

Modelos de Regressão Linear Mistos para dados discretos: Uma abordagem utilizando MCMC através do Stan integrado ao R.

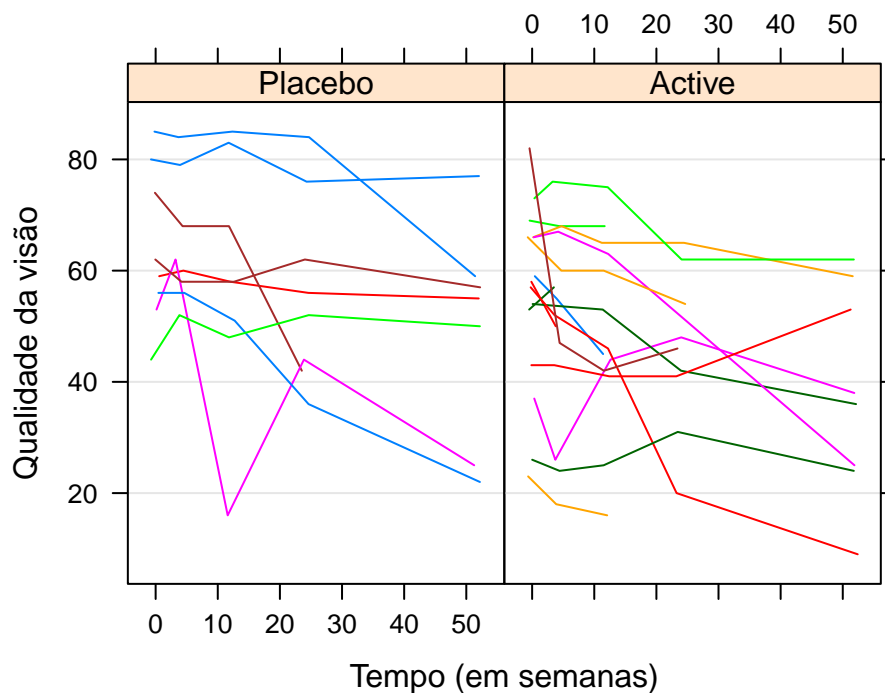
Guilherme Artoni - RA160318

Citizen4 - RAXXXXXXX

Student - RAXXXXXXX

Malba Tahan - RAXXXXXXX

The Stan project develops a probabilistic programming language that implements full Bayesian statistical inference via Markov Chain Monte Carlo, rough Bayesian inference via ‘variational’ approximation, and (optionally penalized) maximum likelihood estimation via optimization. In all three cases, automatic differentiation is used to quickly and accurately evaluate gradients without burdening the user with the need to derive the partial derivatives.

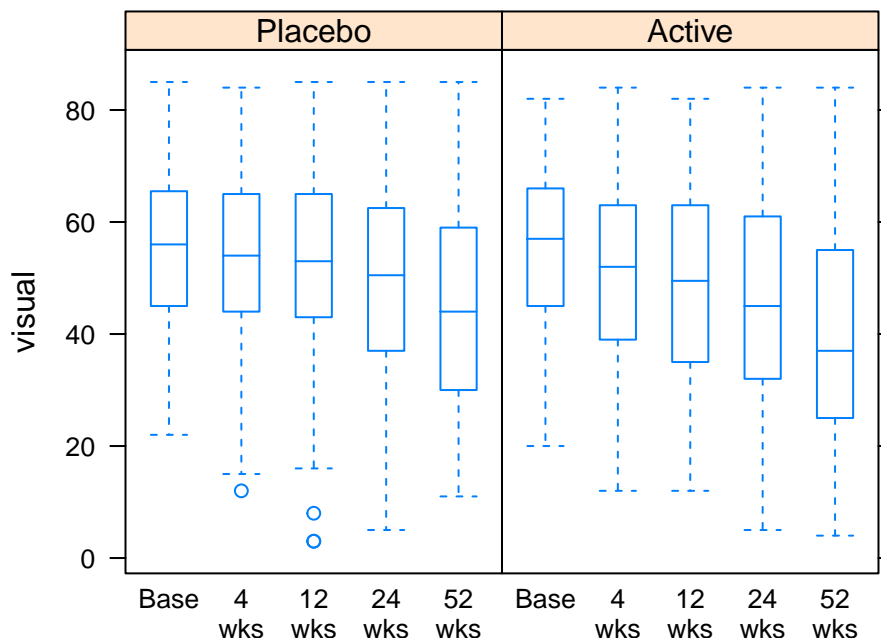


```
##           Placebo Active
## Baseline    119    121
## 4wks         117    114
## 12wks        117    110
## 24wks        112    102
## 52wks        105     90

## [1] "Placebo" "Active"  "Placebo" "Active"  "Placebo" "Active"

##           P:n A:n    P:Mean    A:Mean P:Mdn A:Mdn
## Baseline 119 121 55.33613 54.57851 56.0 57.0
```

```
## 4wks      117 114 53.96581 50.91228 54.0 52.0
## 12wks     117 110 52.87179 48.67273 53.0 49.5
## 24wks     112 102 49.33036 45.46078 50.5 45.0
## 52wks     105 90 44.43810 39.10000 44.0 37.0
```



```
##          visual0  visual4  visual12  visual24  visual52
## visual0  220.3055 206.7096 196.2439 193.3099 152.7141
## visual4   206.7096 246.2204 224.7933 221.2677 179.2284
## visual12  196.2439 224.7933 286.2072 257.7738 222.6830
## visual24  193.3099 221.2677 257.7738 334.4456 285.2327
## visual52  152.7141 179.2284 222.6830 285.2327 347.4311
```

```
##          visual0  visual4  visual12  visual24  visual52
## visual0         1.00    0.89    0.78    0.71    0.55
## visual4         0.89    1.00    0.85    0.77    0.61
## visual12        0.78    0.85    1.00    0.83    0.71
## visual24        0.71    0.77    0.83    1.00    0.84
## visual52        0.55    0.61    0.71    0.84    1.00
```

1 MODELO NORMAL INDEPENDENTE HOMOCEDASTICO

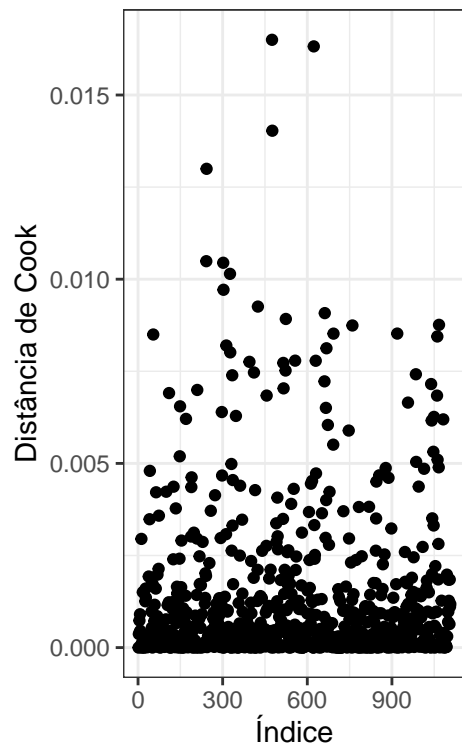
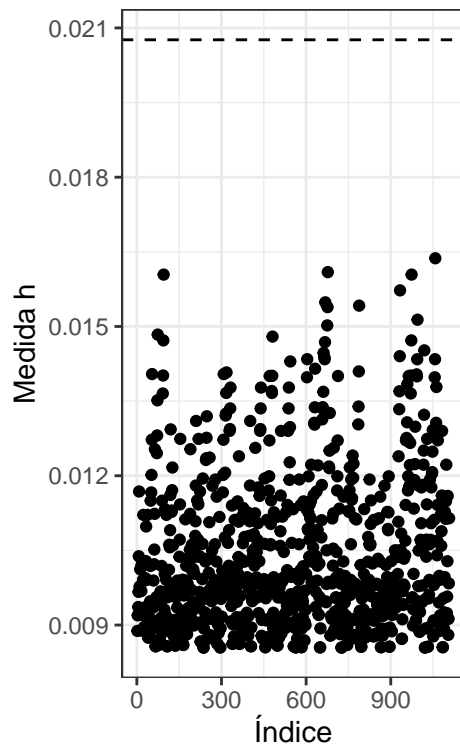
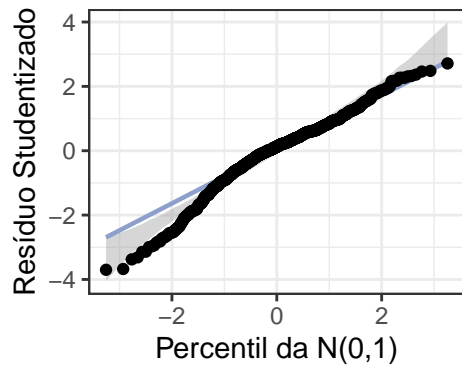
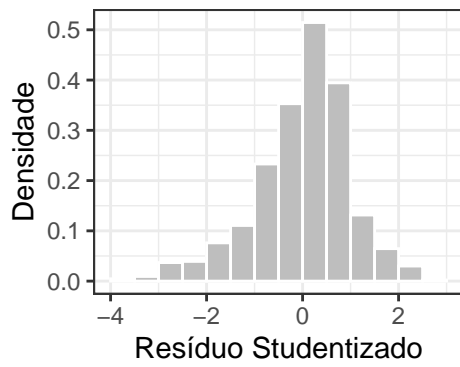
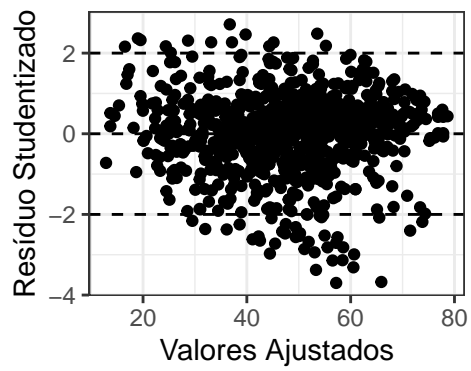
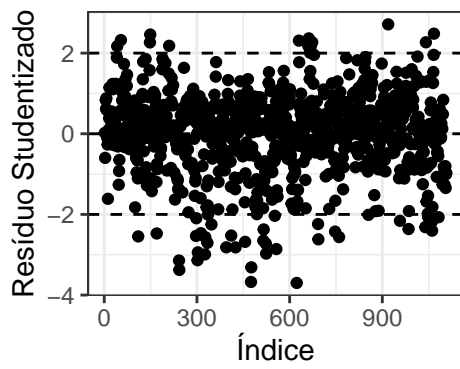
```
##
## Call:
## lm(formula = lm.form, data = armd)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -45.210  -6.459   1.532   7.512  33.283
```

```
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## visual0          0.83037    0.02842  29.213 < 2e-16 ***
## time.f4wks        8.07531    1.94341   4.155 3.58e-05 ***
## time.f12wks       7.08066    1.94066   3.649 0.00028 ***
## time.f24wks       3.63022    1.95316   1.859 0.06342 .
## time.f52wks      -1.74643    1.98952  -0.878 0.38029
## time.f4wks:treat.fActive -2.35278    1.62894  -1.444 0.14900
## time.f12wks:treat.fActive -3.70852    1.64378  -2.256 0.02432 *
## time.f24wks:treat.fActive -3.44915    1.69399  -2.036 0.04205 *
## time.f52wks:treat.fActive -4.47345    1.77811  -2.516 0.01206 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.38 on 858 degrees of freedom
## Multiple R-squared:  0.9432, Adjusted R-squared:  0.9426
## F-statistic: 1583 on 9 and 858 DF, p-value: < 2.2e-16

## [1] 12.37649

##               2.5 %      97.5 %
## visual0          0.7745832  0.8861617
## time.f4wks        4.2609239 11.8897036
## time.f12wks       3.2716615 10.8896534
## time.f24wks      -0.2033236  7.4637556
## time.f52wks      -5.6513208  2.1584611
## time.f4wks:treat.fActive -5.5499518  0.8443996
## time.f12wks:treat.fActive -6.9348245 -0.4822195
## time.f24wks:treat.fActive -6.7740045 -0.1243018
## time.f52wks:treat.fActive -7.9634126 -0.9834943

## Analysis of Variance Table
##
## Response: visual
##               Df Sum Sq Mean Sq    F value    Pr(>F)
## visual0         1 2165776 2165776 14138.9886 < 2.2e-16 ***
## time.f           4   14434    3608   23.5574 < 2.2e-16 ***
## time.f:treat.f   4    2703     676    4.4109 0.001555 **
## Residuals      858  131426     153
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



2 MODELO MISTO FREQUENTISTA

```
## Linear mixed-effects model fit by REML
## Data: armd
## Log-restricted-likelihood: -3288.986
## Fixed: lme.form
##      (Intercept)          visual0          time          treat.fActive
##      9.28807837          0.82643987         -0.21221595         -2.42200012
## time:treat.fActive
##      -0.04959058
##
## Random effects:
## Formula: ~1 | subject
##      (Intercept) Residual
## StdDev:      8.978212 8.627514
##
## Number of Observations: 867
## Number of Groups: 234

##      Value Std.Error DF t-value p-value
## (Intercept)      9.288078  2.681889 631.000000  3.4633 0.0005698 ***
## visual0          0.826440  0.044667 231.000000 18.5022 < 2.2e-16 ***
## time            -0.212216  0.022929 631.000000 -9.2552 < 2.2e-16 ***
## treat.fActive    -2.422000  1.499967 231.000000 -1.6147 0.1077402
## time:treat.fActive -0.049591  0.033562 631.000000 -1.4776 0.1400155
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## subject 2
## Conditional variance covariance matrix
##      1      2      3      4
## 1 74.434 0.000 0.000 0.000
## 2 0.000 74.434 0.000 0.000
## 3 0.000 0.000 74.434 0.000
## 4 0.000 0.000 0.000 74.434
## Standard Deviations: 8.6275 8.6275 8.6275 8.6275

##      1      2      3      4
## 1 1.0000000 0.5199116 0.5199116 0.5199116
## 2 0.5199116 1.0000000 0.5199116 0.5199116
## 3 0.5199116 0.5199116 1.0000000 0.5199116
## 4 0.5199116 0.5199116 0.5199116 1.0000000

##      AIC      BIC      AICc      SABIC      HQCIC -2log.lik
## 6833.788 6881.438 6833.998 6846.091 6848.200 6813.788

##      AIC      BIC      AICc      SABIC      HQCIC -2log.lik
## 6591.971 6625.286 6592.026 6597.137 6597.444 6577.971
```

3 MODELO MISTO BAYESIANO

Segundo Manly (1997) a ideia básica por trás da Inferência Bayesiana é mudar as probabilidades para os parâmetros tomando valores numéricos particulares para novas probabilidades como um resultado da coleta de mais dados, com essa mudança sendo alcançada através do Teorema de Bayes. Como um exemplo da abordagem Bayesiana, suponha que temos interesse no valor de um parâmetro θ de uma determinada população, e que antes de qualquer informação ser observada é de alguma forma possível afirmar que θ deve assumir um dos valores entre $\theta_1, \theta_2, \dots, \theta_n$ e que a probabilidade de o valor ser θ_i é $\pi(\theta_i)$. Suponha também que alguns dados novos são coletados e a probabilidade de observar estes dados é $\pi(dados|\theta_i)$ se de fato $\theta = \theta_i$

```
##
## SAMPLING FOR MODEL 'matrixModel' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 0.000595 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 5.95 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:    1 / 2000 [  0%] (Warmup)
## Chain 1: Iteration:   200 / 2000 [ 10%] (Warmup)
## Chain 1: Iteration:   400 / 2000 [ 20%] (Warmup)
## Chain 1: Iteration:   600 / 2000 [ 30%] (Warmup)
## Chain 1: Iteration:   800 / 2000 [ 40%] (Warmup)
## Chain 1: Iteration:  1000 / 2000 [ 50%] (Warmup)
## Chain 1: Iteration: 1001 / 2000 [ 50%] (Sampling)
## Chain 1: Iteration:  1200 / 2000 [ 60%] (Sampling)
## Chain 1: Iteration:  1400 / 2000 [ 70%] (Sampling)
## Chain 1: Iteration:  1600 / 2000 [ 80%] (Sampling)
## Chain 1: Iteration:  1800 / 2000 [ 90%] (Sampling)
## Chain 1: Iteration:  2000 / 2000 [100%] (Sampling)
## Chain 1:
## Chain 1: Elapsed Time: 30.2139 seconds (Warm-up)
## Chain 1:           16.0806 seconds (Sampling)
## Chain 1:           46.2946 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'matrixModel' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 0.000438 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 4.38 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:    1 / 2000 [  0%] (Warmup)
## Chain 2: Iteration:   200 / 2000 [ 10%] (Warmup)
## Chain 2: Iteration:   400 / 2000 [ 20%] (Warmup)
## Chain 2: Iteration:   600 / 2000 [ 30%] (Warmup)
## Chain 2: Iteration:   800 / 2000 [ 40%] (Warmup)
## Chain 2: Iteration:  1000 / 2000 [ 50%] (Warmup)
## Chain 2: Iteration: 1001 / 2000 [ 50%] (Sampling)
## Chain 2: Iteration:  1200 / 2000 [ 60%] (Sampling)
```

```

## Chain 2: Iteration: 1400 / 2000 [ 70%] (Sampling)
## Chain 2: Iteration: 1600 / 2000 [ 80%] (Sampling)
## Chain 2: Iteration: 1800 / 2000 [ 90%] (Sampling)
## Chain 2: Iteration: 2000 / 2000 [100%] (Sampling)
## Chain 2:
## Chain 2: Elapsed Time: 30.4907 seconds (Warm-up)
## Chain 2: 15.9974 seconds (Sampling)
## Chain 2: 46.4881 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'matrixModel' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 0.000988 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 9.88 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration: 1 / 2000 [ 0%] (Warmup)
## Chain 3: Iteration: 200 / 2000 [ 10%] (Warmup)
## Chain 3: Iteration: 400 / 2000 [ 20%] (Warmup)
## Chain 3: Iteration: 600 / 2000 [ 30%] (Warmup)
## Chain 3: Iteration: 800 / 2000 [ 40%] (Warmup)
## Chain 3: Iteration: 1000 / 2000 [ 50%] (Warmup)
## Chain 3: Iteration: 1001 / 2000 [ 50%] (Sampling)
## Chain 3: Iteration: 1200 / 2000 [ 60%] (Sampling)
## Chain 3: Iteration: 1400 / 2000 [ 70%] (Sampling)
## Chain 3: Iteration: 1600 / 2000 [ 80%] (Sampling)
## Chain 3: Iteration: 1800 / 2000 [ 90%] (Sampling)
## Chain 3: Iteration: 2000 / 2000 [100%] (Sampling)
## Chain 3:
## Chain 3: Elapsed Time: 28.8419 seconds (Warm-up)
## Chain 3: 15.6728 seconds (Sampling)
## Chain 3: 44.5147 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL 'matrixModel' NOW (CHAIN 4).
## Chain 4:
## Chain 4: Gradient evaluation took 0.000217 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 2.17 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration: 1 / 2000 [ 0%] (Warmup)
## Chain 4: Iteration: 200 / 2000 [ 10%] (Warmup)
## Chain 4: Iteration: 400 / 2000 [ 20%] (Warmup)
## Chain 4: Iteration: 600 / 2000 [ 30%] (Warmup)
## Chain 4: Iteration: 800 / 2000 [ 40%] (Warmup)
## Chain 4: Iteration: 1000 / 2000 [ 50%] (Warmup)
## Chain 4: Iteration: 1001 / 2000 [ 50%] (Sampling)

```

```
## Chain 4: Iteration: 1200 / 2000 [ 60%] (Sampling)
## Chain 4: Iteration: 1400 / 2000 [ 70%] (Sampling)
## Chain 4: Iteration: 1600 / 2000 [ 80%] (Sampling)
## Chain 4: Iteration: 1800 / 2000 [ 90%] (Sampling)
## Chain 4: Iteration: 2000 / 2000 [100%] (Sampling)
## Chain 4:
## Chain 4: Elapsed Time: 30.8203 seconds (Warm-up)
## Chain 4:           15.6548 seconds (Sampling)
## Chain 4:           46.4751 seconds (Total)
## Chain 4:
```

Manly, Bryan F. J. 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall.