC_1	BA_1	SC_2	BA_2
Intercepto	19.0291	8.7490	19.1323	6.2119
Branco	1.8778	1.2157	1.6537	1.3721
Sexo (Mulheres)	0.5657	0.6719	0.6502	0.5981
Local (Rural)	1.9376	1.2232	1.9767	1.3931

Para o 2T2012, temos que a razão de chances de pessoas brancas estarem empregadas em SC é de 87,7% maior que os não brancos, enquanto na BA a probabilidade é de 21,5%. Pelo lado geográfico, SC acusou que os trabalhadores da zona rural têm a sua probabilidade multiplicada por 1,93 para o evento ocupação em comparação a área urbana. Já na BA, tal chance cai para 1,22 de estarem empregados frente aos trabalhadores urbanos. No tocante ao sexo, SC registrou que ser mulher amplia em 0,56 as chances de estarem empregadas ou tem 43,4% menos chances que os homens. Para BA, tal resultado se reduz para 32,8% menos de chance.

Quando se tratamos do 2T2022, no que se refere à raça, observamos dinâmicas diferentes entre as UFs. SC apresentou uma queda das chances de não brancos estarem empregados, enquanto a BA uma alta, porém, a probabilidade de os brancos estarem empregados seguem maiores no primeiro Estado. Geograficamente, as chances de estar na zona rural e empregado em SC e na BA evoluíram, ainda assim, SC apresentou maior probabilidade de se estar fora da zona urbana e empregado que a UF nordestina. Finalmente, as chances de ser mulher em SC e estar empregada aumentou e na BA diminuiu, contudo, os homens seguem tendo maiores chances de emprego que as mulheres em ambas as UFs.

### 1.33 Modelos Probit

Nesta seção será apresentado os resultados dos modelos, cuja aplicação foi a mesma dos modelos logit, e para interpretação dos resultados foi adotado o método do efeito marginal médio para dos coeficientes. Por fim, como a aplicação contou com a mesma base de dados e o processo de estimação dos parâmetros é similar, verificou-se a ausência multicolinearidade e de pontos influentes.

### Santa Catarina 2T2012



```
##
## Call:
## glm(formula = desemprego ~ raça + sexo + local, family = binomial(link = "probit"),
##      data = sc_1)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.9286   0.2348   0.2348   0.3114   0.4154
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.63970    0.04752  34.504 < 2e-16 ***
## raça         0.28405    0.04782   5.940 2.86e-09 ***
## sexoMulher  -0.25230    0.03595  -7.018 2.25e-12 ***
## localRural   0.28126    0.05064   5.554 2.80e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 5508.5  on 17913  degrees of freedom
## Residual deviance: 5389.9  on 17910  degrees of freedom
## AIC: 5397.9
##
## Number of Fisher Scoring iterations: 6
```

## Bahia 2T2012

```
##
## Call:
## glm(formula = desemprego ~ raça + sexo + local, family = binomial(link = "probit"),
##      data = ba_1)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3016   0.4225   0.4657   0.5130   0.5596
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.26587    0.02396  52.843 < 2e-16 ***
## raça         0.10047    0.03803   2.642 0.008239 **
## sexoMulher  -0.20756    0.02901  -7.154 8.45e-13 ***
## localRural   0.10395    0.03081   3.374 0.000741 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 9372.0  on 13514  degrees of freedom
## Residual deviance: 9295.7  on 13511  degrees of freedom
## AIC: 9303.7
##
## Number of Fisher Scoring iterations: 4
```

## Santa Catarina 2T2022

```
##
## Call:
## glm(formula = desemprego ~ raça + sexo + local, family = binomial(link = "probit"),
##      data = sc_2)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8979   0.2494   0.2494   0.3088   0.3911
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.64163    0.04019  40.846 < 2e-16 ***
## raça         0.23006    0.04141   5.556 2.76e-08 ***
## sexoMulher  -0.19243    0.03689  -5.217 1.82e-07 ***
## localRural   0.29803    0.05464   5.455 4.91e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 5225.6  on 15979  degrees of freedom
## Residual deviance: 5129.9  on 15976  degrees of freedom
## AIC: 5137.9
##
## Number of Fisher Scoring iterations: 6
```

## Bahia 2T2022

```
##
## Call:
## glm(formula = desemprego ~ raça + sexo + local, family = binomial(link = "probit"),
##      data = ba_2)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2702   0.4663   0.5476   0.5962   0.6889
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.08376    0.02548  42.537 < 2e-16 ***
## raça         0.16789    0.04275   3.927 8.61e-05 ***
## sexoMulher  -0.28156    0.03069  -9.175 < 2e-16 ***
## localRural   0.18074    0.03258   5.547 2.91e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 8713.3  on 10359  degrees of freedom
## Residual deviance: 8565.6  on 10356  degrees of freedom
## AIC: 8573.6
##
## Number of Fisher Scoring iterations: 4
```

## Resultados

Os efeitos marginais médios calculados seguem na tabale abaixo:

Coefficientes	SC_1	BA_1	SC_2	BA_2
(Intercept)	0.1261	0.2363	0.1353	0.2471
raça	0.0219	0.0188	0.0190	0.0383
sexoMulher	-0.0194	-0.0388	-0.0159	-0.0642
localRural	0.0216	0.0194	0.0246	0.0412

A partir dos resultados, temos que, no primeiro período para SC e na média, o efeito marginal de ser branco e estar empregado é maior do que para os não brancos, assim como na BA, porém o efeito para o Estado é menor do que em SC. A mesma dinâmica se estende para quem vem da zona rural, com isso temos que em SC e na BA trabalhadores fora da zona urbana tem mais chances de estarem empregados, contudo, o efeito médio é maior para quem se encontrado no Estado sulista. Pelo lado do sexo, a dinâmica se inverte, ou seja, as mulheres, tem efeito na margem negativo de estarem empregadas nas UFs, sendo que na BA o efeito negativo é ainda maior do que em SC.

Para o segundo período analisado, os efeitos marginais médio mantiveram o sentido com algumas diferenças frente aos 2T2012. Os brancos para SC reduziram suas chances de estarem empregados contra os não brancos, ao passo que para BA a probabilidade dos brancos estarem empregados teve ligeiro aumento. No tocante à geografia, tanto SC como BA apresentaram resultados maiores em relação ao período passado. Finalmente, para o sexo em SC a probabilidade das mulheres estarem empregadas melhorou, mas na BA a chance de ter um emprego frente aos homens diminuiu.

## 2. Modelos Multinomiais

Diferentemente dos modelos de resposta binária, os modelos multinomiais consideram como variável dependente informações categóricas, porém não ordenadas ou nominais. Em outras palavras, as repostas binárias podem somente considerar variáveis discretas com duas características, ao passo que o método multinomial considera variáveis discretas com mais de duas características. Ademais, outra divergência é a necessidade de modelos nominais apresentarem a independência de alternativas irrelevantes (IIA). Isto é, garantir que variáveis independentes apresentem os mesmos parâmetros estimados para as covariadas mediante a exclusão de alguma categoria das variáveis explicadas. Para tal, será feito o teste de Hausman-McFadden em que a hipótese nula considera a IIA. No mais, as propriedades dos métodos de modelagem logit e probit prevalecem em relação ao arcabouço de modelagem binária.

### 2.1 Aplicação

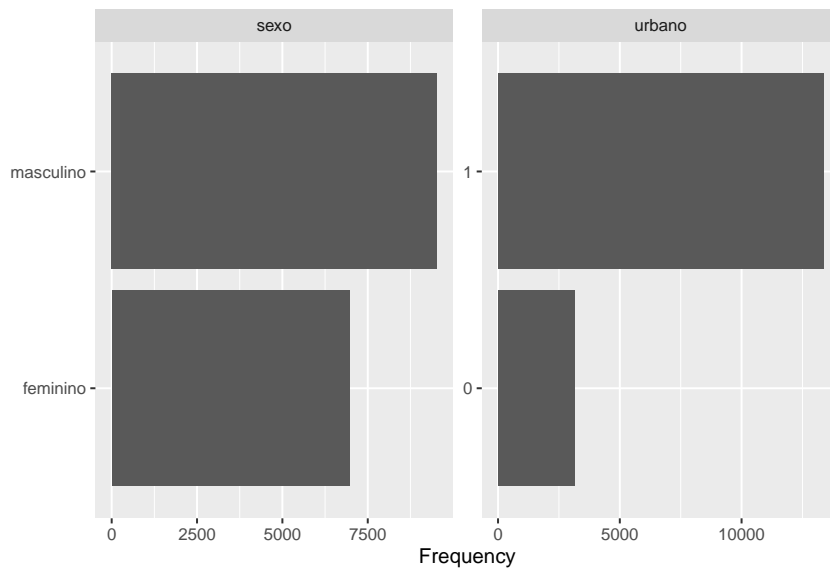
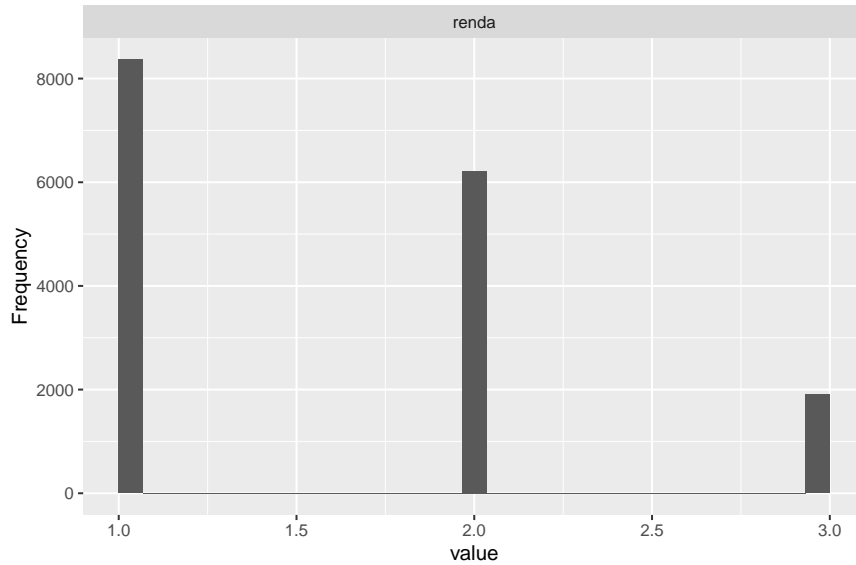
Para a seção em questão, tendo em vista a necessidade de aumentar a dimensão da variável dependente, foi escolhido as mesmas UFs e períodos da aplicação anterior, contudo, o exercício contemplou os microdados da PNAD referentes à renda total do trabalho, sexo e localização, sendo essa última uma dummy onde a zona urbana é 1 e a rural 0. Dito isto, tomou-se a renda como variável dependente, em que foi aplicado recortes para rendimentos abaixo 1000,00, rendimentos entre 1000,01 e 2500,00 e, por fim, rendimentos acima de 2500,01, sendo que tais recortes configuram as categorias 1, 2 e 3, respectivamente

#### 2.1.1 Análise Descritiva

##### Santa Catarina 2T2012

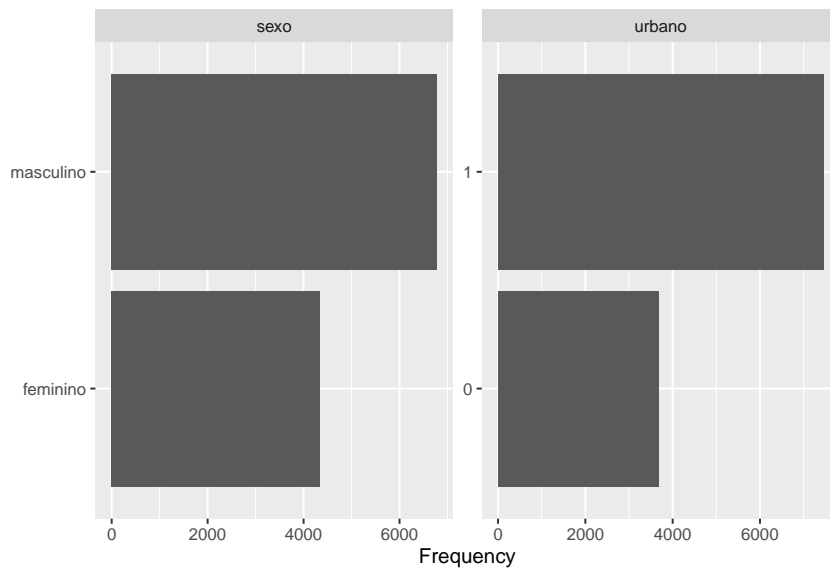
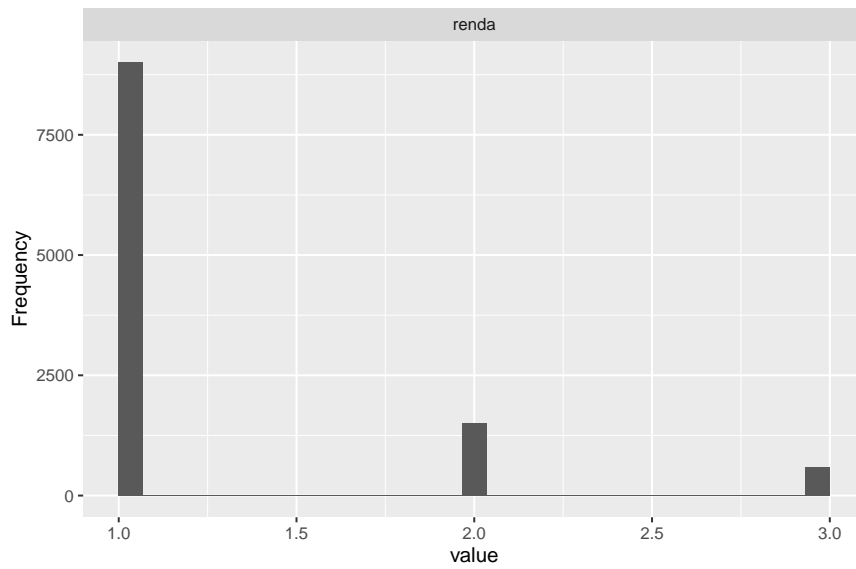
```
##      renda      sexo      urbano
## Min.    :1.000  Length:16484  Min.    :0.0000
```

```
## 1st Qu.:1.000   Class :character   1st Qu.:1.0000
## Median :1.000   Mode  :character   Median :1.0000
## Mean   :1.608                                Mean  :0.8098
## 3rd Qu.:2.000                                3rd Qu.:1.0000
## Max.   :3.000                                Max.   :1.0000
```



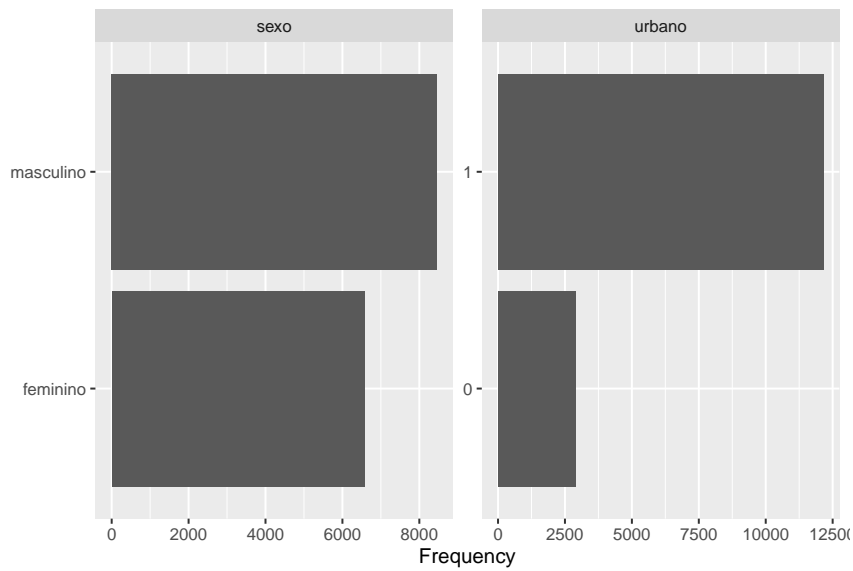
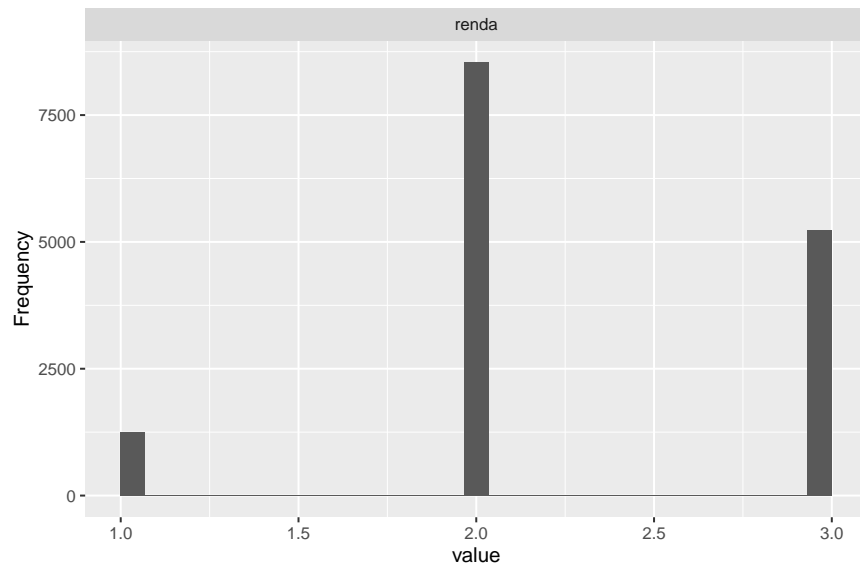
## Bahia 2T2012

```
##      renda      sexo      urbano
## Min.   :1.000   Length:11118   Min.    :0.0000
## 1st Qu.:1.000   Class :character   1st Qu.:0.0000
## Median :1.000   Mode  :character   Median :1.0000
## Mean   :1.243                                Mean  :0.6697
## 3rd Qu.:1.000                                3rd Qu.:1.0000
## Max.   :3.000                                Max.   :1.0000
```



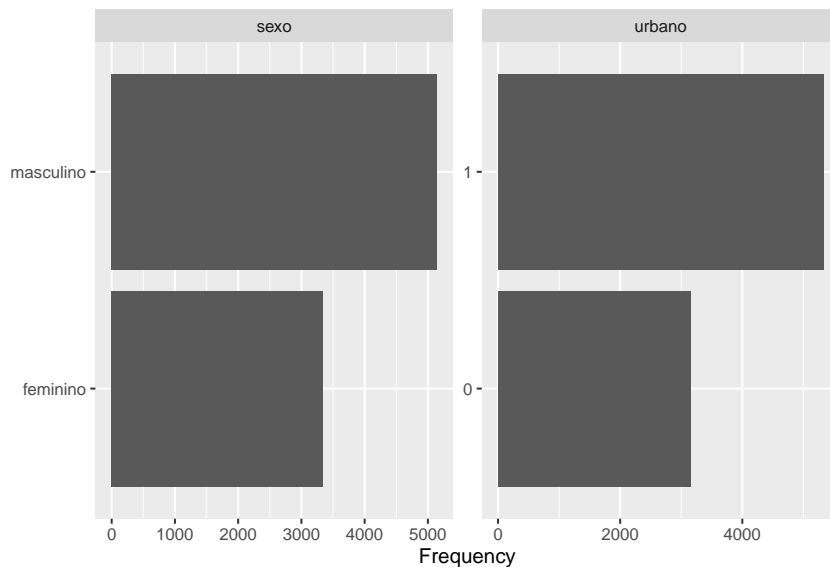
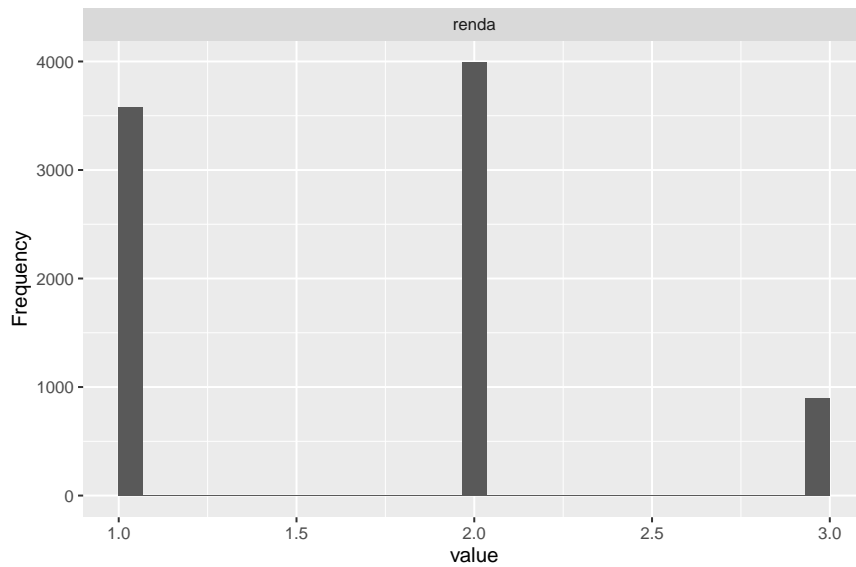
### Santa Catarina 2T2022

##	renda	sexo	urbano
##	Min. :1.000	Length:15029	Min. :0.000
##	1st Qu.:2.000	Class :character	1st Qu.:1.000
##	Median :2.000	Mode :character	Median :1.000
##	Mean :2.265		Mean :0.808
##	3rd Qu.:3.000		3rd Qu.:1.000
##	Max. :3.000		Max. :1.000



## Bahia 2T2022

##	renda	sexo	urbano
##	Min. :1.000	Length:8482	Min. :0.0000
##	1st Qu.:1.000	Class :character	1st Qu.:0.0000
##	Median :2.000	Mode :character	Median :1.0000
##	Mean :1.684		Mean :0.6276
##	3rd Qu.:2.000		3rd Qu.:1.0000
##	Max. :3.000		Max. :1.0000



O comportamento dos dados, para o primeiro período, aponta que a maioria dos indivíduos se concentram na primeira categoria de renda, sendo que para as demais categorias, SC apresenta maior frequência que a BA. Pelo lado da localização, ambas as UFs demonstraram maior concentração na área urbana, porém a BA acusou uma diferença menor de indivíduos entre as geografias. Enquanto, o gênero que mais apareceu na amostra foi o masculino com distribuição similar entre os Estados.

Para o segundo período, temos que a 2ª categoria de renda é quem mais aparece na amostra, sendo que SC acusou mais indivíduos na 3ª categoria que a BA, ao passo que o Estado nordestino teve uma distribuição semelhante entre as duas primeiras categorias. Para localização, a dinâmica das UFs permanece similar ao período passado, onde a BA apresenta mais indivíduos na área rural do que SC, porém a zona urbana segue com maior concentração para ambas. Finalmente, no tocante ao gênero, as mulheres em SC ganharam ligeira parcela de participação na amostra, ao passo que na BA a distribuição segue semelhante ao 2T2012.

## 2.12 Modelos Logit

Os resultados dos modelos se encontram abaixo. Além disso, para fins de avaliar a IIA, retirou-se a categoria 3 e aplicou-se o teste de Hausman-McFadden, sendo que ao final pudemos aceitar a hipótese nula de forma generalizada.

### Santa Catarina 2T2012

```
##
## Call:
## mlogit(formula = renda ~ 1 | sexo + urbano, data = msc_1, reflevel = "1",
##       shape = "wide", method = "nr")
##
## Frequencies of alternatives:choice
##      1      2      3
## 0.50740 0.37709 0.11551
##
## nr method
## 5 iterations, 0h:0m:1s
## g'(-H)^-1g = 2.27E-07
## gradient close to zero
##
## Coefficients :
##              Estimate Std. Error z-value Pr(>|z|)
## (Intercept):2  -1.348930   0.047741 -28.255 < 2.2e-16 ***
## (Intercept):3  -3.116476   0.086483 -36.035 < 2.2e-16 ***
## sexomasculino:2  0.984743   0.035566  27.688 < 2.2e-16 ***
## sexomasculino:3  1.255062   0.056912  22.053 < 2.2e-16 ***
## urbano:2        0.602178   0.044636  13.491 < 2.2e-16 ***
## urbano:3        1.055862   0.078563  13.440 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -15215
## McFadden R^2:  0.039842
## Likelihood ratio test : chisq = 1262.7 (p.value = < 2.22e-16)
##
## Hausman-McFadden test
##
## data:  msc_1
## chisq = 0.22433, df = 3, p-value = 0.9736
## alternative hypothesis: IIA is rejected
```

### Bahia 2T2012

```
##
## Call:
## mlogit(formula = renda ~ 1 | sexo + urbano, data = mba_1, reflevel = "1",
##       shape = "wide", method = "nr")
##
## Frequencies of alternatives:choice
##      1      2      3
## 0.810667 0.135186 0.054146
##
## nr method
```



```

## 7 iterations, 0h:0m:1s
## g'(-H)^-1g = 2.05E-05
## successive function values within tolerance limits
##
## Coefficients :
##           Estimate Std. Error z-value Pr(>|z|)
## (Intercept):2  -3.249906   0.087907 -36.9700 < 2.2e-16 ***
## (Intercept):3  -4.889403   0.181768 -26.8991 < 2.2e-16 ***
## sexomasculino:2  0.506987   0.060313   8.4059 < 2.2e-16 ***
## sexomasculino:3  0.437908   0.089047   4.9177 8.755e-07 ***
## urbano:2        1.501024   0.080401  18.6692 < 2.2e-16 ***
## urbano:3        2.347929   0.173928  13.4994 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -6268.2
## McFadden R^2:  0.058116
## Likelihood ratio test : chisq = 773.51 (p.value = < 2.22e-16)
##
## Hausman-McFadden test
##
## data:  mba_1
## chisq = 0.37179, df = 3, p-value = 0.946
## alternative hypothesis: IIA is rejected

```

## Santa Catarina 2T2022

```

##
## Call:
## mlogit(formula = renda ~ 1 | sexo + urbano, data = msc_2, reflevel = "1",
##        shape = "wide", method = "nr")
##
## Frequencies of alternatives:choice
##           1           2           3
## 0.083039 0.568634 0.348327
##
## nr method
## 6 iterations, 0h:0m:1s
## g'(-H)^-1g = 3.93E-05
## successive function values within tolerance limits
##
## Coefficients :
##           Estimate Std. Error z-value Pr(>|z|)
## (Intercept):2   1.2482051   0.0653717 19.0940 < 2.2e-16 ***
## (Intercept):3  -0.0072744   0.0740200 -0.0983   0.9217
## sexomasculino:2  0.5648986   0.0622815   9.0701 < 2.2e-16 ***
## sexomasculino:3  1.1530352   0.0656120  17.5735 < 2.2e-16 ***
## urbano:2        0.5527957   0.0681531   8.1111 4.441e-16 ***
## urbano:3        1.0560278   0.0748490  14.1088 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -13148
## McFadden R^2:  0.022518

```

```

## Likelihood ratio test : chisq = 605.78 (p.value = < 2.22e-16)

##
## Hausman-McFadden test
##
## data: msc_2
## chisq = -12.3, df = 3, p-value = 1
## alternative hypothesis: IIA is rejected

Bahia 2T2022

##
## Call:
## mlogit(formula = renda ~ 1 | sexo + urbano, data = mba_2, reflevel = "1",
## shape = "wide", method = "nr")
##
## Frequencies of alternatives:choice
##      1      2      3
## 0.42254 0.47100 0.10646
##
## nr method
## 6 iterations, 0h:0m:1s
## g'(-H)^-1g = 0.000751
## successive function values within tolerance limits
##
## Coefficients :
##              Estimate Std. Error z-value Pr(>|z|)
## (Intercept):2  -0.780557   0.052413 -14.8923 < 2.2e-16 ***
## (Intercept):3  -3.114541   0.114605 -27.1763 < 2.2e-16 ***
## sexomascuino:2  0.269201   0.050021   5.3818 7.375e-08 ***
## sexomascuino:3  0.324582   0.079528   4.0814 4.477e-05 ***
## urbano:2       1.219916   0.050011  24.3929 < 2.2e-16 ***
## urbano:3       2.223643   0.108282  20.5356 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -7644.3
## McFadden R^2: 0.058356
## Likelihood ratio test : chisq = 947.47 (p.value = < 2.22e-16)

##
## Hausman-McFadden test
##
## data: mba_2
## chisq = 0.31673, df = 3, p-value = 0.9569
## alternative hypothesis: IIA is rejected

```

## Resultados

A seguir, temos a tabela com a razão de chances calculadas a partir da exponencialização dos coeficientes estimados dos modelos:

Coefficientes	SC_1	BA_1	SC_2	BA_2
(Intercept):2	0.2595	0.0388	3.4841	0.4582

Coeficientes	SC_1	BA_1	SC_2	BA_2
(Intercept):3	0.0443	0.0075	0.9928	0.0444
sexomascuino:2	2.6771	1.6603	1.7593	1.3089
sexomascuino:3	3.5081	1.5495	3.1678	1.3835
urbano:2	1.8261	4.4863	1.7381	3.3869
urbano:3	2.8745	10.4639	2.8749	9.2409

No que tange à interpretação dos resultados, primeiramente, vale ressaltar que todos os preditores apresentaram significância aos níveis de 1%, 5% e 10%, à exceção do intercepto do 2T2022 para SC na categoria 3. No mais, pelo lado das razões de chances calculadas, nota-se, de forma generalizada, que a chance de alguma covariada estar na faixa de renda 2 ou 3, dado um incremento de 1 unidade nos regressores, é sempre maior do que estar na faixa referente. Em outras palavras, para o 2T2012, temos que quando colocamos 1 homem a mais na amostra de SC a probabilidade dele estar na faixa 2 em relação a 1 é multiplicada em 2,6 e, em relação 3 frente 1, 3,5. Ademais, quando comparamos as duas UFs, verificou-se que o impacto do sexo masculino na probabilidade é sempre maior em SC do que na BA, para ambos os períodos, e que o efeito na probabilidade da área urbana é sempre maior na BA ante SC. Finalmente, é importante salientar que houve uma redução das chances quando olhamos para o 2T2022.

### 3. Modelos de Contagem

As variáveis aleatórias discretas podem apresentar a característica de contagem, uma vez que assumem somente valores positivos inteiros cuja incidência de zeros reflete a baixa probabilidade do evento ocorrer em conjunto a alta concentração de observações em valores específicos, bem como a grande heterocedasticidade que acelera a medida que nos aproximamos da média. Desta maneira, os métodos lineares não configuram uma boa opção de estimação para tais dados. O trabalho, por sua vez, pretende explorar os modelos de Poisson e Binomial Negativo

#### 3.1 Poisson

O modelo em destaque pressupõe que a variável de interesse apresente média e variância idêntica, a equidispersão. A sua estimação é feita por meio da máxima verossimilhança, em que a parametrização busca a relação entre a média do processo de contagem e as covariadas independentes entre si. Após calcular as condições de primeira ordem da função de máxima verossimilhança, encontramos relações não lineares que serão resolvidas pelo algoritmo de Newton-Raphson.

#### 3.2 Binomial Negativa

Quando estamos fora do mundo gaussiano muitas vezes a modelagem se depara com a sobredispersão de uma variável de contagem, tendo em vista a alta variância do processo gerador de dados. A resposta para esse tipo de problema surge, em grande medida, para métodos que envolvem o lado exponencial e funções de controle da variância. O modelo Binomial negativo se apresenta como um deles. Nesse sentido, quando a regressão de Poisson foge da equidispersão, a modelagem traz um elemento de controle para variância para assegurar uma robustez maior.

#### 3.3 Aplicação

Esta seção consiste em uma primeira estimação de Poisson para número de filhos até 14 anos(nfilhos014) versus anos de estudo (anos de estudos), renda do trabalho (rendtrabt) e idade a partir dos microdados da PNAD para os mesmos períodos e UFs analisados anteriormente. Posteriormente, será feita uma regressão

Binomial Negativo, dado os resultados de sobredispersão dos modelos de Poisson, e uma comparação entre os resultados dos métodos.

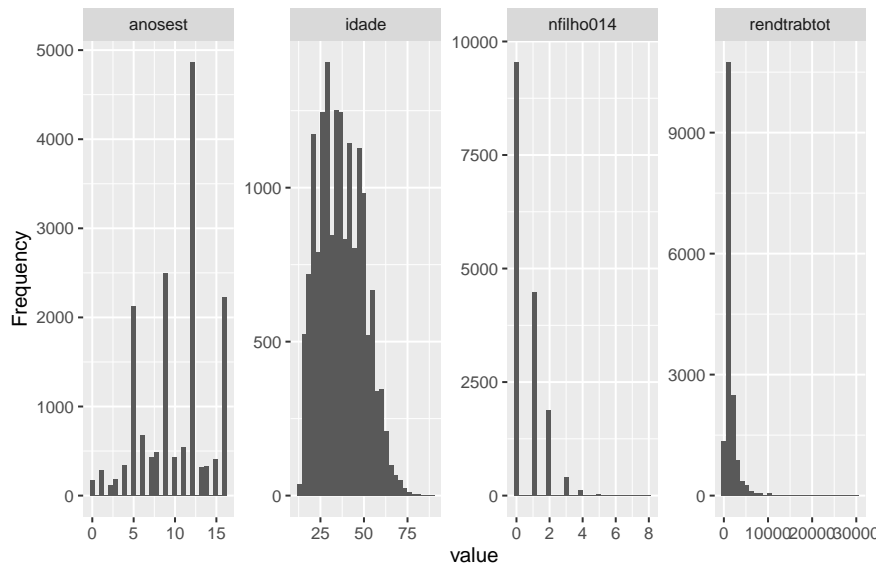
### 3.31 Análise Descritiva

Tal parte apresenta a distribuição das variáveis, bem como as suas estatística em que a média e a variância da variável de interesse estão indicadas pelos itens [1], respectivamente:

#### Santa Catarina 2T2012

```
##      nfilho014      idade      anosest      rendtrabtot
## Min.      :0.0000   Min.      :14.00   Min.      : 0.00   Min.      : 10
## 1st Qu.:0.0000   1st Qu.:27.00   1st Qu.: 7.00   1st Qu.: 750
## Median :0.0000   Median :36.00   Median :11.00   Median : 1000
## Mean    :0.6187   Mean    :36.93   Mean    :10.07   Mean    : 1526
## 3rd Qu.:1.0000   3rd Qu.:46.00   3rd Qu.:12.00   3rd Qu.: 1650
## Max.    :8.0000   Max.    :90.00   Max.    :16.00   Max.    :30000

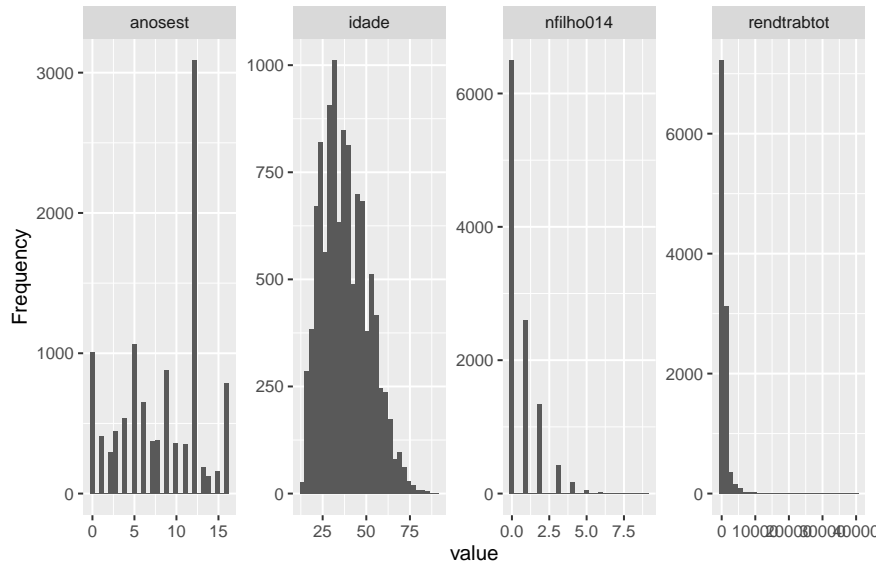
## [1] 0.6186605
## [1] 0.7728508
```



#### Bahia 2T2012

```
##      nfilho014      idade      anosest      rendtrabtot
## Min.      :0.0000   Min.      :14.00   Min.      : 0.000   Min.      : 4.0
## 1st Qu.:0.0000   1st Qu.:28.00   1st Qu.: 5.000   1st Qu.: 300.0
## Median :0.0000   Median :36.00   Median : 9.000   Median : 622.0
## Mean    :0.6872   Mean    :37.99   Mean    : 8.193   Mean    : 903.5
## 3rd Qu.:1.0000   3rd Qu.:47.00   3rd Qu.:12.000   3rd Qu.: 872.0
## Max.    :9.0000   Max.    :91.00   Max.    :16.000   Max.    :40000.0

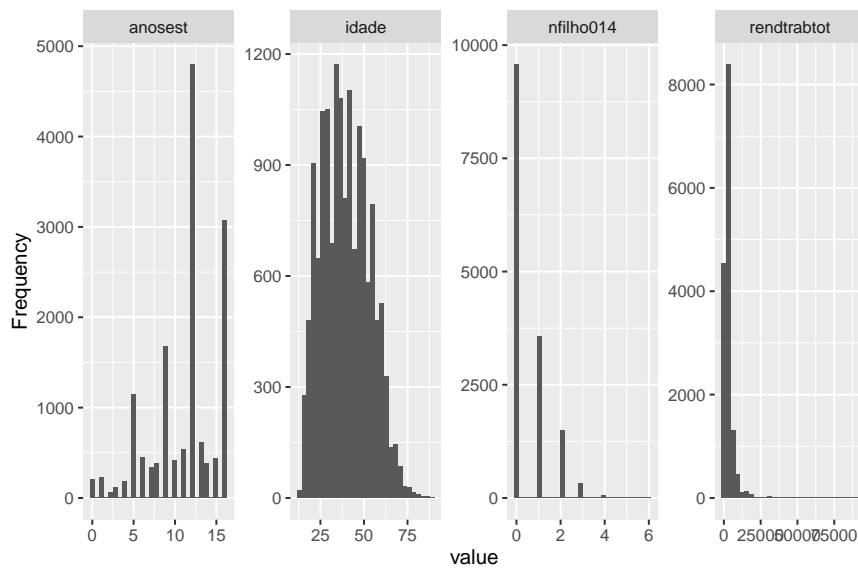
## [1] 0.687174
## [1] 1.025456
```



### Santa Catarina 2T2022

```
##      nfilho014      idade      anosest      rendtrabtot
## Min.   :0.0000   Min.   :14.0   Min.   : 0.00   Min.   :  40
## 1st Qu.:0.0000   1st Qu.:29.0   1st Qu.: 9.00   1st Qu.: 1500
## Median :0.0000   Median :39.0   Median :12.00   Median : 2000
## Mean   :0.5193   Mean   :39.5   Mean   :11.11   Mean   : 2985
## 3rd Qu.:1.0000   3rd Qu.:50.0   3rd Qu.:14.00   3rd Qu.: 3000
## Max.   :6.0000   Max.   :90.0   Max.   :16.00   Max.   :90000

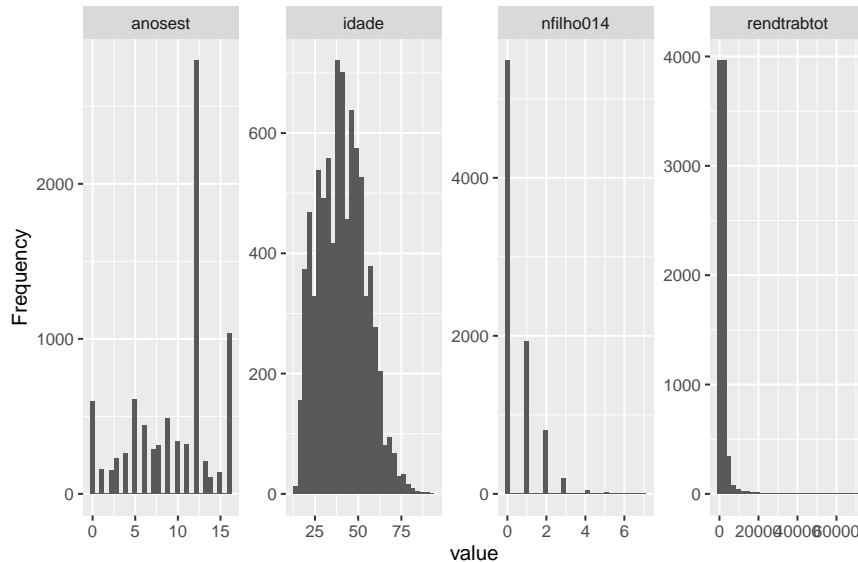
## [1] 0.5193293
## [1] 0.6363877
```



### Bahia 2T2022

```
##      nfilho014      idade      anosest      rendtrabtot
## Min.   :0.0000   Min.   :14.00   Min.   : 0.00   Min.   :    5
## 1st Qu.:0.0000   1st Qu.:30.00   1st Qu.: 6.00   1st Qu.:   500
## Median :0.0000   Median :40.00   Median :12.00   Median : 1212
## Mean   :0.5185   Mean   :40.57   Mean   : 9.45   Mean   : 1525
## 3rd Qu.:1.0000   3rd Qu.:50.00   3rd Qu.:12.00   3rd Qu.: 1500
## Max.   :7.0000   Max.   :93.00   Max.   :16.00   Max.   :70000

## [1] 0.5185098
## [1] 0.6840696
```



Para 2T2012, os dados apontam que, na média, o número de filhos até 14 anos das duas UFs é semelhante, assim como a idade dos indivíduos. Contudo, pelo lado da escolaridade e dos rendimentos também na média, os dados acusam que SC tem uma concentração em números mais elevados do que a BA.

Para o segundo período em questão, temos que na média o número de filhos se reduziu para ambos os Estados, enquanto a idade na média aumentou, ainda que siga similar entre SC e BA. Já a escolaridade e os rendimentos, seguiram a mesma dinâmica anterior, porém ambas as UFs apresentaram avanços em suas médias. Por fim, vale ressaltar que de forma generalizada não houve equidispersão dos dados, sendo que a BA reduz consideravelmente a diferença entre média e variância para o 2º período.

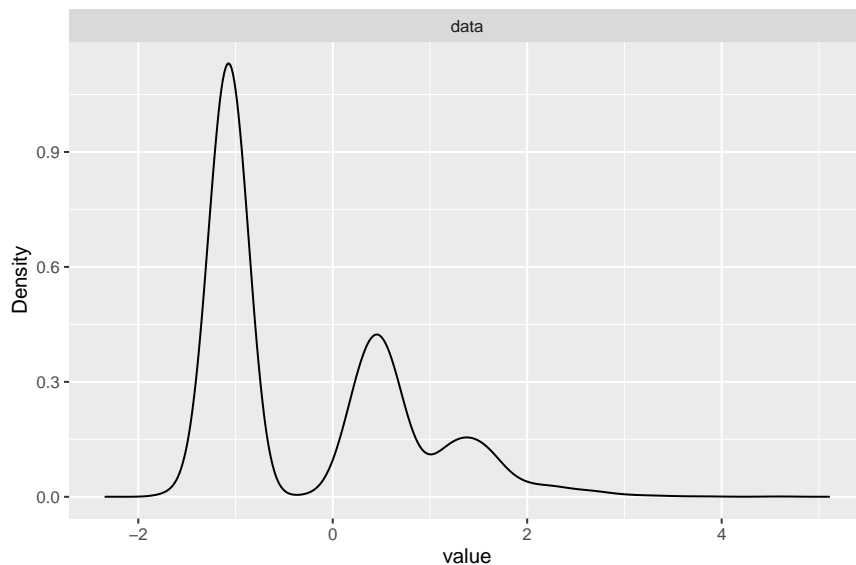
### 3.32 Modelos de Poisson

A seguir segue o *output* dos modelos estimados, bem como a dispersão dos resíduos e os respectivos *fits*

#### Santa Catarina 2T2012

```
##
## Call:
## glm(formula = nfilho014 ~ idade + anosest + rendtrabtot, family = poisson(link = "log"),
##      data = po_sc_1)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3389  -1.1018  -0.9091   0.5150   5.1058
##
```

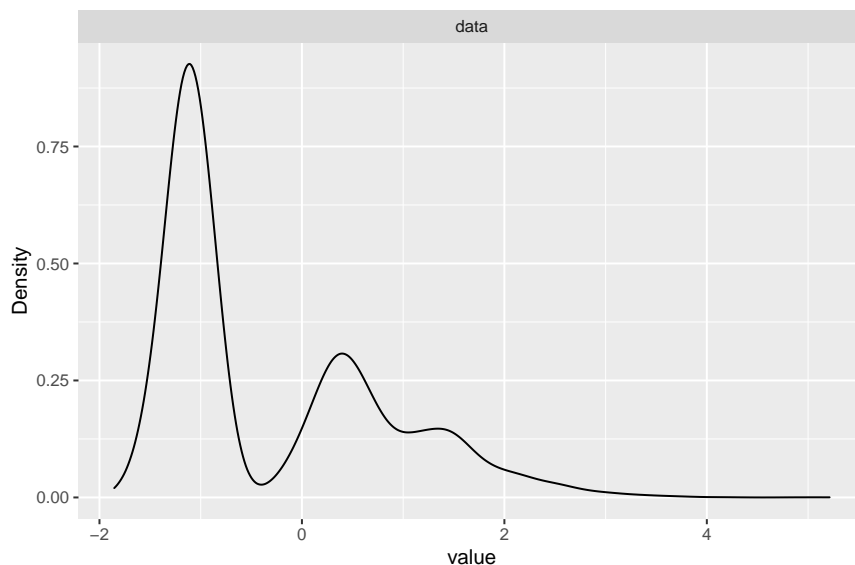
```
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  7.831e-01  4.575e-02  17.118   <2e-16 ***
## idade        -2.115e-02  8.754e-04 -24.159   <2e-16 ***
## anosest      -5.929e-02  2.754e-03 -21.529   <2e-16 ***
## rendtrabtot  5.028e-05  5.724e-06   8.784   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 19898  on 16483  degrees of freedom
## Residual deviance: 19123  on 16480  degrees of freedom
## AIC: 34785
##
## Number of Fisher Scoring iterations: 5
##
## Overdispersion test
##
## data: mpo_sc_1
## z = 11.855, p-value < 2.2e-16
## alternative hypothesis: true dispersion is greater than 1
## sample estimates:
## dispersion
## 1.191516
```



## Bahia 2T2012

```
##
## Call:
## glm(formula = nfilho014 ~ idade + anosest + rendtrabtot, family = poisson(link = "log"),
##      data = po_ba_1)
##
## Deviance Residuals:
```

```
##      Min      1Q   Median      3Q      Max
## -1.8540 -1.1560 -0.9119  0.5050  5.2128
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  9.475e-01  4.574e-02  20.715  < 2e-16 ***
## idade       -2.514e-02  9.794e-04 -25.672  < 2e-16 ***
## anosest     -5.574e-02  2.752e-03 -20.254  < 2e-16 ***
## rendtrabtot  3.570e-05  7.661e-06   4.659  3.18e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 15465  on 11117  degrees of freedom
## Residual deviance: 14621  on 11114  degrees of freedom
## AIC: 25437
##
## Number of Fisher Scoring iterations: 6
##
## Overdispersion test
##
## data: mpo_ba_1
## z = 17.985, p-value < 2.2e-16
## alternative hypothesis: true dispersion is greater than 1
## sample estimates:
## dispersion
## 1.384558
```

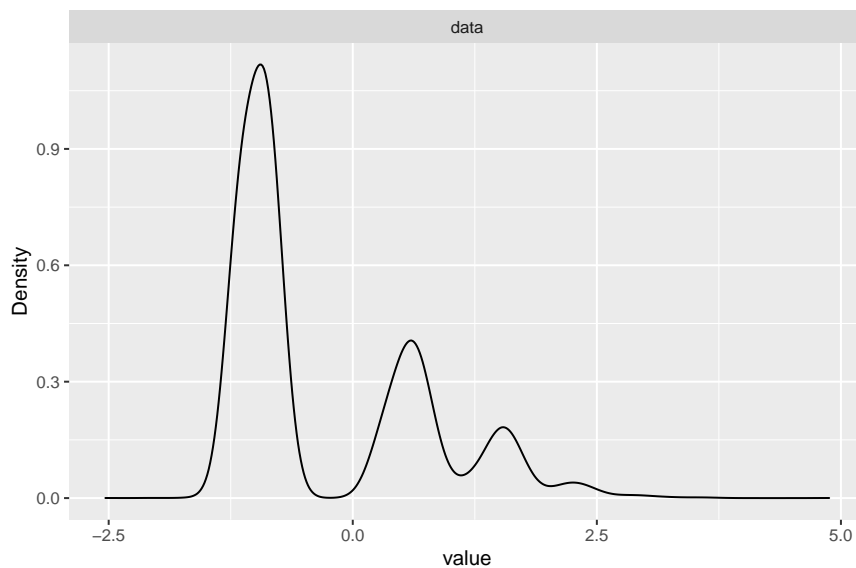


## Santa Catarina 2T2022

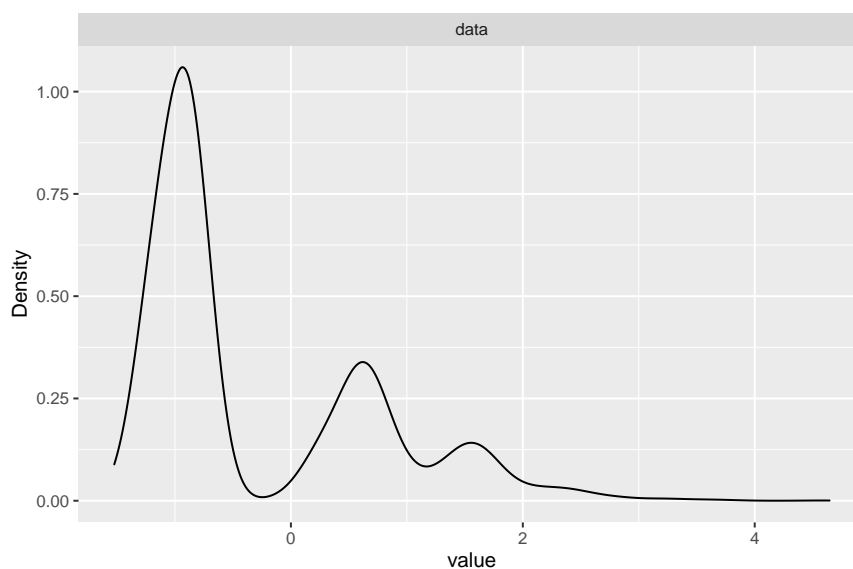
```
##
## Call:
## glm(formula = nfilho014 ~ idade + anosest + rendtrabtot, family = poisson(link = "log"),
```



```
##      data = po_sc_2)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5341  -1.0411  -0.8388   0.5635   4.8796
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.386e-01  5.490e-02   4.346 1.38e-05 ***
## idade        -2.313e-02  9.326e-04 -24.798 < 2e-16 ***
## anosest       -7.400e-03  3.198e-03  -2.314  0.0207 *
## rendtrabtot   1.936e-05  2.442e-06   7.926 2.27e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 17290  on 15028  degrees of freedom
## Residual deviance: 16620  on 15025  degrees of freedom
## AIC: 28862
##
## Number of Fisher Scoring iterations: 5
##
## Overdispersion test
##
## data:  mpo_sc_2
## z = 13.601, p-value < 2.2e-16
## alternative hypothesis: true dispersion is greater than 1
## sample estimates:
## dispersion
##  1.192393
```



```
##
## Call:
## glm(formula = nfilho014 ~ idade + anosest + rendtrabtot, family = poisson(link = "log"),
##      data = po_ba_2)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5233  -1.0180  -0.8279   0.5651   4.6459
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  6.437e-01  6.680e-02   9.637  < 2e-16 ***
## idade        -2.780e-02  1.270e-03 -21.899  < 2e-16 ***
## anosest       -2.701e-02  3.808e-03  -7.093 1.31e-12 ***
## rendtrabtot   1.530e-05  7.097e-06   2.156  0.0311 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 10085.6  on 8481  degrees of freedom
## Residual deviance:  9579.5  on 8478  degrees of freedom
## AIC: 16349
##
## Number of Fisher Scoring iterations: 6
##
## Overdispersion test
##
## data:  mpo_ba_2
## z = 11.603, p-value < 2.2e-16
## alternative hypothesis: true dispersion is greater than 1
## sample estimates:
## dispersion
##  1.265321
```

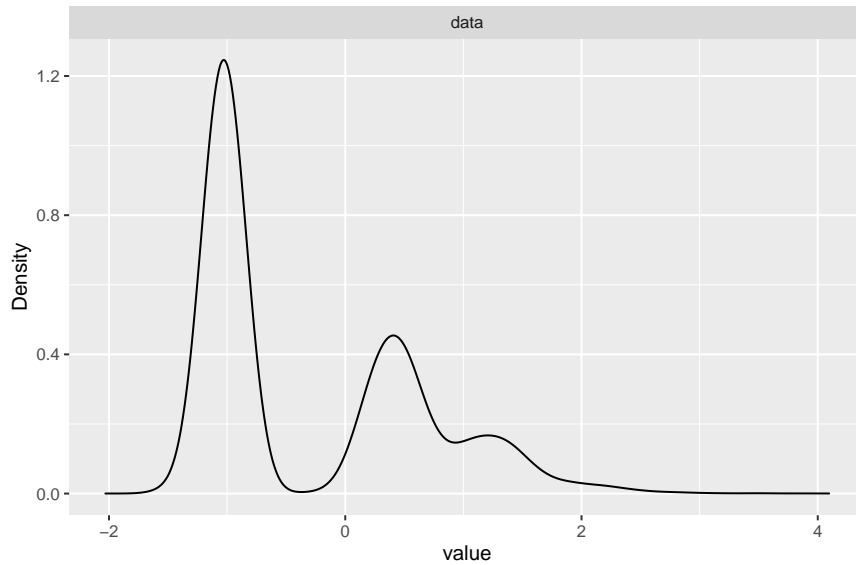


### 3.33 Modelos Binomial Negativa

A seguir segue o *output* dos modelos estimados, bem como a dispersão dos resíduos e os respectivos *fits*

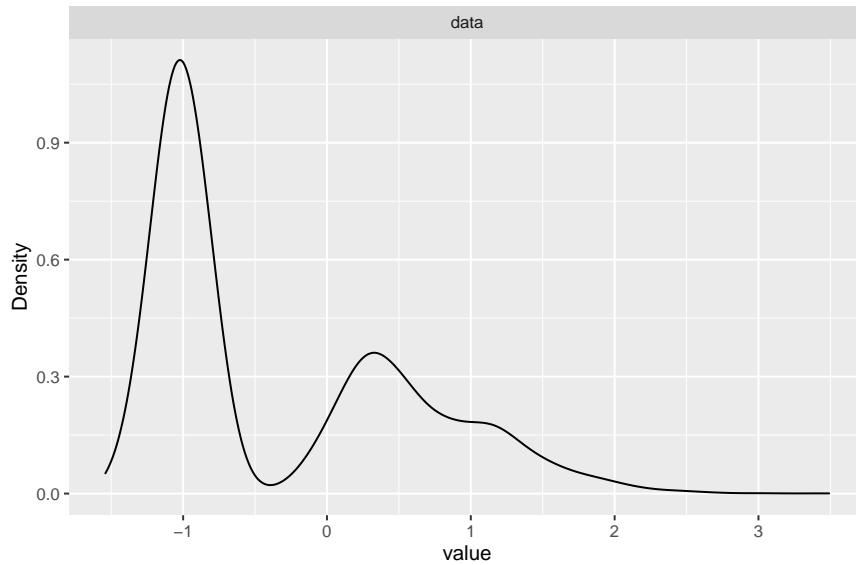
#### Santa Catarina 2T2012

```
##
## Call:
## glm.nb(formula = nfilho014 ~ idade + anosest + rendtrabtot, data = po_sc_1,
##       init.theta = 3.149017284, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0310  -1.0537  -0.8758   0.4667   4.0966
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  8.068e-01  5.094e-02  15.839  < 2e-16 ***
## idade       -2.197e-02  9.637e-04 -22.802  < 2e-16 ***
## anosest     -5.909e-02  3.055e-03 -19.342  < 2e-16 ***
## rendtrabtot  5.289e-05  6.493e-06   8.145  3.8e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(3.149) family taken to be 1)
##
##      Null deviance: 16828  on 16483  degrees of freedom
## Residual deviance: 16169  on 16480  degrees of freedom
## AIC: 34522
##
## Number of Fisher Scoring iterations: 1
##
##              Theta:  3.149
##      Std. Err.:  0.236
##
## 2 x log-likelihood: -34511.569
```



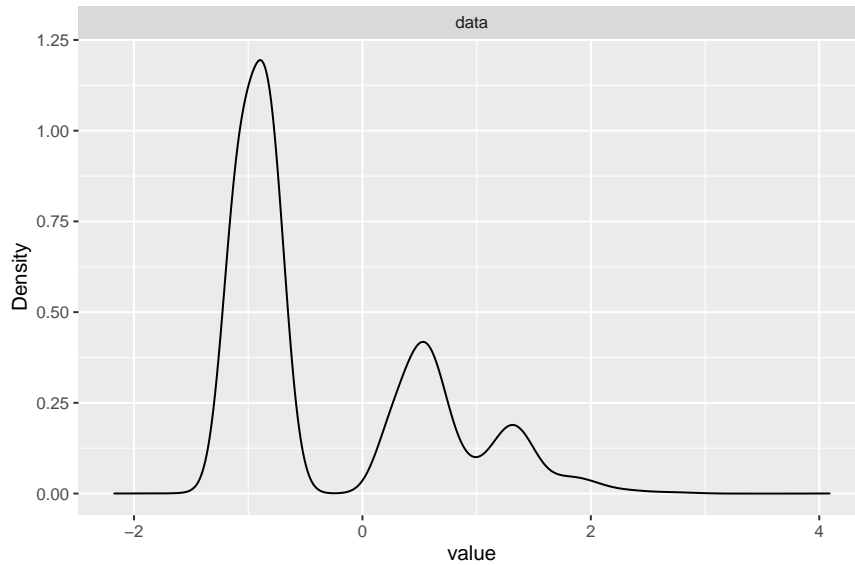
## Bahia 2T2012

```
##
## Call:
## glm.nb(formula = nfilho014 ~ idade + anosest + rendtrabtot, data = po_ba_1,
##       init.theta = 1.548645761, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5409  -1.0536  -0.8432   0.4164   3.4916
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.020e+00  5.689e-02  17.933  < 2e-16 ***
## idade       -2.732e-02  1.187e-03 -23.010  < 2e-16 ***
## anosest     -5.528e-02  3.359e-03 -16.458  < 2e-16 ***
## rendtrabtot  3.907e-05  9.336e-06   4.185 2.85e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(1.5486) family taken to be 1)
##
##    Null deviance: 11111  on 11117  degrees of freedom
## Residual deviance: 10497  on 11114  degrees of freedom
## AIC: 24747
##
## Number of Fisher Scoring iterations: 1
##
##              Theta:  1.5486
##             Std. Err.:  0.0828
##
## 2 x log-likelihood:  -24737.2440
```



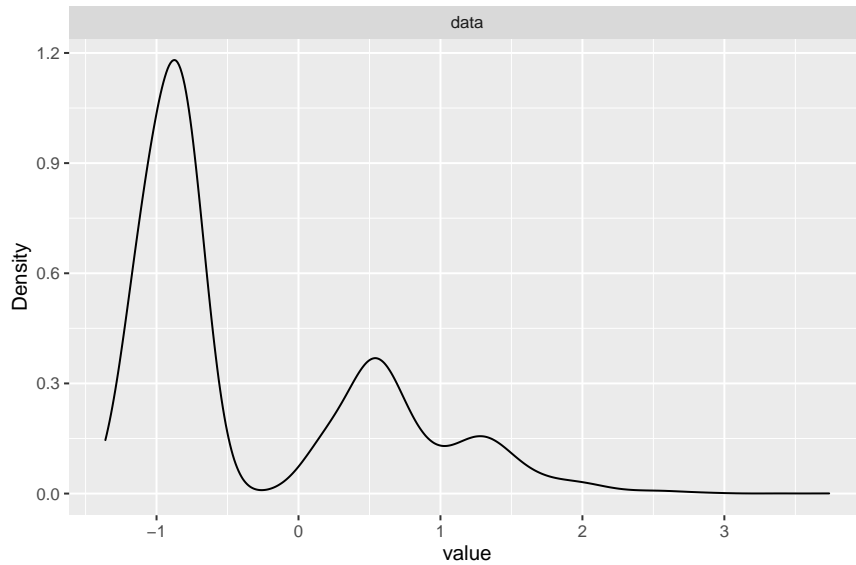
## Santa Catarina 2T2022

```
##
## Call:
## glm.nb(formula = nfilho014 ~ idade + anosest + rendtrabtot, data = po_sc_2,
##       init.theta = 2.357546679, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.1731  -0.9896  -0.7962   0.4954   4.0904
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.960e-01  6.132e-02  4.827 1.39e-06 ***
## idade       -2.510e-02  1.037e-03 -24.200 < 2e-16 ***
## anosest     -6.642e-03  3.554e-03  -1.869  0.0616 .
## rendtrabtot  2.203e-05  2.918e-06   7.550 4.37e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(2.3575) family taken to be 1)
##
##      Null deviance: 14303  on 15028  degrees of freedom
## Residual deviance: 13707  on 15025  degrees of freedom
## AIC: 28601
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta:  2.358
##              Std. Err.:  0.177
##
## 2 x log-likelihood:  -28591.465
```



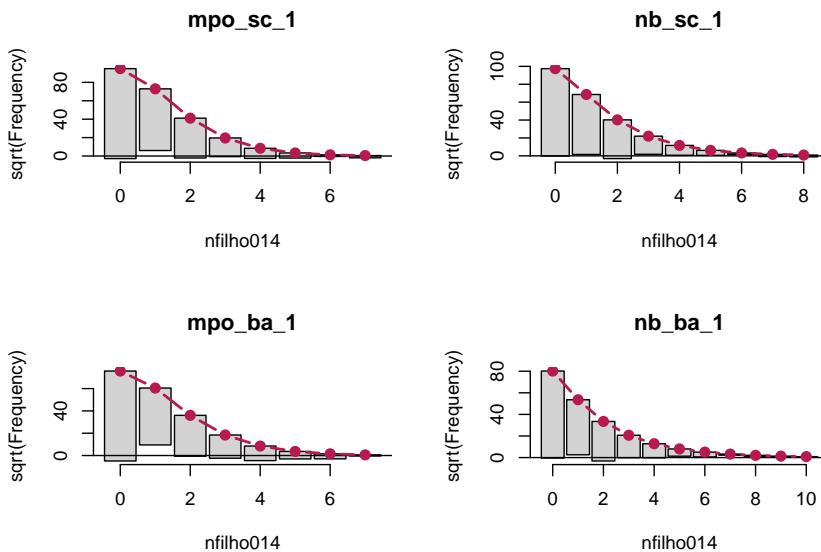
## Bahia 2T2022

```
##
## Call:
## glm.nb(formula = nfilho014 ~ idade + anosest + rendtrabtot, data = po_ba_2,
##       init.theta = 1.789675907, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3606  -0.9538  -0.7783   0.4804   3.7383
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  7.268e-01  7.748e-02   9.380  < 2e-16 ***
## idade       -3.018e-02  1.456e-03 -20.730  < 2e-16 ***
## anosest     -2.637e-02  4.353e-03  -6.057  1.39e-09 ***
## rendtrabtot  1.669e-05  8.026e-06   2.079   0.0376 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(1.7897) family taken to be 1)
##
##      Null deviance: 7905.9  on 8481  degrees of freedom
## Residual deviance: 7481.5  on 8478  degrees of freedom
## AIC: 16101
##
## Number of Fisher Scoring iterations: 1
##
##              Theta:  1.790
##             Std. Err.:  0.147
##
## 2 x log-likelihood:  -16091.468
```

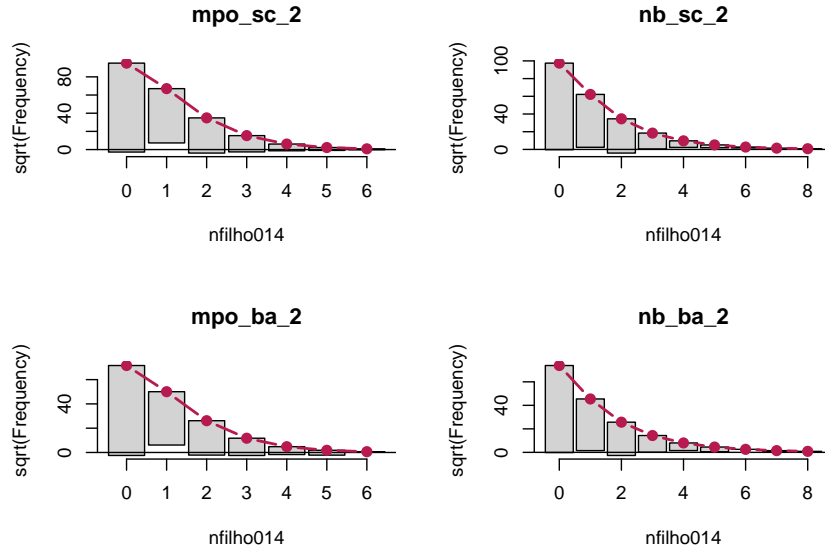


## Resultados

Como a interpretação é bem similar entre os modelos de Poisson e Binomial Negativa, cabe, primeiramente, comentar sobre a robustez de cada um e, assim, realizar uma análise sobre o método que melhor se adequou aos dados. Desta maneira, é sabido que as estimações de Binomial Negativa relaxam alguns pressupostos em relação a Poisson, no sentido de ignorar a sobredispersão dos resultados e gerar estimadores menos eficientes. O trabalho, por sua vez, optou por prevalecer com os modelos de menor *deviance* e melhor *fit* via *rootogram* que, basicamente, demonstram a qualidade de previsão do modelo para cada ponto da variável, em que podemos pensar no *rootograma* como a relação entre as contagens ajustadas - se uma barra não atingir a linha zero, o modelo prediz uma determinada caixa de contagem e, se a barra exceder a linha zero, ela sob prevê. O *fit* para Poisson e Binomial Negativa, estão indicadas pelas variáveis *mpo* e *nb*, respectivamente, assim como os períodos 1 e 2 e a tabela de *deviance* na sequência para cada um deles.



Poisson_SC	Binomial_SC	Poisson_BA	Binomial_BA
19,122.94	16,169.47	14,620.65	10,497.03



Poisson_SC	Binomial_SC	Poisson_BA	Binomial_BA
16,620.11	13,706.57	9,579.543	7,481.495

Em linhas gerais, observou-se que os modelos estimados pela Binomial Negativa apresentaram melhor acomodação dos dados e menor *deviance*. Sendo assim, o trabalho seguiu com seus resultados e a tabela de razão de chances calculadas a partir da exponencialização dos coeficientes dos modelos segue abaixo.

Odds	SC_1	BA_1	SC_2	BA_2
(Intercept)	2.2406	2.7739	1.3444	2.0685
idade	0.9783	0.9730	0.9752	0.9703
anosest	0.9426	0.9462	0.9934	0.9740
rendtrabtot	1.0001	1.0000	1.0000	1.0000

Os resultados dos parâmetros exponencializados apontam, para idade e escolaridade, que à medida que eles avançam a probabilidade de se ter um filho até 14 anos é menor, isso para ambas as UFs nos dois recortes temporais. Enquanto a renda, por mais que tenha apresentado significância para todos os níveis testados, não gera influências na razão de chances do evento. Por fim, a leitura dos resultados sugere que a decisão de ter filhos pouco mudou depois de 10 anos e que a geografia quase não interfere no processo de decisão.

## Referências

- Cameron, A. C. and P. K. Trivedi (2005), Microeconometrics: Methods and Applications, Cambridge University Press, New York, Chapters 14-16