2 - Regress?o Log?stica em R com Stepwise

May 6, 2019

0.1 Biblioteca

```
In [10]: # PACOTES NECESSÁRIOS

         # install.packages("boot")
         # install.packages("ggplot2")
         # install.packages("caret")
         # install.packages("gmodels")
         # install.packages("hmeasure")
         # install.packages("pROC")

         library(gmodels)
         library(boot)
         library(hmeasure)
         library(pROC)
```

0.2 Dados

```
In [12]: dados.treino <- read.csv("porto_seguro_limpo_treino.csv")
         dados.teste <- read.csv("porto_seguro_limpo_teste.csv")
```

0.3 Regressão Logística

```
In [13]: fit1=glm(data = dados.treino, target~ps_ind_01+ps_ind_03+ps_ind_06_bin+ps_ind_07_bin+
         summary(fit1)


Call:
glm(formula = target ~ ps_ind_01 + ps_ind_03 + ps_ind_06_bin +
    ps_ind_07_bin + ps_ind_08_bin + ps_ind_09_bin + ps_ind_12_bin +
    ps_ind_14 + ps_ind_15 + ps_ind_16_bin + ps_ind_17_bin + ps_ind_18_bin +
    ps_reg_01 + ps_reg_02 + ps_car_08_cat + ps_car_11 + ps_car_15 +
    ps_calc_01 + ps_calc_02 + ps_calc_03 + ps_calc_04 + ps_calc_05 +
    ps_calc_06 + ps_calc_07 + ps_calc_08 + ps_calc_09 + ps_calc_10 +
    ps_calc_11 + ps_calc_12 + ps_calc_13 + ps_calc_14 + ps_calc_15_bin +
    ps_calc_16_bin + ps_calc_17_bin + ps_calc_18_bin + ps_calc_19_bin +
    ps_calc_20_bin + ps_car_07_cat_null + ps_car_07_cat_1 + ps_car_05_cat_null +
    ps_car_05_cat_1 + ps_car_03_cat_null + ps_car_03_cat_1 +
```

```
    ps_car_02_cat_1 + ps_ind_04_cat_1 + ps_ind_02_cat_null +
    ps_ind_02_cat_2_3_4 + ps_ind_05_cat_0 + ps_ind_05_cat_1_3_4_5_6 +
    ps_ind_05_cat_2 + ps_car_01_cat_6_7 + ps_car_01_cat_3_4_5_10 +
    ps_car_01_cat_0_1_2_8_11 + ps_car_01_cat_9 + ps_car_04_cat_0_4 +
    ps_car_04_cat_1_2 + ps_car_04_cat_3_8 + ps_car_04_cat_6_9 +
    ps_car_06_cat_0_1_3_4_6_7_11_14 + ps_car_06_cat_10_12_15_16 +
    ps_car_06_cat_2_5_8_17 + ps_car_09_cat_0_2_3 + ps_car_09_cat_1_4 +
    ps_car_11_cat_A + ps_car_11_cat_B + ps_car_11_cat_C + ps_reg_03_no_out +
    ps_car_12_no_out + ps_car_13_no_out + ps_car_14_no_out, family = binomial(),
    data = dados.treino)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.3018  -0.2949  -0.2510  -0.2160   2.9879

Coefficients: (1 not defined because of singularities)
                               Estimate Std. Error z value Pr(>|z|)
(Intercept)                  -1.2620372  0.5630182  -2.242 0.024990 *
ps_ind_01                     0.0108361  0.0053194   2.037 0.041641 *
ps_ind_03                     0.0231360  0.0039247   5.895 3.75e-09 ***
ps_ind_06_bin                -0.0601917  0.0307845  -1.955 0.050552 .
ps_ind_07_bin                 0.2199380  0.0314285   6.998 2.60e-12 ***
ps_ind_08_bin                 0.1832236  0.0331771   5.523 3.34e-08 ***
ps_ind_09_bin                        NA         NA      NA       NA
ps_ind_12_bin                 0.0477556  0.1750686   0.273 0.785020
ps_ind_14                    -0.0129034  0.1315490  -0.098 0.921863
ps_ind_15                    -0.0270594  0.0030906  -8.755  < 2e-16 ***
ps_ind_16_bin                -0.0991407  0.0398582  -2.487 0.012870 *
ps_ind_17_bin                 0.2633460  0.0447906   5.879 4.12e-09 ***
ps_ind_18_bin                -0.0533171  0.0469793  -1.135 0.256415
ps_reg_01                     0.2839617  0.0420040   6.760 1.38e-11 ***
ps_reg_02                     0.0660732  0.0362115   1.825 0.068054 .
ps_car_08_cat                 0.0086981  0.0294590   0.295 0.767793
ps_car_11                    -0.0340714  0.0136661  -2.493 0.012662 *
ps_car_15                     0.0486911  0.0238613   2.041 0.041292 *
ps_calc_01                    0.0495885  0.0328048   1.512 0.130629
ps_calc_02                    0.0533068  0.0327967   1.625 0.104083
ps_calc_03                    0.0323703  0.0327710   0.988 0.323263
ps_calc_04                    0.0047886  0.0084307   0.568 0.570041
ps_calc_05                    0.0086510  0.0082741   1.046 0.295766
ps_calc_06                    0.0066857  0.0070824   0.944 0.345181
ps_calc_07                    0.0034187  0.0066482   0.514 0.607095
ps_calc_08                   -0.0088200  0.0064410  -1.369 0.170886
ps_calc_09                    0.0033372  0.0075509   0.442 0.658520
ps_calc_10                   -0.0019192  0.0032439  -0.592 0.554106
ps_calc_11                    0.0003161  0.0040297   0.078 0.937468
ps_calc_12                    0.0025493  0.0078257   0.326 0.744603
ps_calc_13                   -0.0016366  0.0055582  -0.294 0.768412
```

```
ps_calc_14                          0.0023972  0.0034235   0.700 0.483803
ps_calc_15_bin                     -0.0135597  0.0288637  -0.470 0.638511
ps_calc_16_bin                     -0.0047453  0.0194603  -0.244 0.807351
ps_calc_17_bin                     -0.0269042  0.0189210  -1.422 0.155049
ps_calc_18_bin                     -0.0003948  0.0207950  -0.019 0.984853
ps_calc_19_bin                     -0.0110038  0.0197979  -0.556 0.578343
ps_calc_20_bin                     -0.0176357  0.0262754  -0.671 0.502103
ps_car_07_cat_null                 -0.2527938  0.0877471  -2.881 0.003965 **
ps_car_07_cat_1                    -0.2299564  0.0391682  -5.871 4.33e-09 ***
ps_car_05_cat_null                  0.0386755  0.0328151   1.179 0.238563
ps_car_05_cat_1                     0.0303818  0.0253100   1.200 0.229989
ps_car_03_cat_null                 -0.0619009  0.0461809  -1.340 0.180115
ps_car_03_cat_1                     0.0810402  0.0345466   2.346 0.018985 *
ps_car_02_cat_1                     0.0103849  0.0304706   0.341 0.733242
ps_ind_04_cat_1                     0.0478639  0.0215409   2.222 0.026283 *
ps_ind_02_cat_null                  0.6836904  0.4030278   1.696 0.089813 .
ps_ind_02_cat_2_3_4                 0.0783489  0.0225918   3.468 0.000524 ***
ps_ind_05_cat_0                    -0.8321385  0.1034962  -8.040 8.96e-16 ***
ps_ind_05_cat_1_3_4_5_6            -0.3912865  0.1062517  -3.683 0.000231 ***
ps_ind_05_cat_2                    -0.0597311  0.1315028  -0.454 0.649671
ps_car_01_cat_6_7                  -1.4612785  0.4813467  -3.036 0.002399 **
ps_car_01_cat_3_4_5_10             -1.2828634  0.4814646  -2.665 0.007710 **
ps_car_01_cat_0_1_2_8_11           -1.3013753  0.4812676  -2.704 0.006850 **
ps_car_01_cat_9                    -1.1262505  0.4827297  -2.333 0.019644 *
ps_car_04_cat_0_4                  -0.1872128  0.2015961  -0.929 0.353069
ps_car_04_cat_1_2                  -0.2095091  0.1994766  -1.050 0.293583
ps_car_04_cat_3_8                  -0.3025040  0.2032587  -1.488 0.136679
ps_car_04_cat_6_9                  -0.2430141  0.2015039  -1.206 0.227817
ps_car_06_cat_0_1_3_4_6_7_11_14 -0.0363701  0.0488643  -0.744 0.456690
ps_car_06_cat_10_12_15_16           0.0090601  0.0522475   0.173 0.862332
ps_car_06_cat_2_5_8_17              0.0700389  0.0794996   0.881 0.378320
ps_car_09_cat_0_2_3                -0.2560795  0.2387785  -1.072 0.283515
ps_car_09_cat_1_4                  -0.0455662  0.2405115  -0.189 0.849736
ps_car_11_cat_A                    -0.1368269  0.0279010  -4.904 9.39e-07 ***
ps_car_11_cat_B                     0.0212361  0.0296786   0.716 0.474277
ps_car_11_cat_C                     0.2112941  0.0404172   5.228 1.72e-07 ***
ps_reg_03_no_out                    0.1081263  0.0450490   2.400 0.016386 *
ps_car_12_no_out                    0.3944833  0.3434668   1.149 0.250748
ps_car_13_no_out                    0.4857958  0.1305792   3.720 0.000199 ***
ps_car_14_no_out                   -0.6331301  0.4588025  -1.380 0.167598
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 101574  on 324637  degrees of freedom
Residual deviance:  98916  on 324568  degrees of freedom
  (1 observation deleted due to missingness)
```

```
AIC: 99056

Number of Fisher Scoring iterations: 6
```

0.4 Stepwise p-value < 15%

```
In [ ]: fit2=step(fit1)
```

```
In [15]: summary(fit2)
```

```
Call:
glm(formula = target ~ ps_ind_01 + ps_ind_03 + ps_ind_06_bin +
    ps_ind_07_bin + ps_ind_08_bin + ps_ind_15 + ps_ind_16_bin +
    ps_ind_17_bin + ps_reg_01 + ps_reg_02 + ps_car_11 + ps_calc_01 +
    ps_calc_02 + ps_calc_17_bin + ps_car_07_cat_null + ps_car_07_cat_1 +
    ps_car_03_cat_1 + ps_ind_04_cat_1 + ps_ind_02_cat_null +
    ps_ind_02_cat_2_3_4 + ps_ind_05_cat_0 + ps_ind_05_cat_1_3_4_5_6 +
    ps_car_01_cat_6_7 + ps_car_01_cat_3_4_5_10 + ps_car_01_cat_0_1_2_8_11 +
    ps_car_01_cat_9 + ps_car_04_cat_3_8 + ps_car_09_cat_0_2_3 +
    ps_car_11_cat_A + ps_car_11_cat_C + ps_reg_03_no_out + ps_car_13_no_out +
    ps_car_14_no_out, family = binomial(), data = dados.treino)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2734  -0.2946  -0.2509  -0.2163   2.9673

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)             -1.265263   0.476265  -2.657 0.007892 **
ps_ind_01                0.010797   0.005087   2.122 0.033805 *
ps_ind_03                0.023211   0.003857   6.018 1.76e-09 ***
ps_ind_06_bin           -0.058666   0.030329  -1.934 0.053077 .
ps_ind_07_bin            0.220217   0.031095   7.082 1.42e-12 ***
ps_ind_08_bin            0.186570   0.032929   5.666 1.46e-08 ***
ps_ind_15               -0.025614   0.002903  -8.822  < 2e-16 ***
ps_ind_16_bin           -0.062434   0.025139  -2.484 0.013009 *
ps_ind_17_bin            0.304083   0.031099   9.778  < 2e-16 ***
ps_reg_01                0.284107   0.041901   6.780 1.20e-11 ***
ps_reg_02                0.069882   0.036085   1.937 0.052796 .
ps_car_11               -0.033181   0.011842  -2.802 0.005078 **
ps_calc_01               0.049595   0.032803   1.512 0.130553
ps_calc_02               0.053513   0.032793   1.632 0.102718
ps_calc_17_bin          -0.026981   0.018918  -1.426 0.153823
ps_car_07_cat_null      -0.209529   0.071177  -2.944 0.003242 **
ps_car_07_cat_1         -0.230130   0.038883  -5.918 3.25e-09 ***
```

```
ps_car_03_cat_1           0.094502   0.028288   3.341 0.000836 ***
ps_ind_04_cat_1           0.045514   0.021252   2.142 0.032222 *
ps_ind_02_cat_null        0.693621   0.399509   1.736 0.082531 .
ps_ind_02_cat_2_3_4       0.077975   0.022418   3.478 0.000505 ***
ps_ind_05_cat_0          -0.790578   0.063400 -12.470  < 2e-16 ***
ps_ind_05_cat_1_3_4_5_6  -0.349745   0.067614  -5.173 2.31e-07 ***
ps_car_01_cat_6_7        -1.509322   0.460871  -3.275 0.001057 **
ps_car_01_cat_3_4_5_10   -1.337838   0.460803  -2.903 0.003693 **
ps_car_01_cat_0_1_2_8_11 -1.343422   0.460786  -2.916 0.003551 **
ps_car_01_cat_9          -1.164354   0.462315  -2.519 0.011785 *
ps_car_04_cat_3_8        -0.092372   0.048346  -1.911 0.056051 .
ps_car_09_cat_0_2_3      -0.218061   0.035831  -6.086 1.16e-09 ***
ps_car_11_cat_A          -0.152848   0.025853  -5.912 3.38e-09 ***
ps_car_11_cat_C           0.196728   0.032300   6.091 1.12e-09 ***
ps_reg_03_no_out          0.113169   0.044920   2.519 0.011757 *
ps_car_13_no_out          0.683958   0.071440   9.574  < 2e-16 ***
ps_car_14_no_out         -0.998395   0.305143  -3.272 0.001068 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 101574  on 324637  degrees of freedom
Residual deviance:  98938  on 324604  degrees of freedom
  (1 observation deleted due to missingness)
AIC: 99006

Number of Fisher Scoring iterations: 6
```

```
In [17]: # PROB
         dados.teste$pred=predict(fit2, newdata = dados.teste, type = "response")
         head(dados.teste$pred, 5)
```

1. 0.0246787308236698 2. 0.0231731609768565 3. 0.015733950788423 4. 0.0222012030507004 5. 0.0496793524806326

```
In [18]: # CROSS VALIDATION
         cv.glm(data=dados.teste, glmfit=fit2, K=10)$delta[1]
```

```
Warning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
```

```
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object length
```

0.0356786393674526

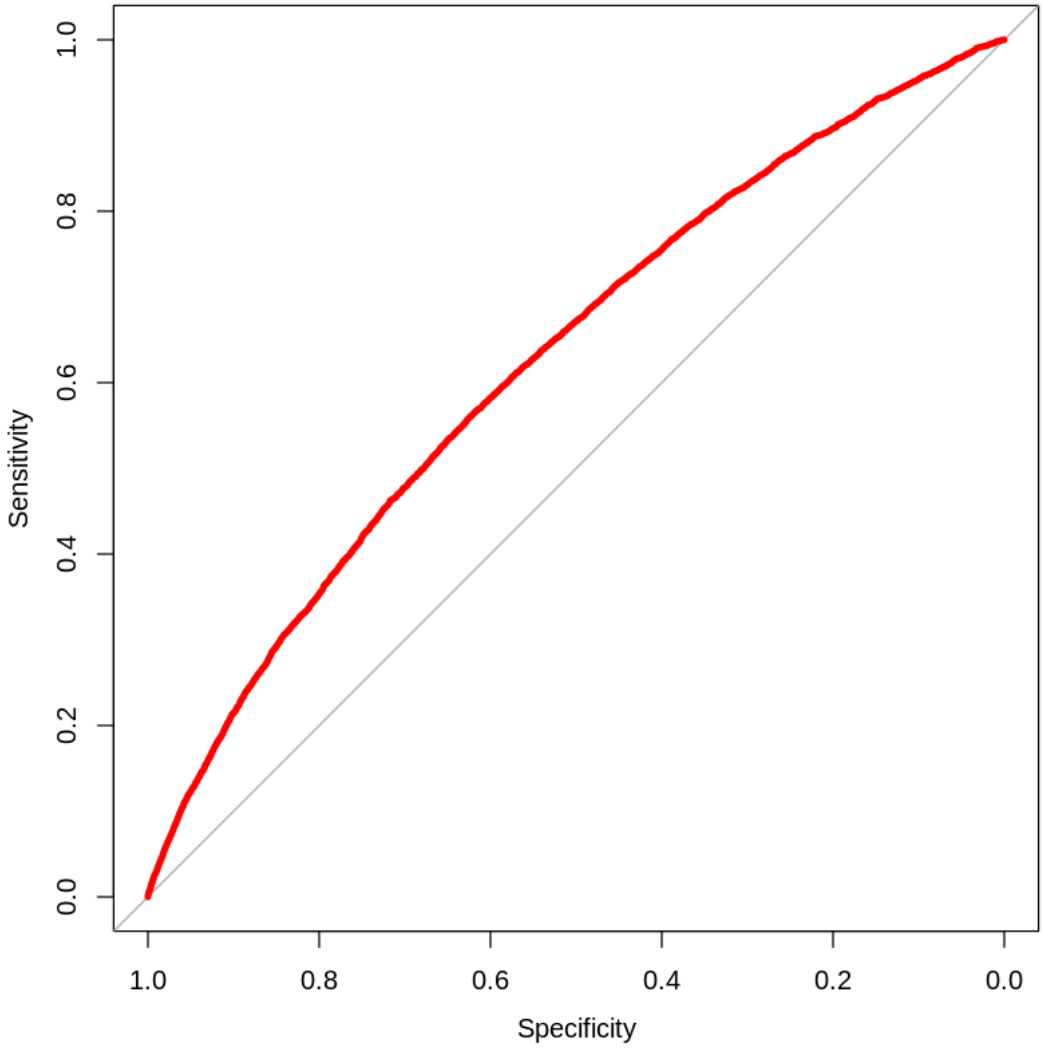
```
In [23]: # ROC E AUC
         medidah=HMeasure(dados.teste$target, dados.teste$pred)
         medidah$metrics
```

	H	Gini	AUC	AUCH	KS	MER	MWL	Spec.Sens95
scores	0.05026603	0.2530683	0.6265341	0.628246	0.184126	0.03593109	0.0565408	0.1075657

```
In [20]: # CURVA ROC
         roc1 = roc(dados.teste$target, dados.teste$pred)
         plot(roc1, lwd=4, col="red")
         roc1
```

```
Call:
roc.default(response = dados.teste$target, predictor = dados.teste$pred)

Data: dados.teste$pred in 172146 controls (dados.teste$target 0) < 6418 cases (dados.teste$targ
Area under the curve: 0.6265
```



0.5 Stepwise p-value<50%

```
In [ ]: fit3=step(fit1, k=0.45)
```

```
In [25]: summary(fit3)
```

```
Call:
glm(formula = target ~ ps_ind_01 + ps_ind_03 + ps_ind_06_bin +
    ps_ind_07_bin + ps_ind_08_bin + ps_ind_15 + ps_ind_16_bin +
    ps_ind_17_bin + ps_ind_18_bin + ps_reg_01 + ps_reg_02 + ps_car_11 +
    ps_car_15 + ps_calc_01 + ps_calc_02 + ps_calc_03 + ps_calc_05 +
    ps_calc_06 + ps_calc_08 + ps_calc_14 + ps_calc_17_bin + ps_calc_20_bin +
```

```
    ps_car_07_cat_null + ps_car_07_cat_1 + ps_car_05_cat_null +
    ps_car_05_cat_1 + ps_car_03_cat_null + ps_car_03_cat_1 +
    ps_ind_04_cat_1 + ps_ind_02_cat_null + ps_ind_02_cat_2_3_4 +
    ps_ind_05_cat_0 + ps_ind_05_cat_1_3_4_5_6 + ps_car_01_cat_6_7 +
    ps_car_01_cat_3_4_5_10 + ps_car_01_cat_0_1_2_8_11 + ps_car_01_cat_9 +
    ps_car_04_cat_0_4 + ps_car_04_cat_1_2 + ps_car_04_cat_3_8 +
    ps_car_04_cat_6_9 + ps_car_06_cat_0_1_3_4_6_7_11_14 + ps_car_06_cat_2_5_8_17 +
    ps_car_09_cat_0_2_3 + ps_car_11_cat_A + ps_car_11_cat_B +
    ps_car_11_cat_C + ps_reg_03_no_out + ps_car_12_no_out + ps_car_13_no_out +
    ps_car_14_no_out, family = binomial(), data = dados.treino)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2910  -0.2948  -0.2510  -0.2161   2.9855

Coefficients:
                             Estimate Std. Error z value Pr(>|z|)
(Intercept)                -1.260246   0.540211  -2.333 0.019654 *
ps_ind_01                   0.011062   0.005246   2.109 0.034963 *
ps_ind_03                   0.022905   0.003887   5.892 3.81e-09 ***
ps_ind_06_bin              -0.060944   0.030693  -1.986 0.047081 *
ps_ind_07_bin               0.219822   0.031383   7.004 2.48e-12 ***
ps_ind_08_bin               0.183878   0.033154   5.546 2.92e-08 ***
ps_ind_15                  -0.027134   0.003075  -8.823  < 2e-16 ***
ps_ind_16_bin              -0.099147   0.039846  -2.488 0.012837 *
ps_ind_17_bin               0.263182   0.044784   5.877 4.19e-09 ***
ps_ind_18_bin              -0.053011   0.046949  -1.129 0.258851
ps_reg_01                   0.283958   0.041962   6.767 1.31e-11 ***
ps_reg_02                   0.065937   0.036184   1.822 0.068411 .
ps_car_11                  -0.034536   0.013457  -2.566 0.010277 *
ps_car_15                   0.049718   0.023435   2.122 0.033878 *
ps_calc_01                  0.049462   0.032803   1.508 0.131597
ps_calc_02                  0.053244   0.032796   1.623 0.104489
ps_calc_03                  0.032345   0.032769   0.987 0.323620
ps_calc_05                  0.008668   0.008274   1.048 0.294811
ps_calc_06                  0.006674   0.007082   0.942 0.345998
ps_calc_08                 -0.008804   0.006441  -1.367 0.171622
ps_calc_14                  0.002396   0.003424   0.700 0.484066
ps_calc_17_bin             -0.026846   0.018920  -1.419 0.155922
ps_calc_20_bin             -0.017678   0.026275  -0.673 0.501073
ps_car_07_cat_null         -0.235002   0.075314  -3.120 0.001807 **
ps_car_07_cat_1            -0.229703   0.039148  -5.868 4.42e-09 ***
ps_car_05_cat_null          0.038352   0.032719   1.172 0.241127
ps_car_05_cat_1             0.030594   0.025281   1.210 0.226216
ps_car_03_cat_null         -0.063906   0.045674  -1.399 0.161767
ps_car_03_cat_1             0.077379   0.033575   2.305 0.021185 *
ps_ind_04_cat_1             0.047130   0.021508   2.191 0.028434 *
ps_ind_02_cat_null          0.693405   0.399938   1.734 0.082957 .
```

```
ps_ind_02_cat_2_3_4              0.078848   0.022503   3.504 0.000459 ***
ps_ind_05_cat_0                 -0.796674   0.064207 -12.408  < 2e-16 ***
ps_ind_05_cat_1_3_4_5_6         -0.355835   0.068367  -5.205 1.94e-07 ***
ps_car_01_cat_6_7               -1.498419   0.461258  -3.249 0.001160 **
ps_car_01_cat_3_4_5_10          -1.319380   0.461427  -2.859 0.004245 **
ps_car_01_cat_0_1_2_8_11        -1.338904   0.461136  -2.903 0.003690 **
ps_car_01_cat_9                 -1.163682   0.462664  -2.515 0.011897 *
ps_car_04_cat_0_4               -0.187147   0.200166  -0.935 0.349810
ps_car_04_cat_1_2               -0.210200   0.198747  -1.058 0.290226
ps_car_04_cat_3_8               -0.302316   0.203140  -1.488 0.136695
ps_car_04_cat_6_9               -0.242662   0.200813  -1.208 0.226895
ps_car_06_cat_0_1_3_4_6_7_11_14 -0.041047   0.030065  -1.365 0.172169
ps_car_06_cat_2_5_8_17           0.059936   0.069116   0.867 0.385848
ps_car_09_cat_0_2_3             -0.211404   0.036145  -5.849 4.95e-09 ***
ps_car_11_cat_A                 -0.137574   0.027772  -4.954 7.28e-07 ***
ps_car_11_cat_B                  0.020978   0.029651   0.707 0.479258
ps_car_11_cat_C                  0.210385   0.039932   5.269 1.37e-07 ***
ps_reg_03_no_out                 0.107741   0.045021   2.393 0.016705 *
ps_car_12_no_out                 0.401549   0.336788   1.192 0.233148
ps_car_13_no_out                 0.464366   0.117749   3.944 8.02e-05 ***
ps_car_14_no_out                -0.637351   0.456873  -1.395 0.163007
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 101574  on 324637  degrees of freedom
Residual deviance:  98919  on 324586  degrees of freedom
  (1 observation deleted due to missingness)
AIC: 99023

Number of Fisher Scoring iterations: 6
```

```
In [26]: dados.teste$predfit3=predict(fit3, newdata = dados.teste, type = "response")
         head(dados.teste$predfit3, 5)
```

1. 0.0244398194928463 2. 0.0243418388992513 3. 0.0150835588042613 4. 0.0223305931692943 5. 0.0507207442264915

```
In [29]: # CROSS VALIDATION
         cv.glm(data=dados.teste, glmfit=fit3, K=10)$delta[1]
```

```
Warning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
```

```
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object lengthWarning message in y - yhat:
longer object length is not a multiple of shorter object length
```

0.0356910823930923

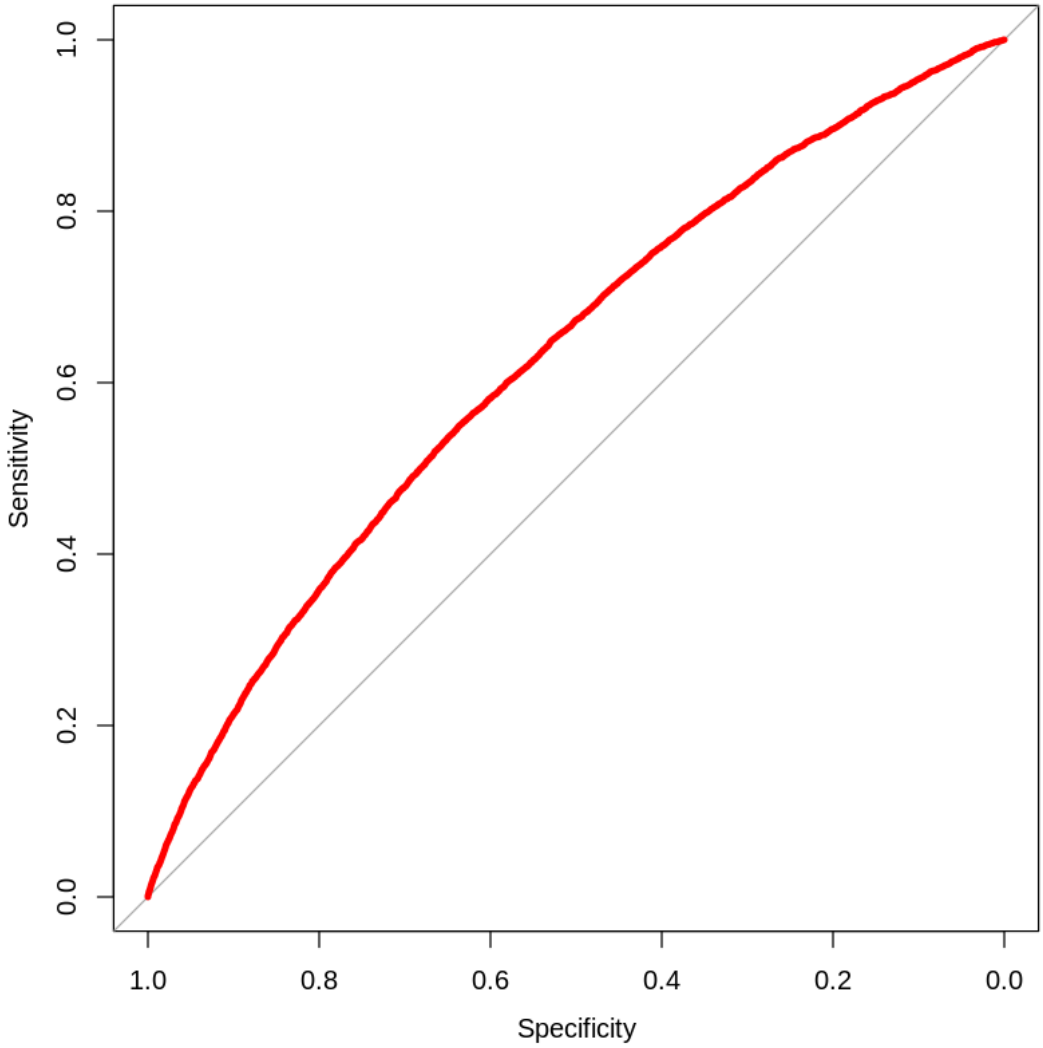
In [28]:
```
# ROC E AUC
medidah=HMeasure(dados.teste$target, dados.teste$predfit3)
medidah$metrics
```

	H	Gini	AUC	AUCH	KS	MER	MWL	Spec.Sens9
scores	0.05066024	0.2544372	0.6272186	0.6288169	0.1863633	0.03591989	0.05638575	0.107159

In [27]:
```
roc2 = roc(dados.teste$target, dados.teste$predfit3)
plot(roc2, lwd=4, col="red")
roc2
```

```
Call:
roc.default(response = dados.teste$target, predictor = dados.teste$predfit3)

Data: dados.teste$predfit3 in 172146 controls (dados.teste$target 0) < 6418 cases (dados.teste$
Area under the curve: 0.6272
```


Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
1.0
0.8
0.6
0.4
0.2
0.0
Specificity