Regressão Beta

Éder David Borges da Silva, Rogerio, Shirley

June 19, 2011

1 A distribuição Beta

$$f(y;p,q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{(p-1)}(1-y)^{(q-1)}, 0 < y < 1 \tag{1}$$

onde p $>$ 0,q $>$ 0 e $\Gamma(.)$ é a função gamma. A média e variância é dada por:

$$E(y) = \frac{p}{(p+q)} \tag{2}$$

e

$$VAR(y) = \frac{pq}{(p+q)^2(p+q+1)} \tag{3}$$

2 Modelo Beta com reparametrização

Para a estimativa vamos reparametrizar o modelo, onde: $\mu = p/(p+q)$ e $\phi = p+q$ equivalente a $p = \mu\phi$ e $q = (1-\mu)\phi$, desta forma a distribuição fica:

$$f(y;p,q) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{(\mu\phi-1)}(1-y)^{((1-\mu)\phi-1)}, 0 < y < 1 \tag{4}$$

onde $0 < \mu < 1$ e $\phi > 0$, desta forma μ e o parametro de média e ϕ de variância do modelo.
Podemos escrever a média no modo como:

$$g(\mu_t) = \sum_{i=1}^{k} x_{ti}\beta_i \tag{5}$$

podemos usar como link a funçao logit, desta forma:

$$\mu_t = \frac{e^{x_t\beta}}{1+e^{x_t\beta}} \tag{6}$$

3 log-verossimilhança

$$l_t(\mu_t,\phi) = log\Gamma(\phi) - log\Gamma(\mu_t\phi) - log\Gamma((1-\mu_t)\phi) + (\mu_t\phi - 1)logy_t + ((1-\mu_t)\phi - 1)log(1-y_t) \tag{7}$$

4 Ajuste do modelo

4.1 dados

Para este ajuste usaremos o conjunto de dados presente no pacote betareg, denominado ”Proportion of Household Income Spent on Food” que contem dados sobre proporção da renda gasta com comida para uma amostra aleatória de 38 famílias em uma cidade dos EUA. onde food é gastos domesticos para a alimentaçao, income é renda familiar e persons é o número de pessoas na familia.

```
     food income persons
1  15.998 62.476       1
2  16.652 82.304       5
3  21.741 74.679       3
4   7.431 39.151       3
5  10.481 64.724       5
6  13.548 36.786       3
7  23.256 83.052       4
8  17.976 86.935       1
9  14.161 88.233       2
10  8.825 38.695       2
11 14.184 73.831       7
12 19.604 77.122       3
13 13.728 45.519       2
14 21.141 82.251       2
15 17.446 59.862       3
16  9.629 26.563       3
17 14.005 61.818       2
18  9.160 29.682       1
19 18.831 50.825       5
20  7.641 71.062       4
21 13.882 41.990       4
22  9.670 37.324       3
23 21.604 86.352       5
24 10.866 45.506       2
25 28.980 69.929       6
26 10.882 61.041       2
27 18.561 82.469       1
28 11.629 44.208       2
29 18.067 49.467       5
30 14.539 25.905       5
31 19.192 79.178       5
32 25.918 75.811       3
33 28.833 82.718       6
34 15.869 48.311       4
35 14.910 42.494       5
36  9.550 40.573       4
37 23.066 44.872       6
38 14.751 27.167       7
```

4.2 Modelo

Como resposta utilizaremos a razão entre food/income e como covariaveis income e persons, desta forma o modelo fica:

$$\mu_t = \frac{e^{\beta_0 + \beta_1 income_t + \beta_2 persons_t}}{1 + e^{\beta_0 + \beta_1 income_t + \beta_2 persons_t}} \tag{8}$$

4.3 Ajuste com o pacote oficial

Primeiramente vamos ajustar o modelo com o pacote oficial do R, betareg

```
Call:
betareg(formula = I(food/income) ~ income + persons, data = FoodExpenditure)

Standardized weighted residuals 2:
    Min      1Q  Median      3Q     Max 
-2.7818 -0.4445  0.2024  0.6852  1.8755 

Coefficients (mean model with logit link):
             Estimate Std. Error z value Pr(>|z|)    
(Intercept) -0.622547   0.223853  -2.781 0.005418 ** 
income      -0.012299   0.003036  -4.052 5.09e-05 ***
persons      0.118462   0.035341   3.352 0.000802 ***

Phi coefficients (precision model with identity link):
      Estimate Std. Error z value Pr(>|z|)    
(phi)    35.61       8.08   4.407 1.05e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Log-likelihood: 45.33 on 4 Df
Pseudo R-squared: 0.3878
Number of iterations in BFGS optimization: 28
```

```
                  2.5 %       97.5 %
(Intercept) -1.06129125 -0.183801787
income      -0.01824850 -0.006349226
persons      0.04919561  0.187728438
(phi)       19.77404055 51.445502974
```

4.4 Ajuste do modelo

Primeiramente escreveremos a log-verossimilhança do modelo com as duas covariaveis e usando como função de ligação a logit.

```
        B0          B1          B2         phi 
-0.62225257 -0.01232139  0.11863161 35.85834022 
```

```
        B0          B1          B2         phi 
0.220704548 0.003075865 0.035623127 8.140336639 
```

	B0	B1	B2	phi
Limite_inferior	-1.0548255	-0.018349974	0.04881157	19.90357
Estimador_pontual	-0.6222526	-0.012321389	0.11863161	35.85834
Limite_superior	-0.1896796	-0.006292805	0.18845166	51.81311

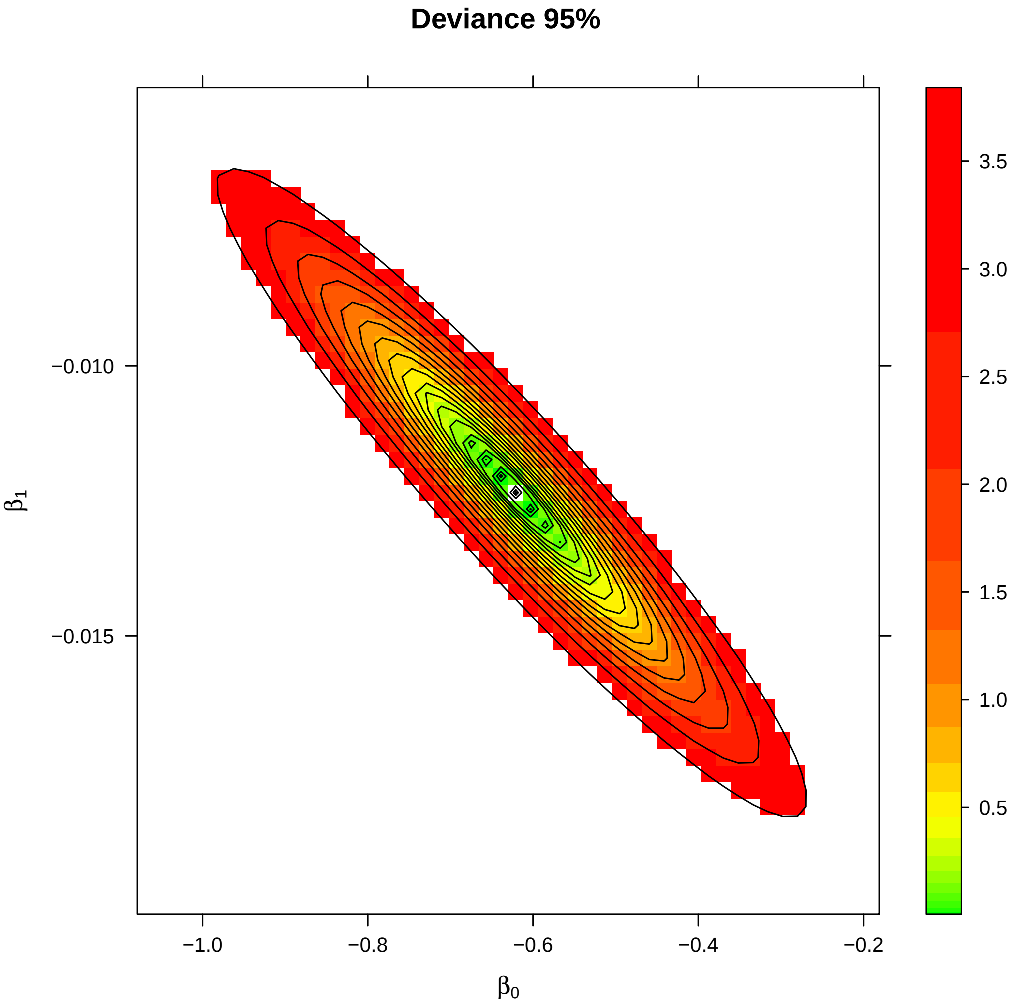

Figure 1: Deviance conjunta β_0 e β_1

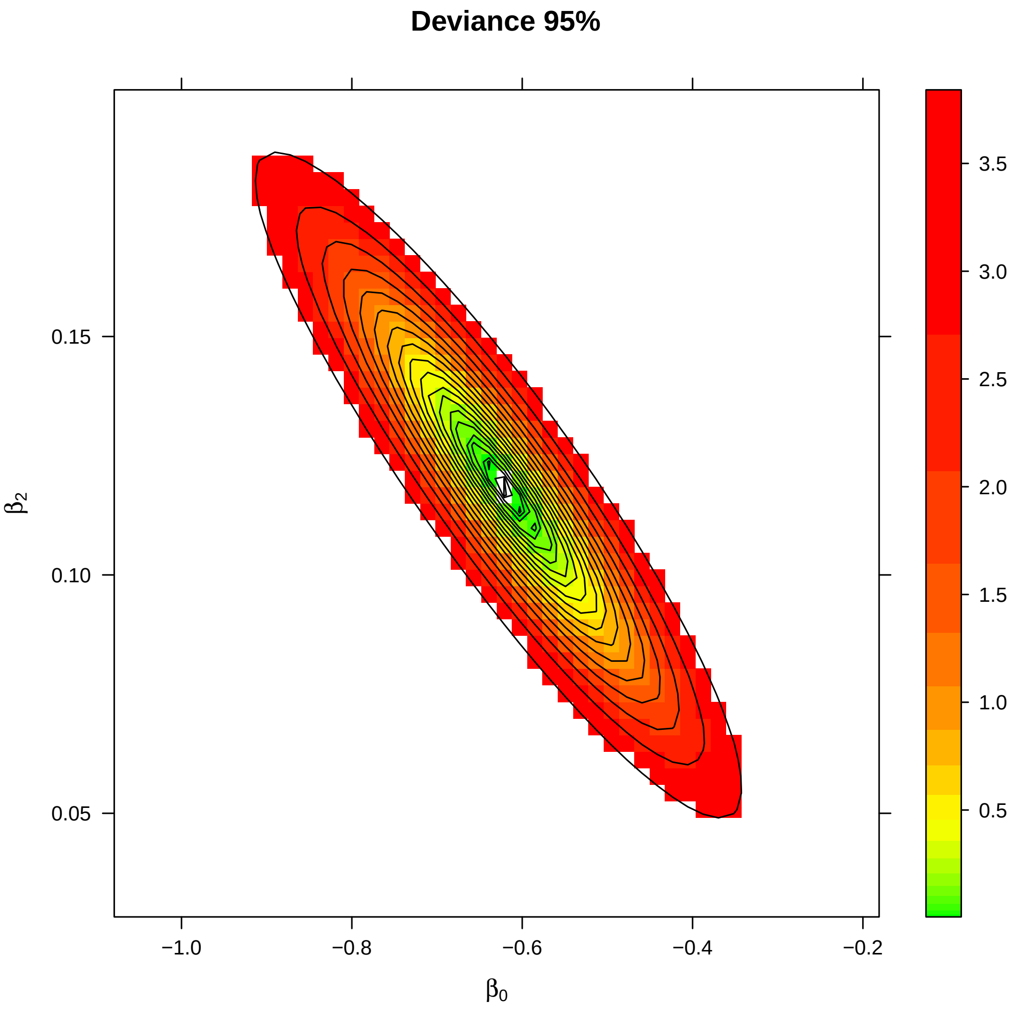

Figure 2: Deviance conjunta β_0 e β_2

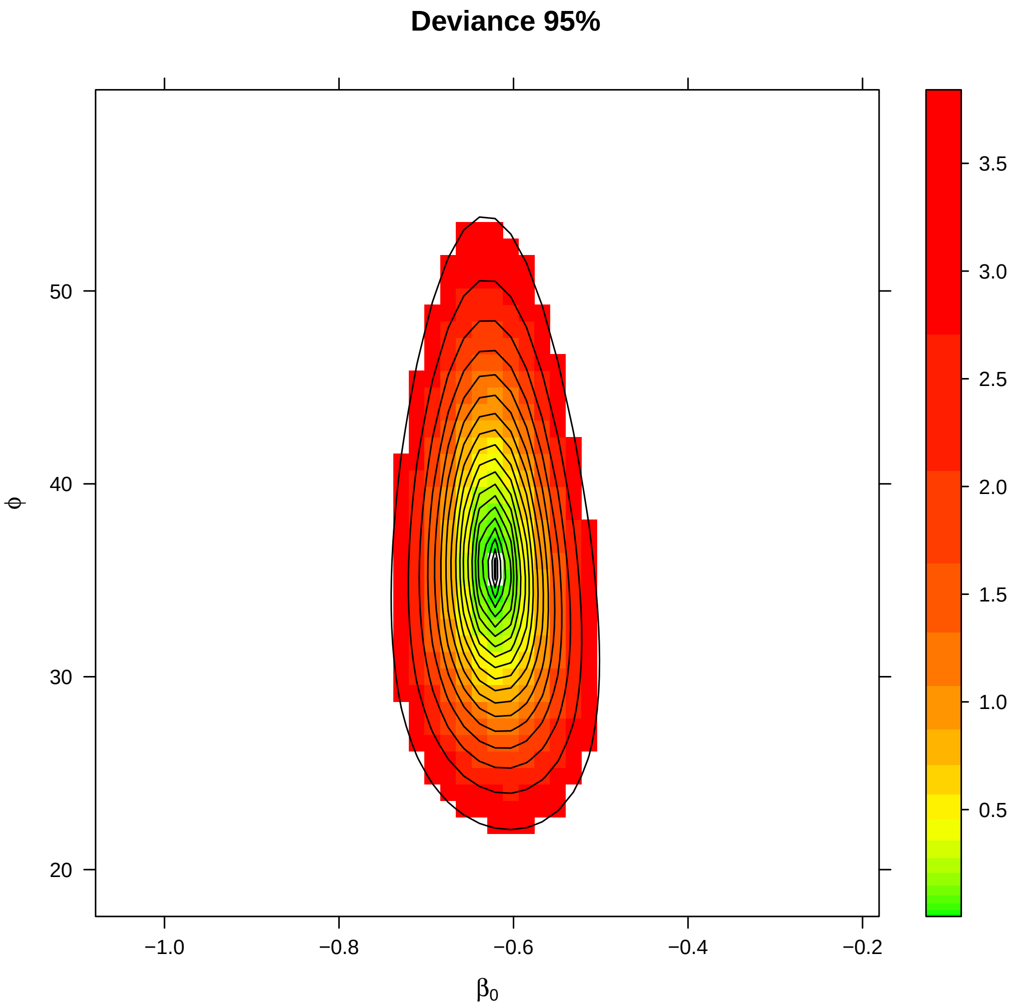

Figure 3: Deviance conjunta β_0 e ϕ

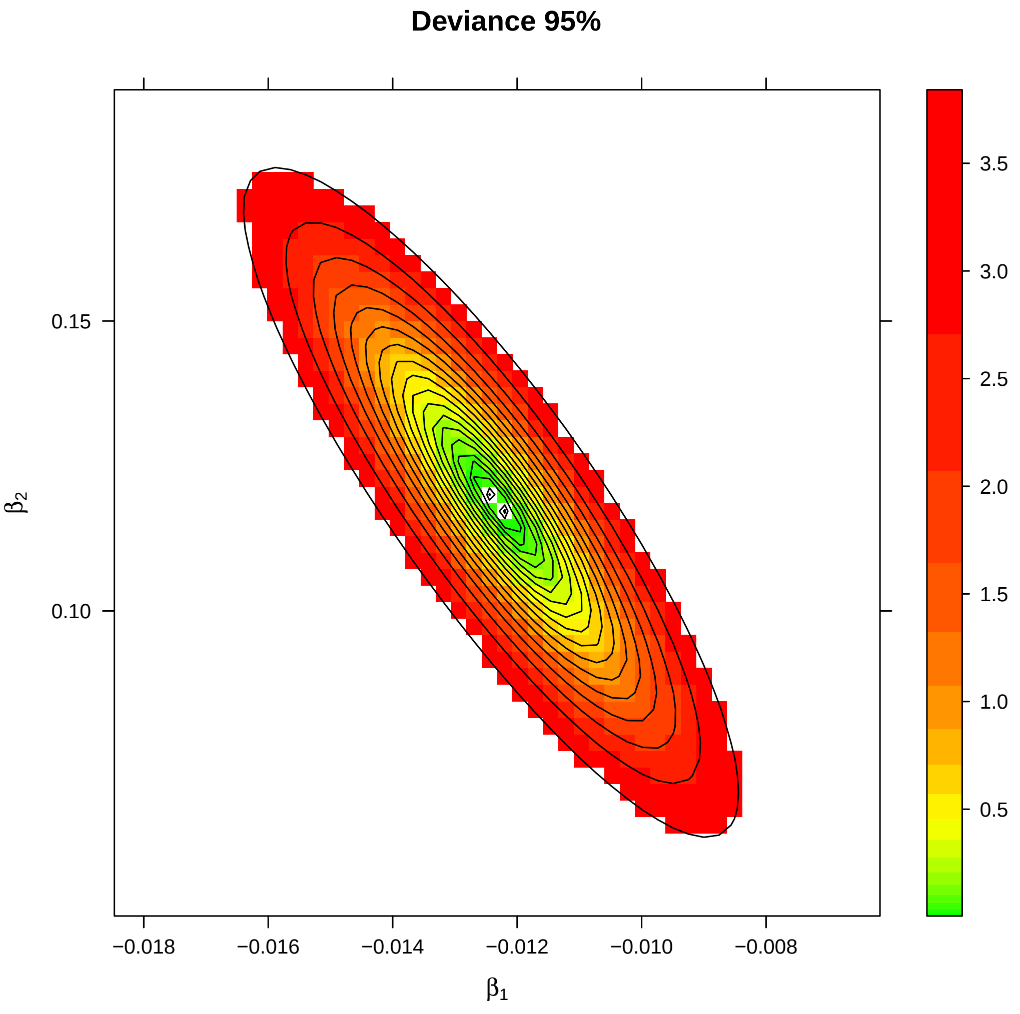

Figure 4: Deviance conjunta β_0 e β_2

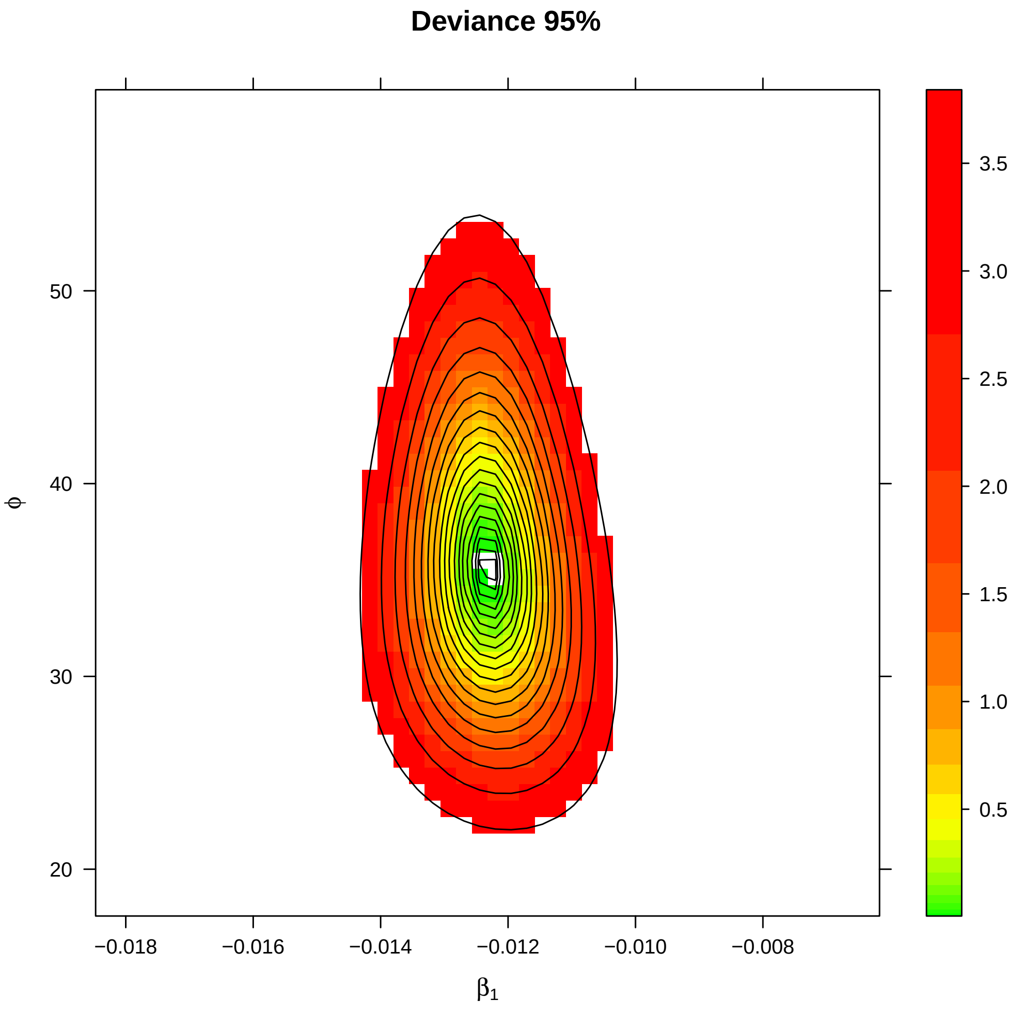

Figure 5: Deviance conjunta β_1 e ϕ

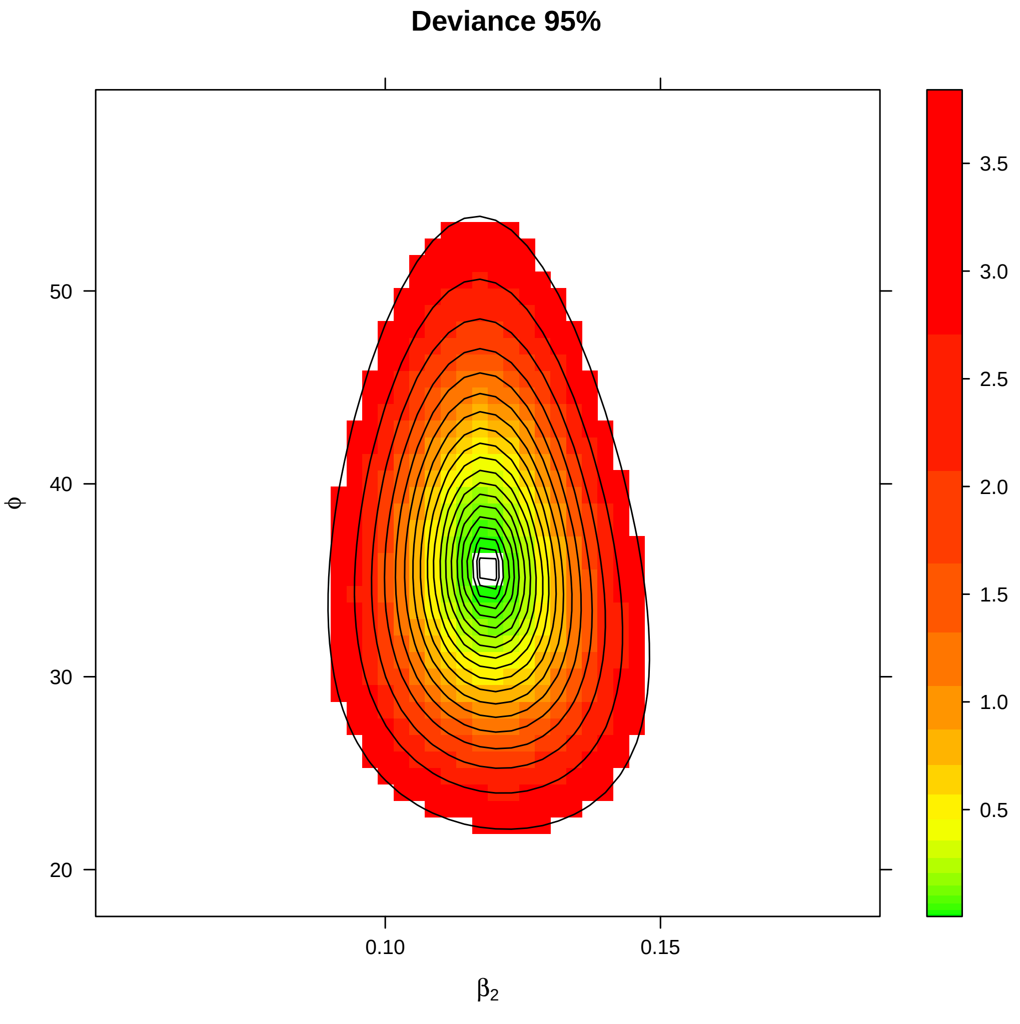

Figure 6: Deviance conjunta β_2 e ϕ

4.5 Verossimilhança perfilhada para ϕ

```
[1] 22.21106 53.76884
```

```
[1] 22.21106 53.76884
```

```
                          B0           B1         B2      phi
Limite_inferior   -1.0548255 -0.018349974 0.04881157 19.90357
Estimador_pontual -0.6222526 -0.012321389 0.11863161 35.85834
Limite_superior   -0.1896796 -0.006292805 0.18845166 51.81311
```

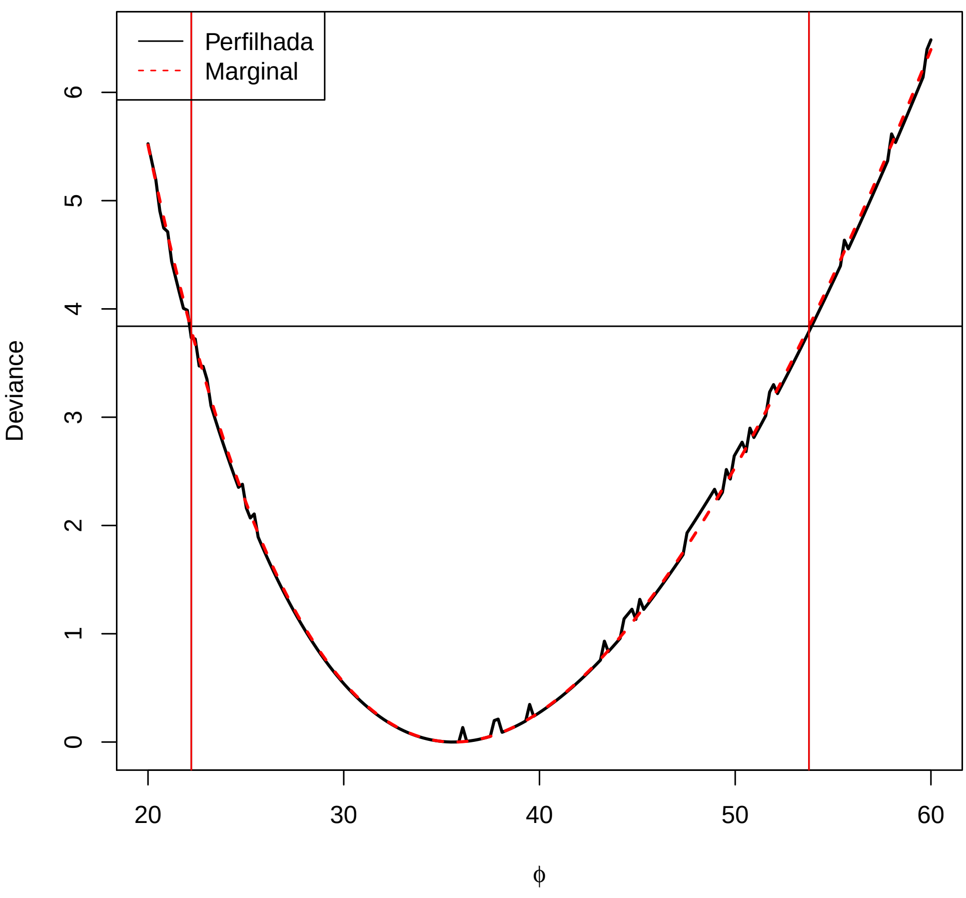

Figure 7: Deviance conjunta β_2 e ϕ

4.6 Ajustando com o mle()

```
Maximum likelihood estimation

Call:
mle2(minuslogl = log.vero1, start = list(B0 = -0.5, B1 = -0.51,
    B2 = 0.11, phi = 35), data = data.frame(y = FoodExpenditure$food/FoodExpenditure$income,
    x1 = FoodExpenditure$income, x2 = FoodExpenditure$persons))

Coefficients:
      Estimate Std. Error z value     Pr(z)
B0  -0.6215900  0.2222179 -2.7972 0.0051546 **
B1  -0.0123004  0.0030957 -3.9734 7.086e-05 ***
B2   0.1182984  0.0358738  3.2976 0.0009751 ***
phi 35.3196474  8.0151774  4.4066 1.050e-05 ***
---
Signif. codes:  0 Ś***Š 0.001 Ś**Š 0.01 Ś*Š 0.05 Ś.Š 0.1 Ś Š 1

-2 log L: -90.66571
```

```
          2.5 %       97.5 %
B0  -1.06908032 -0.179638653
B1  -0.01849124 -0.006104434
B2   0.04651763  0.190087921
phi 22.04497531 53.914037686
```

```
            B0          B1          B2         phi
B0   1.00000000 -0.75026741 -0.56387257 -0.01928734
B1  -0.75026741  1.00000000 -0.05214355 -0.04237211
B2  -0.56387257 -0.05214355  1.00000000  0.03675506
phi -0.01928734 -0.04237211  0.03675506  1.00000000
```

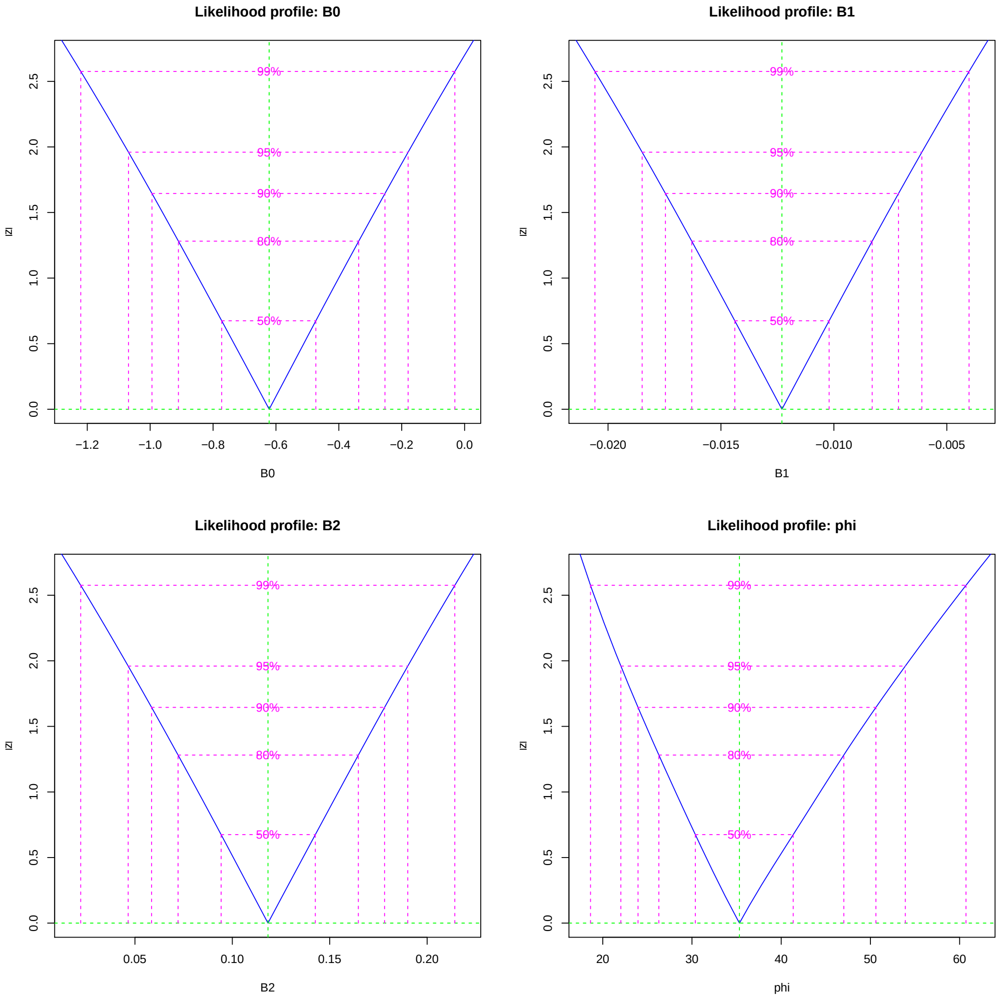

Figure 8: Parametros Perfilhados