

Laboratorio statistico-informatico A.A 2017/2018

*Bertolini Federico, Bonanno Ezio Alberto, Cavalieri Tommaso, Cesaro Giada,
Malalan Sara*

REGRESSIONE LOGISTICA - PROBABILITA' DI AVER RICEVUTO AIUTO

Carichiamo i dati:

```
dati <- read.table("datiCoppie.csv", header = TRUE, sep = ";")  
head(dati)
```

```
## PROFAM1 ETA SESSO ISTRUZIONE REGIONE REDDITO SALUTE AIUTIDATI  
## 1      10  75     1         10      10       3      3      NA  
## 2      10  74     2         12      10       3      3      NA  
## 3      29  71     2         10      10       3      2       4  
## 4      29  74     1          9      10       3      2       4  
## 5      38  81     1          9      10       3      2      11  
## 6      38  69     2         10      10       3      2      NA  
## AIUTIRICEVUTI FRATELLI FIGLI NIPOTI PARENTI AMICI VICINI NUMNODI  
## 1           NA         1     0     0       1     1     1     1  
## 2           NA         0     0     0       1     1     1     0  
## 3           NA         0     1     2       1     2     3     4  
## 4           NA         0     1     2       1     2     3     4  
## 5           NA         0     1     1       2     2     3     5  
## 6           NA         0     1     1       2     2     3     5  
##          RETE  
## 1 IMMEDIATE  
## 2          0  
## 3 IMM+EXT+NO  
## 4 IMM+EXT+NO  
## 5 IMM+EXT+NO  
## 6 IMM+EXT+NO
```

Codifichiamo le variabili, in particolare:

ETA':

65-69, 70-74, 75-79, 80-84, 85+

SESSO:

M, F

ISTRUZIONE:

ALTO - Dottorato di ricerca o specializzazione post-laurea, Laurea di 4 anni o più, Diploma universitario, Laurea breve vecchio ordinamento, Scuola diretta a fini speciali, Scuola Parau-niversitaria, Accademia Belle Arti, Istituto Superiore Industrie Artistiche, Accademia di Arte Drammatica, Scuola interpreti e Traduttori

MEDIO - Diploma di scuola superiore di 4-5 anni che permette l'iscrizione all'università, Diploma di scuola superiore di 2-3 anni che non permette l'iscrizione all'università, Licenza media

BASSO - Licenza elementare, Nessun titolo-sa leggere e scrivere, Nessun titolo-non sa leggere e/o scrivere

REGIONE:

NORD - Piemonte, Lombardia, Trentino Alto Adige, Veneto, Friuli Venezia Giulia, Liguria, Emilia-Romagna

CENTRO - Toscana, Umbria, Marche, Lazio, Abruzzo, Molise

SUD - Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna

REDDITO:

DA LAVORO- reddito da lavoro dipendente, reddito da lavoro autonomo

PENSIONE - pensione

ALTRO - indennità e provvidenze varie, reddito patrimoniale

MANTENIMENTO FAMILIARI - mantenimento da parte dei familiari

SALUTE:

BUONA SALUTE - molto bene, bene

NE' BENE NE' MALE - nè bene nè male

CATTIVA SALUTE - male, molto male

AIUTI DATI:

1 - un qualsiasi aiuto dato

0 - non ha dato aiuti

AIUTI RICEVUTI:

1 - un qualsiasi aiuto ricevuto

0 - non ha ricevuto aiuti

RETE:

NESSUNO - 0

FAMIGLIA E NON - extended + no-family, immediate+ extended + no-family, immediate+ no-family

FAMIGLIA - extended, immediate+extended, immediate

NO FAMIGLIA - no-family

```

#eta
dati$eta.factor <- as.factor(dati$ETA)
levels(dati$eta.factor) <- c(rep("65-69", 5),
  rep("70-74", 5), rep("75-79", 5), rep("80-84", 5), rep("85+", 12))

#sesso
dati$sesto.factor <- as.factor(dati$SESSO)
levels(dati$sesto.factor) <- c("M", "F")

#istruzione
dati$istruzione.factor <- as.factor(dati$ISTRUZIONE)
levels(dati$istruzione.factor) <- c("ALTO", "ALTO", "ALTO", "ALTO", "MEDIO",
  "MEDIO", "MEDIO", "BASSO", "BASSO", "BASSO")

#regione
dati$regione.factor <- as.factor(dati$REGIONE)
levels(dati$regione.factor) <- c("NORD", "NORD", "NORD", "NORD", "NORD",
  "NORD", "NORD", "CENTRO", "CENTRO", "CENTRO", "CENTRO", "CENTRO", "CENTRO",
  "SUD", "SUD", "SUD", "SUD", "SUD", "SUD")

#reddito
dati$reddito.factor <- as.factor(dati$REDDITO)
levels(dati$reddito.factor) <- c("DA LAVORO", "DA LAVORO", "PENSIONE",
  "ALTRO", "ALTRO", "MANTENIMENTO FAMILIARI")

#salute
dati$salute.factor <- as.factor(dati$SALUTE)
levels(dati$salute.factor) <- c("BUONA SALUTE", "BUONA SALUTE",
  "NE' BENE, N'E MALE", "CATTIVA SALUTE", "CATTIVA SALUTE")

#aiuti dati
dati$AIUTIDATI[is.na(dati$AIUTIDATI)] = "NO"
#sovrascrivo "NA" con "NO" perché "NA" viene letto da R come dato mancante
dati$aiutodato.factor <- as.factor(dati$AIUTIDATI)
levels(dati$aiutodato.factor) <- c(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0)
#1 si, 0 no

#aiuti ricevuti
dati$AIUTIRICEVUTI[is.na(dati$AIUTIRICEVUTI)] = "NO"
#sovrascrivo "NA" con "NO" perché "NA" viene letto da R come dato mancante
dati$aiutoricevuto.factor <- as.factor(dati$AIUTIRICEVUTI)
levels(dati$aiutoricevuto.factor) <- c(1, 1, 1, 1, 1, 1, 1, 1, 0)
#1 si, 0 no

```

```

#creiamo una nuova variabile "aiutoricevutnew" che avrà come livelli
#sempre 0 e 1 ma saranno visti come interi e non come fattori
#(ci serve per la regressione logistica)
dati$aiutoricevutnew<-abs(as.integer(dati$aiutoricevuto.factor)-2)

#rete
dati$rete.factor <- as.factor(dati$RETE)
levels(dati$rete.factor) <- c("NESSUNO", "FAMIGLIA E NON", "FAMIGLIA",
    "FAMIGLIA", "FAMIGLIA E NON", "FAMIGLIA E NON", "FAMIGLIA", "NO FAMIGLIA")

```

Il data frame contiene informazioni su 1722 coppie (in totale 3444 osservazioni).

	n
65-69	762
70-74	1060
75-79	897
80-84	508
85+	217

Table 1: ETA'

	n
ALTO	135
MEDIO	1006
BASSO	2303

Table 2: ISTRUZIONE

	n
NORD	1522
CENTRO	926
SUD	996

Table 3: REGIONE

	n
DA LAVORO	69
PENSIONE	2854
ALTRO	90
MANTENIMENTO FAMILIARI	431

Table 4: REDDITO

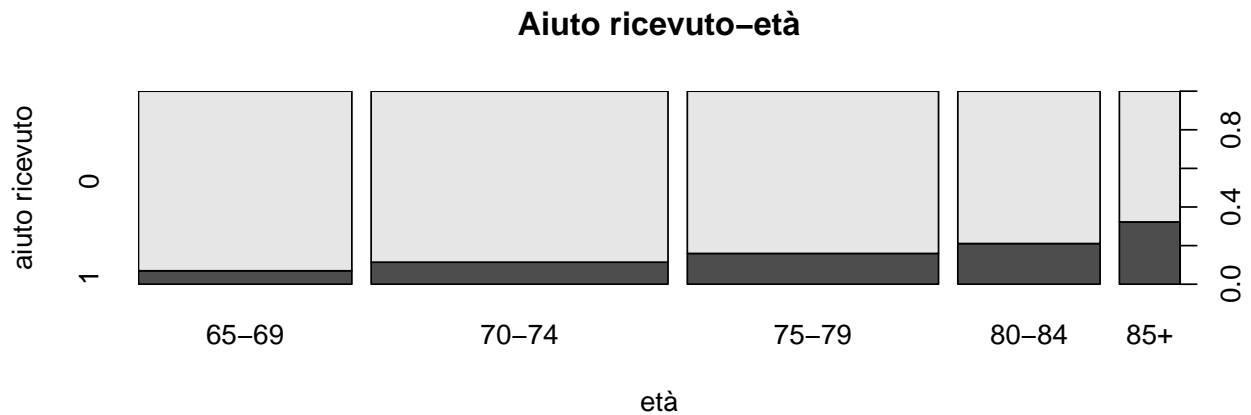
	n
BUONA SALUTE	1223
NE' BENE, N'E MALE	1518
CATTIVA SALUTE	703

Table 5: SALUTE

1. ANALISI DESCRITTIVE

Analizziamo con dei grafici l'eventuale effetto delle variabili età, regione, rete... sulla variabile risposta.

```
plot(dati$eta.factor, dati$aiutoricevuto.factor,
     main= "Aiuto ricevuto-età", xlab="età", ylab="aiuto ricevuto")
```



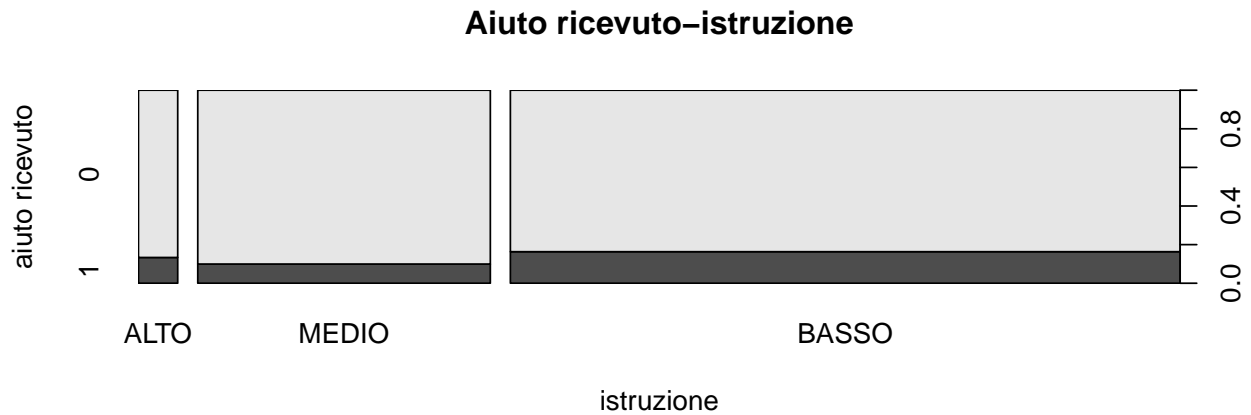
	n
SI	1015
NO	2429

Table 6: AIUTO DATO

	n
SI	494
NO	2950

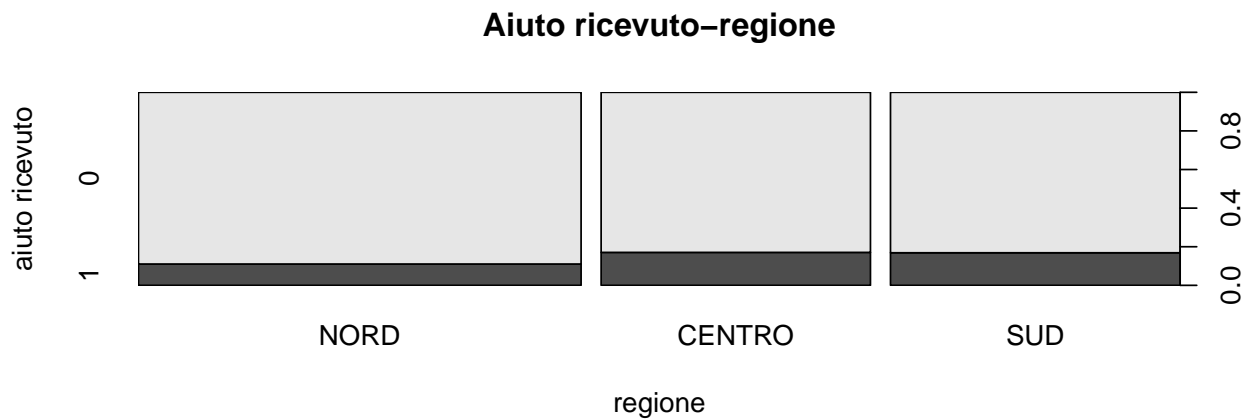
Table 7: AIUTO RICEVUTO

```
plot(dati$istruzione.factor, dati$aiutoricevuto.factor,
     main= "Aiuto ricevuto-istruzione", xlab="istruzione", ylab="aiuto ricevuto")
```

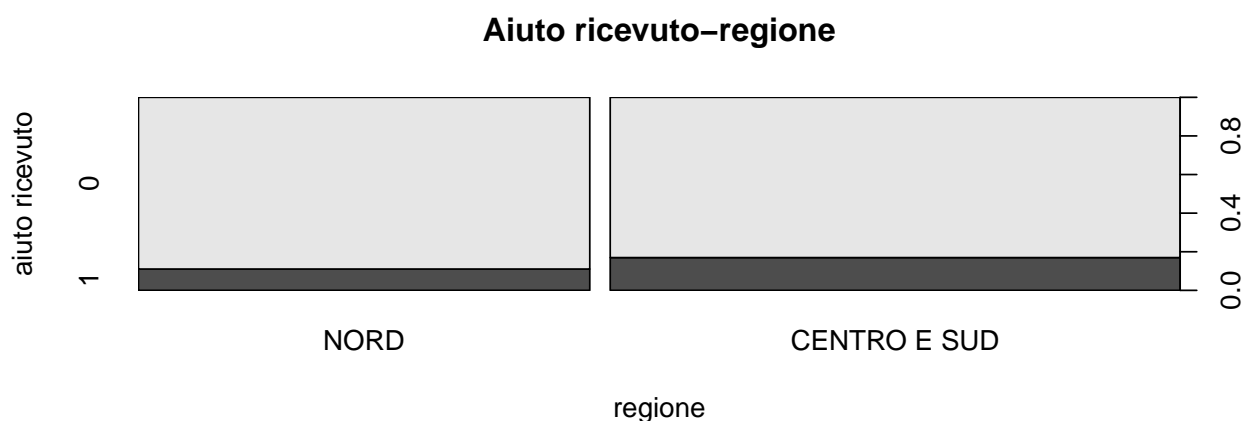


Verifichiamo se la regione possa essere influente per spiegare l'aiuto ricevuto o se la differenza tra nord e centro-sud nell'aiuto ricevuto sia spiegata, ad esempio, da altre variabili, come una diversa distribuzione della condizione di salute.

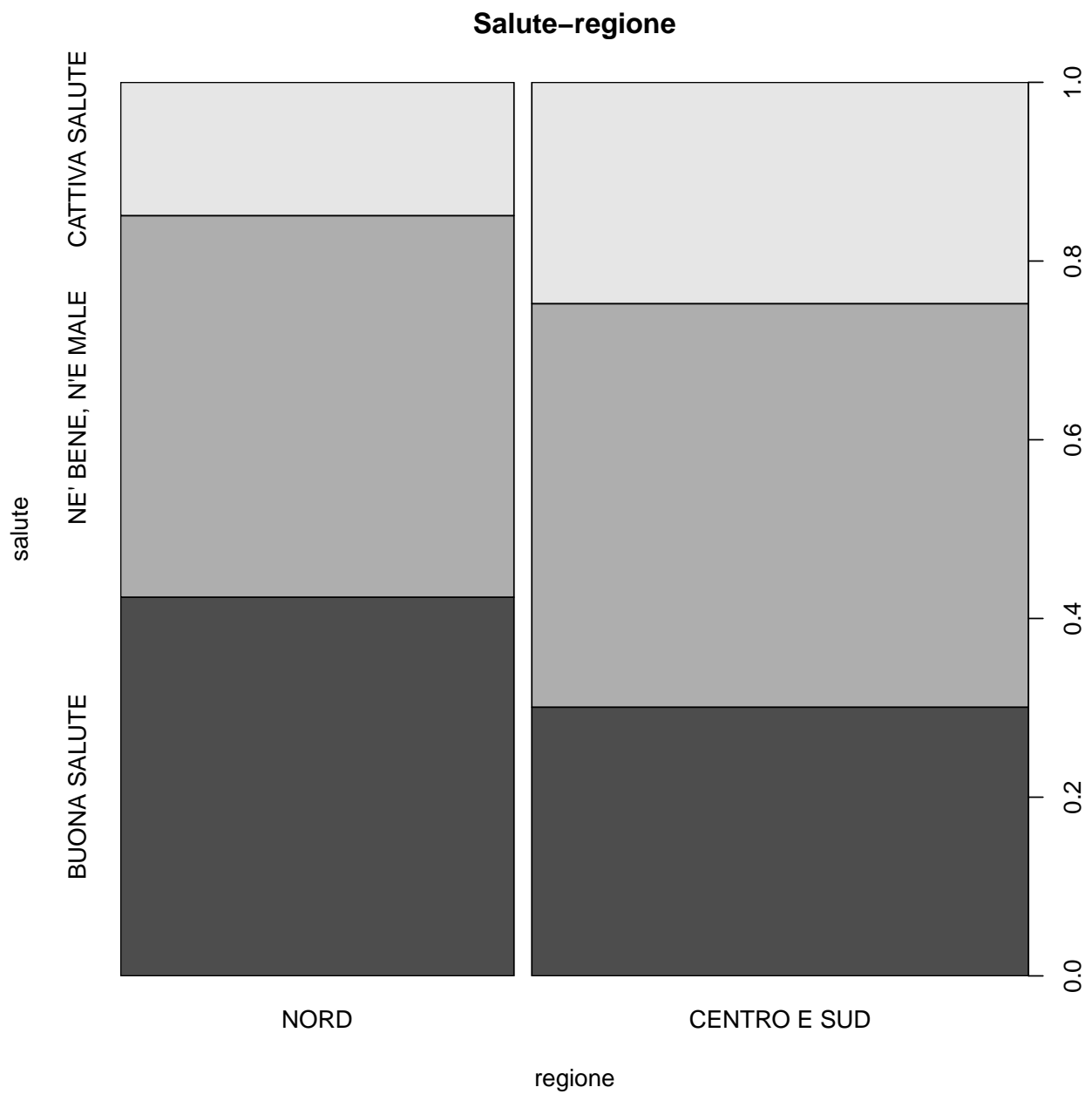
```
plot(dati$regione.factor, dati$aiutoricevuto.factor,
     main= "Aiuto ricevuto-regione", xlab="regione", ylab="aiuto ricevuto")
```



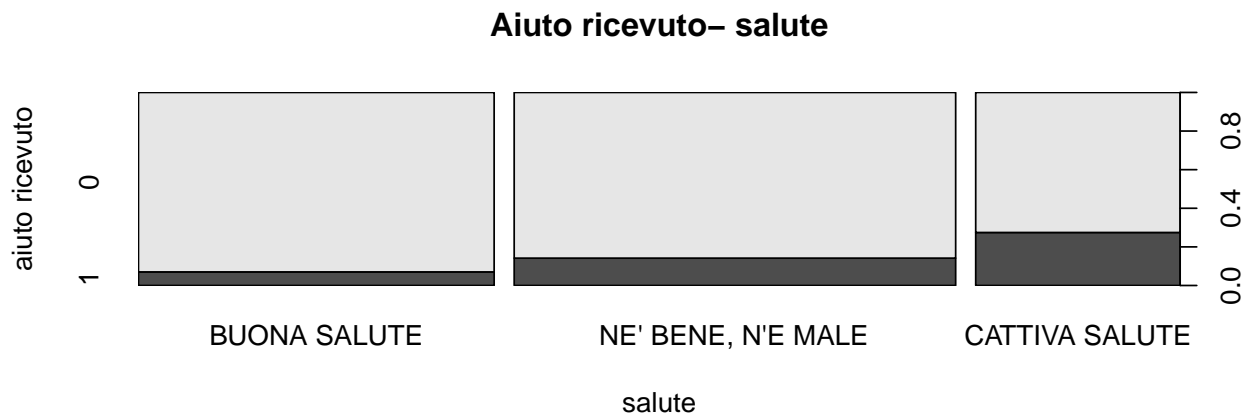
```
#modifichiamo regione aggregando centro e sud
dati$regione1.factor <- as.factor(dati$REGIONE)
levels(dati$regione1.factor) <- c("NORD", "NORD", "NORD", "NORD", "NORD",
  "NORD", "NORD", "CENTRO E SUD", "CENTRO E SUD", "CENTRO E SUD",
  "CENTRO E SUD", "CENTRO E SUD", "CENTRO E SUD", "CENTRO E SUD",
  "CENTRO E SUD", "CENTRO E SUD", "CENTRO E SUD", "CENTRO E SUD")
plot(dati$regione1.factor, dati$aiutoricevuto.factor,
     main= "Aiuto ricevuto-regione", xlab="regione", ylab="aiuto ricevuto")
```



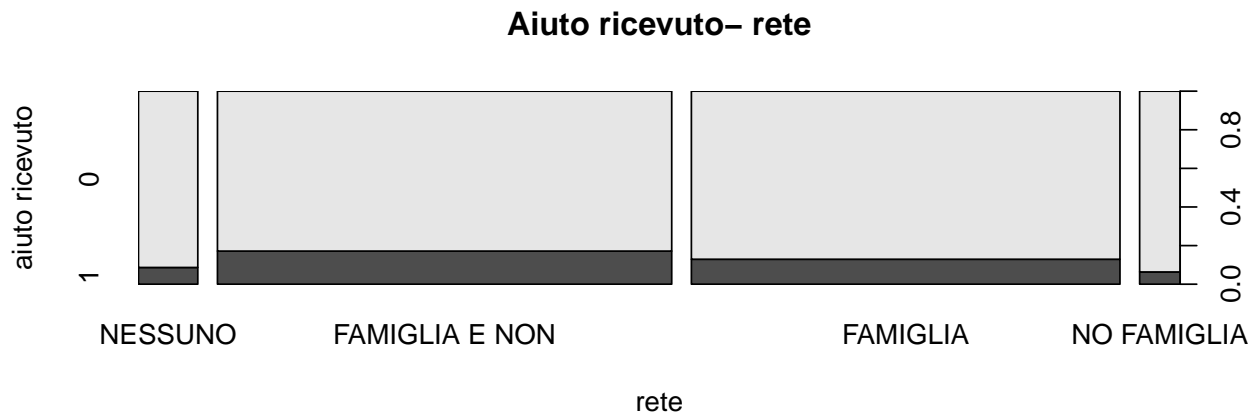
```
plot(dati$regione1.factor, dati$salute.factor,
     main= "Salute-regione", xlab="regione", ylab="salute")
```



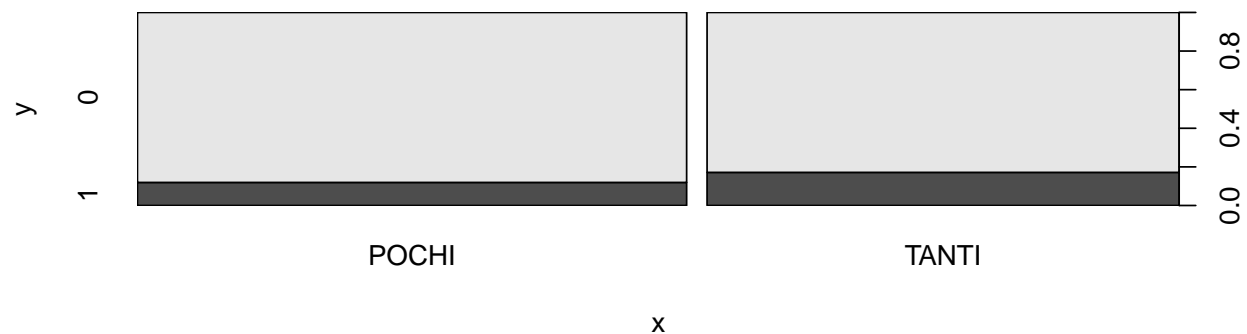
```
plot(dati$salute.factor, dati$aiutoricevuto.factor,
main="Aiuto ricevuto- salute", xlab="salute",ylab="aiuto ricevuto")
```

```
plot(dati$rete.factor, dati$aiutoricevuto.factor,
main="Aiuto ricevuto- rete", xlab="rete",ylab="aiuto ricevuto")
```



```
#modifica variabile rete aggregando (tanti hanno sia famiglia che non famiglia,
#pochi hanno o famiglia o non famiglia)
dati$rete1.factor <- as.factor(dati$RETE)
levels(dati$rete1.factor) <- c("POCHI", "TANTI", "POCHI", "POCHI", "TANTI",
  "TANTI", "POCHI", "POCHI")
plot(dati$rete1.factor,dati$aiutoricevuto.factor)
```

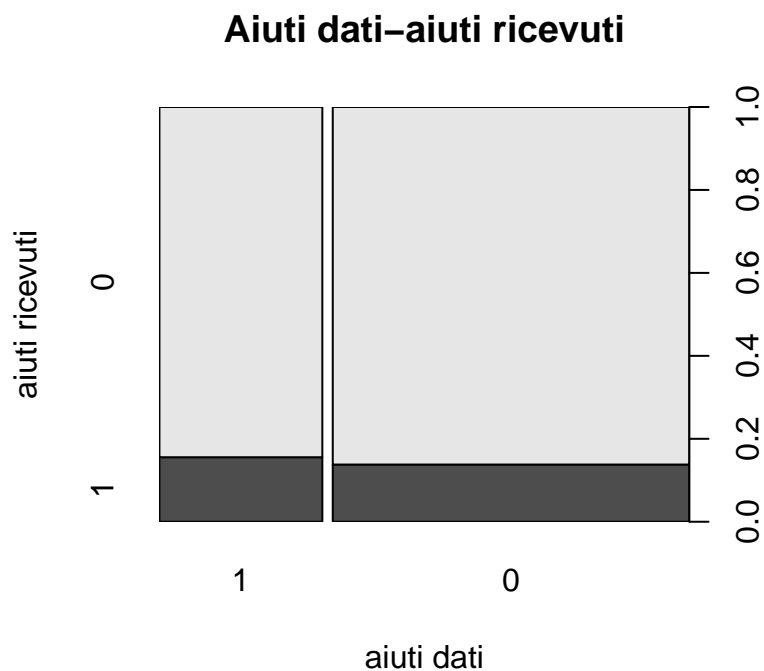


2. MODELLO LOGISTICO

Facciamo le prime analisi con il campione costituito dalle sole donne.

Vediamo graficamente, per le sole femmine, la relazione tra aiuti dati e aiuti ricevuti:

```
plot(dati[dati$sezzo.factor == "F", ]$aiutodato.factor,
     dati[dati$sezzo.factor == "F", ]$aiutoricevuto.factor,
     main="Aiuti dati-aiuti ricevuti", xlab="aiuti dati", ylab="aiuti ricevuti")
```



```

#regressione con tutte le variabili esplicative
fit.f <- glm(aiutoricevutoneu ~ eta.factor + istruzione.factor + regione.factor
  + reddito.factor + salute.factor + aiutodato.factor+ rete.factor,
data = dati[dati$sezzo.factor == "F", ], family = binomial(link = logit))
summary(fit.f)

##
## Call:
## glm(formula = aiutoricevutoneu ~ eta.factor + istruzione.factor +
##     regione.factor + reddito.factor + salute.factor + aiutodato.factor +
##     rete.factor, family = binomial(link = logit), data = dati[dati$sezzo.factor ==
##     "F", ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3710  -0.5751  -0.4234  -0.2782   2.8354
##
## Coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -3.69594      1.18430   -3.121 0.001804
## eta.factor70-74      0.24150      0.21111    1.144 0.252655
## eta.factor75-79      0.72602      0.21582    3.364 0.000768
## eta.factor80-84      0.96213      0.24638    3.905 9.42e-05
## eta.factor85+       1.66076      0.37005    4.488 7.19e-06
## istruzione.factorMEDIO -0.54405      0.47698   -1.141 0.254031
## istruzione.factorBASSO -0.04976      0.44871   -0.111 0.911701
## regione.factorCENTRO  0.38018      0.17974    2.115 0.034416
## regione.factorSUD     0.27511      0.17998    1.529 0.126372
## reddito.factorPENSIONE 0.25848      1.05871    0.244 0.807118
## reddito.factorALTRO    0.86284      1.11503    0.774 0.439031
## reddito.factorMANTENIMENTO FAMILIARI 0.25741      1.06615    0.241 0.809218
## salute.factorNE' BENE, N'E MALE 0.71401      0.20797    3.433 0.000596
## salute.factorCATTIVA SALUTE 1.58802      0.22123    7.178 7.07e-13
## aiutodato.factor0     -0.54248      0.16574   -3.273 0.001064
## rete.factorFAMIGLIA E NON 0.95012      0.42124    2.256 0.024100
## rete.factorFAMIGLIA    0.52235      0.42323    1.234 0.217122
## rete.factorNO FAMIGLIA  0.30075      0.57853    0.520 0.603167
##
## (Intercept)          **
## eta.factor70-74
## eta.factor75-79      ***
## eta.factor80-84      ***
## eta.factor85+        ***
## istruzione.factorMEDIO

```

```

## istruzione.factorBASSO
## regione.factorCENTRO *
## regione.factorSUD
## reddito.factorPENSIONE
## reddito.factorALTRO
## reddito.factorMANTENIMENTO FAMILIARI
## salute.factorNE' BENE, N'E MALE ***
## salute.factorCATTIVA SALUTE ***
## aiutodato.factor0 **
## rete.factorFAMIGLIA E NON *
## rete.factorFAMIGLIA
## rete.factorNO FAMIGLIA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1416.0 on 1721 degrees of freedom
## Residual deviance: 1254.7 on 1704 degrees of freedom
## AIC: 1290.7
##
## Number of Fisher Scoring iterations: 5
#nuova regressione senza istruzione, che risulta non significativa
#dalla prima regressione, e sembra non determinante anche
#dal grafico visto in precedenza (per M e F)
fit.f2 <- glm(aiutoricevutone ~ eta.factor + regione.factor + salute.factor +
  aiutodato.factor+rete.factor, data = dati[dati$sezzo.factor == "F", ],
  family = binomial(link = logit))
summary(fit.f2)

##
## Call:
## glm(formula = aiutoricevutone ~ eta.factor + regione.factor +
## salute.factor + aiutodato.factor + rete.factor, family = binomial(link = logit),
## data = dati[dati$sezzo.factor == "F", ])
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -1.3651 -0.5741 -0.4148 -0.2882 2.7169
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.7108 0.4671 -7.944 1.96e-15 ***
## eta.factor70-74 0.2822 0.2093 1.349 0.177488

```

```
## eta.factor75-79          0.7655      0.2138      3.580 0.000343 ***
## eta.factor80-84          1.0157      0.2430      4.180 2.92e-05 ***
## eta.factor85+            1.6810      0.3668      4.583 4.58e-06 ***
## regione.factorCENTRO     0.3966      0.1789      2.218 0.026588 *
## regione.factorSUD        0.3311      0.1778      1.862 0.062644 .
## salute.factorNE' BENE, N'E MALE 0.7607      0.2049      3.712 0.000206 ***
## salute.factorCATTIVA SALUTE    1.6647      0.2181      7.632 2.31e-14 ***
## aiutodato.factor0        -0.5048      0.1643     -3.073 0.002118 **
## rete.factorFAMIGLIA E NON    0.9700      0.4186      2.317 0.020481 *
## rete.factorFAMIGLIA        0.5499      0.4207      1.307 0.191141
## rete.factorNO FAMIGLIA      0.2881      0.5784      0.498 0.618462
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
```

```
##      Null deviance: 1416.0  on 1721  degrees of freedom
```

```
## Residual deviance: 1264.3  on 1709  degrees of freedom
```

```
## AIC: 1290.3
```

```
##
```

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

```
#verifichiamo se, presa singolarmente, l'istruzione sembra  
#avere una qualche influenza sull'aiuto ricevuto
```

```
fit.fist <- glm(aiutoricevutnew ~ istruzione.factor,  
  data = dati[dati$sezzo.factor == "F", ],family = binomial(link = logit))  
summary(fit.fist)
```

```
##
```

```
## Call:
```

```
## glm(formula = aiutoricevutnew ~ istruzione.factor, family = binomial(link = logit),
```

```
##      data = dati[dati$sezzo.factor == "F", ])
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min        1Q      Median        3Q        Max
```

```
## -0.6030  -0.6030  -0.6030  -0.4122   2.2395
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)  
## (Intercept)      -1.8827     0.4057  -4.641 3.47e-06 ***  
## istruzione.factorMEDIO -0.5401     0.4414  -1.224  0.221  
## istruzione.factorBASSO  0.2704     0.4129   0.655  0.513
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```

## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416  on 1721  degrees of freedom
## Residual deviance: 1395  on 1719  degrees of freedom
## AIC: 1401
##
## Number of Fisher Scoring iterations: 5

#verifichiamo se, preso singolarmente, il reddito sembra
#avere una qualche influenza sull'aiuto ricevuto
fit.fred <- glm(aiutoricevutone ~ reddito.factor,
data = dati[dati$ sesso.factor == "F", ],family = binomial(link = logit))
summary(fit.fred)

##
## Call:
## glm(formula = aiutoricevutone ~ reddito.factor, family = binomial(link = logit),
##      data = dati[dati$ sesso.factor == "F", ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7585  -0.5560  -0.5560  -0.5390   2.3548
##
## Coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -2.7081     1.0328  -2.622  0.00874
## reddito.factorPENSIONE      0.9191     1.0360   0.887  0.37497
## reddito.factorALTRO      1.6094     1.0852   1.483  0.13807
## reddito.factorMANTENIMENTO FAMILIARI  0.8523     1.0424   0.818  0.41357
##
## (Intercept)          **
## reddito.factorPENSIONE
## reddito.factorALTRO
## reddito.factorMANTENIMENTO FAMILIARI
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416  on 1721  degrees of freedom
## Residual deviance: 1411  on 1718  degrees of freedom
## AIC: 1419
##
## Number of Fisher Scoring iterations: 5

```

```
#verifichiamo se, presa singolarmente, la rete(due livelli) sembra
#avere una qualche influenza sull'aiuto ricevuto
fit.fret1 <- glm(aiutoricevutoneu ~ retel.factor,
  data = dati[dati$ sesso.factor == "F", ],family = binomial(link = logit))
summary(fit.fret1)
```

```
##
## Call:
## glm(formula = aiutoricevutoneu ~ retel.factor, family = binomial(link = logit),
##     data = dati[dati$ sesso.factor == "F", ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.6117  -0.6117  -0.5050  -0.5050   2.0605
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.9953     0.1016 -19.635 < 2e-16 ***
## retel.factorTANTI  0.4140     0.1383   2.994  0.00276 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416  on 1721  degrees of freedom
## Residual deviance: 1407  on 1720  degrees of freedom
## AIC: 1411
##
## Number of Fisher Scoring iterations: 4
```

```
#La differenza tra "POCHI" (nessuno,no fam,fam)
#è significativa rispetto a "TANTI" (famiglia e non)
```

```
#nuova regressione senza reddito, che risulta non significativa
#dalla prima regressione e dalla regressione singola,
#e sembra non determinante anche dal grafico visto in precedenza (per M e F)
#fit2 si differenzia da fitf in quanto non presenta le variabili
#reddito e istruzione, e inoltre presenta "rete" aggregata su due livelli.
fit.2<-glm(aiutoricevutoneu ~ eta.factor+salute.factor+
  retel.factor+regione.factor+aiutodato.factor,
  data=dati[dati$ sesso.factor=="F",],family = binomial(link = logit))
summary(fit.2)
```

```
##
```

```
## Call:
## glm(formula = aiutoricevutone ~ eta.factor + salute.factor +
##       rete1.factor + regione.factor + aiutodato.factor, family = binomial(link = logit),
##       data = dati[dati$sezzo.factor == "F", ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3622  -0.5721  -0.4156  -0.2924   2.7423
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -3.2174     0.2529 -12.720 < 2e-16 ***
## eta.factor70-74     0.2767     0.2091   1.324 0.185653
## eta.factor75-79     0.7528     0.2136   3.525 0.000424 ***
## eta.factor80-84     1.0087     0.2427   4.156 3.23e-05 ***
## eta.factor85+       1.6724     0.3657   4.573 4.82e-06 ***
## salute.factorNE' BENE, N'E MALE  0.7679     0.2048   3.750 0.000177 ***
## salute.factorCATTIVA SALUTE      1.6772     0.2179   7.699 1.37e-14 ***
## rete1.factorTANTI      0.4826     0.1481   3.259 0.001119 **
## regione.factorCENTRO    0.4040     0.1786   2.263 0.023654 *
## regione.factorSUD       0.3290     0.1776   1.853 0.063941 .
## aiutodato.factor0      -0.5193     0.1640  -3.166 0.001546 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416.0  on 1721  degrees of freedom
## Residual deviance: 1266.4  on 1711  degrees of freedom
## AIC: 1288.4
##
## Number of Fisher Scoring iterations: 5
```

```
# sostituiamo regione.factor con la macroaggregazione nord-centroSud
fit.4<-glm(aiutoricevutone ~ eta.factor+salute.factor+rete1.factor+
  regione1.factor+aiutodato.factor,data=dati[dati$sezzo.factor=="F",],
  family = binomial(link = logit))
summary(fit.4)
```

```
##
## Call:
## glm(formula = aiutoricevutone ~ eta.factor + salute.factor +
##       rete1.factor + regione1.factor + aiutodato.factor, family = binomial(link = logit),
##       data = dati[dati$sezzo.factor == "F", ])
##
```



```
## Deviance Residuals:
##      Min        1Q      Median        3Q        Max
## -1.3781   -0.5781   -0.4240   -0.2981    2.7424
##
## Coefficients:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   -3.2105     0.2521 -12.737 < 2e-16 ***
## eta.factor70-74                 0.2797     0.2090   1.338 0.180786
## eta.factor75-79                 0.7556     0.2135   3.539 0.000401 ***
## eta.factor80-84                 1.0128     0.2425   4.176 2.96e-05 ***
## eta.factor85+                   1.6798     0.3654   4.597 4.28e-06 ***
## salute.factorNE' BENE, N'E MALE  0.7626     0.2044   3.730 0.000191 ***
## salute.factorCATTIVA SALUTE      1.6706     0.2173   7.689 1.48e-14 ***
## rete1.factorTANTI                0.4809     0.1480   3.249 0.001158 **
## regione1.factorCENTRO E SUD       0.3658     0.1528   2.393 0.016692 *
## aiutodato.factor0               -0.5263     0.1631  -3.226 0.001253 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416.0  on 1721  degrees of freedom
## Residual deviance: 1266.6  on 1712  degrees of freedom
## AIC: 1286.6
##
## Number of Fisher Scoring iterations: 5
```

Regione1 (a due livelli) è poco significativo nel modello perchè parte dell'effetto che ha sulla variabile risposta è spiegato in realtà da diversi stati medi di salute nelle due macro aree (peggiore centro-sud) tuttavia, pur essendo significativa al 99% la teniamo perchè il modello risulta migliore in base al criterio AIC(1286.6 mentre AIC nel modello con regione su tre livelli è 1288.4).

L'influenza che mantiene la regione, al netto della variabile salute, potrebbe essere spiegata da una diversa struttura delle famiglie o diversa attitudine delle persone (è da escludere, invece, una differenza della rete sulla base delle macroregioni, come abbiamo verificato graficamente). Il nostro modello finale quindi contiene le seguenti variabili esplicative: età, salute, rete(su due livelli), regione(su due livelli) e aiuto dato.

Ripetiamo analisi simili per il gruppo costituito dai maschi:

```
#regressione con tutte le variabili esplicative
fit.m <- glm(aiutoricevutoneu ~ eta.factor + istruzione.factor + regione.factor +
reddito.factor + salute.factor + aiutodato.factor+ rete.factor,
data = dati[dati$sezzo.factor == "M", ], family = binomial(link = logit))
summary(fit.m)
```

```
##
## Call:
## glm(formula = aiutoricevutone ~ eta.factor + istruzione.factor +
##       regione.factor + reddito.factor + salute.factor + aiutodato.factor +
##       rete.factor, family = binomial(link = logit), data = dati[dati$sezzo.factor ==
##       "M", ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.6322  -0.5995  -0.4446  -0.2606   2.9439
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -3.77921    0.71245  -5.305 1.13e-07
## eta.factor70-74    1.17142    0.41701   2.809 0.00497
## eta.factor75-79    1.34432    0.41554   3.235 0.00122
## eta.factor80-84    1.76050    0.42204   4.171 3.03e-05
## eta.factor85+      2.33808    0.43366   5.392 6.99e-08
## istruzione.factorMEDIO -0.17026    0.37368  -0.456 0.64865
## istruzione.factorBASSO -0.06455    0.36069  -0.179 0.85797
## regione.factorCENTRO  0.41405    0.17772   2.330 0.01982
## regione.factorSUD     0.28560    0.17844   1.601 0.10947
## reddito.factorPENSIONE -0.37945    0.47364  -0.801 0.42305
## reddito.factorALTRO    1.10456    0.58210   1.898 0.05776
## reddito.factorMANTENIMENTO FAMILIARI -12.00924  378.26201  -0.032 0.97467
## salute.factorNE' BENE, N'E MALE  0.60100    0.19330   3.109 0.00188
## salute.factorCATTIVA SALUTE    1.14407    0.21117   5.418 6.04e-08
## aiutodato.factor0     -0.24036    0.17672  -1.360 0.17379
## rete.factorFAMIGLIA E NON    0.59763    0.34568   1.729 0.08383
## rete.factorFAMIGLIA    0.23733    0.34874   0.681 0.49616
## rete.factorNO FAMIGLIA  -1.44205    0.80603  -1.789 0.07360
##
## (Intercept)          ***
## eta.factor70-74      **
## eta.factor75-79      **
## eta.factor80-84      ***
## eta.factor85+        ***
## istruzione.factorMEDIO
## istruzione.factorBASSO
## regione.factorCENTRO  *
## regione.factorSUD
## reddito.factorPENSIONE
## reddito.factorALTRO  .
```

```
## reddito.factorMANTENIMENTO FAMILIARI
## salute.factorNE' BENE, N'E MALE      **
## salute.factorCATTIVA SALUTE          ***
## aiutodato.factor0
## rete.factorFAMIGLIA E NON            .
## rete.factorFAMIGLIA
## rete.factorNO FAMIGLIA                .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416.0  on 1721  degrees of freedom
## Residual deviance: 1271.8  on 1704  degrees of freedom
## AIC: 1307.8
##
## Number of Fisher Scoring iterations: 12
```

```
#senza istruzione

fit.m2 <- glm(aiutoricevutone ~ eta.factor + regione.factor + salute.factor +
  aiutodato.factor+rete.factor, data = dati[dati$sezzo.factor == "M", ],
  family = binomial(link = logit))
summary(fit.m2)
```

```
##
## Call:
## glm(formula = aiutoricevutone ~ eta.factor + regione.factor +
##      salute.factor + aiutodato.factor + rete.factor, family = binomial(link = logit),
##      data = dati[dati$sezzo.factor == "M", ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2390  -0.6056  -0.4529  -0.2765   2.8876
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -4.1064     0.5237  -7.842 4.44e-15 ***
## eta.factor70-74      1.1286     0.4128   2.734 0.00626 **
## eta.factor75-79      1.3089     0.4109   3.185 0.00145 **
## eta.factor80-84      1.6978     0.4170   4.072 4.67e-05 ***
## eta.factor85+        2.2801     0.4290   5.314 1.07e-07 ***
## regione.factorCENTRO  0.4074     0.1761   2.314 0.02065 *
## regione.factorSUD     0.2970     0.1753   1.694 0.09026 .
## salute.factorNE' BENE, N'E MALE  0.5692     0.1887   3.016 0.00256 **
```

```
## salute.factorCATTIVA SALUTE      1.1010      0.2063      5.338 9.40e-08 ***
## aiutodato.factor0                 -0.2438      0.1742     -1.399  0.16167
## rete.factorFAMIGLIA E NON         0.5719      0.3430      1.667  0.09544 .
## rete.factorFAMIGLIA                0.2389      0.3460      0.690  0.48994
## rete.factorNO FAMIGLIA            -1.3394      0.7937     -1.688  0.09149 .
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
```

```
##      Null deviance: 1416.0  on 1721  degrees of freedom
```

```
## Residual deviance: 1287.9  on 1709  degrees of freedom
```

```
## AIC: 1313.9
```

```
##
```

```
## Number of Fisher Scoring iterations: 6
```

```
#istruzione presa singolarmente non è significativa
```

```
fit.mist <- glm(aiutoricevutone ~ istruzione.factor,
  data = dati[dati$sezzo.factor == "M", ],family = binomial(link = logit))
summary(fit.mist)
```

```
##
```

```
## Call:
```

```
## glm(formula = aiutoricevutone ~ istruzione.factor, family = binomial(link = logit),
```

```
##      data = dati[dati$sezzo.factor == "M", ])
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min        1Q      Median        3Q        Max
```

```
## -0.5902  -0.5902  -0.5902  -0.4908   2.0862
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)      -1.8648      0.3240  -5.755 8.66e-09 ***
```

```
## istruzione.factorMEDIO  -0.1909      0.3502  -0.545  0.586
```

```
## istruzione.factorBASSO   0.2054      0.3345   0.614  0.539
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
```

```
##      Null deviance: 1416.0  on 1721  degrees of freedom
```

```
## Residual deviance: 1409.3  on 1719  degrees of freedom
```

```
## AIC: 1415.3
```

```
##
```

```
## Number of Fisher Scoring iterations: 4
```

#reddito preso singolarmente non è significativo

```
fit.mred <- glm(aiutoricevutoneu ~ reddito.factor,  
  data = dati[dati$ sesso.factor == "M", ],family = binomial(link = logit))  
summary(fit.mred)
```

```
##  
## Call:  
## glm(formula = aiutoricevutoneu ~ reddito.factor, family = binomial(link = logit),  
##     data = dati[dati$ sesso.factor == "M", ])  
##  
## Deviance Residuals:  
##      Min       1Q   Median       3Q      Max   
## -0.9005  -0.5486  -0.5486  -0.5486   2.0874   
##  
## Coefficients:  
##                  Estimate Std. Error z value Pr(>|z|)      
## (Intercept)          -2.0584     0.4335  -4.748 2.05e-06      
## reddito.factorPENSIONE      0.2405     0.4394   0.547  0.584        
## reddito.factorALTRO         1.3652     0.5432   2.513  0.012        
## reddito.factorMANTENIMENTO FAMILIARI -11.5077    378.5931  -0.030  0.976        
##  
## (Intercept)                ***  
## reddito.factorPENSIONE  
## reddito.factorALTRO          *  
## reddito.factorMANTENIMENTO FAMILIARI  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## (Dispersion parameter for binomial family taken to be 1)  
##  
##    Null deviance: 1416.0  on 1721  degrees of freedom  
## Residual deviance: 1405.2  on 1718  degrees of freedom  
## AIC: 1413.2  
##  
## Number of Fisher Scoring iterations: 12
```

```
fit.mret1 <- glm(aiutoricevutoneu ~ rete1.factor,  
  data = dati[dati$ sesso.factor == "M", ],family = binomial(link = logit))  
summary(fit.mret1)
```

```
##  
## Call:  
## glm(formula = aiutoricevutoneu ~ rete1.factor, family = binomial(link = logit),  
##     data = dati[dati$ sesso.factor == "M", ])  
##
```

```
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -0.6151  -0.6151  -0.5033  -0.5033   2.0634
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -2.0022     0.1011 -19.800 < 2e-16 ***
## rete1.factorTANTI  0.4333     0.1382   3.134  0.00172 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416.0  on 1721  degrees of freedom
## Residual deviance: 1406.1  on 1720  degrees of freedom
## AIC: 1410.1
##
## Number of Fisher Scoring iterations: 4
```

*#La differenza tra "pochi" (nessuno,no fam,fam)
#è significativa rispetto a "TANTI" (famiglia e non)*

```
fit.3<-glm(aiutoricevutone ~ eta.factor+salute.factor+rete1.factor
+regione.factor+aiutodato.factor,data=dati[dati$ sesso.factor=="M",],
family = binomial(link = logit))
summary(fit.3) #AIC 1317.6
```

```
##
## Call:
## glm(formula = aiutoricevutone ~ eta.factor + salute.factor +
##      rete1.factor + regione.factor + aiutodato.factor, family = binomial(link = logit),
##      data = dati[dati$ sesso.factor == "M", ])
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2429  -0.5962  -0.4559  -0.2990   2.8243
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -3.9695     0.4230  -9.385 < 2e-16 ***
## eta.factor70-74     1.1337     0.4125   2.748  0.00599 **
## eta.factor75-79     1.3118     0.4106   3.195  0.00140 **
## eta.factor80-84     1.7083     0.4165   4.102 4.10e-05 ***
## eta.factor85+       2.2908     0.4281   5.351 8.75e-08 ***
## salute.factorNE' BENE, N'E MALE  0.5743     0.1881   3.053  0.00227 **
```

```
## salute.factorCATTIVA SALUTE      1.1066      0.2058      5.376 7.61e-08 ***
## retel.factorTANTI                 0.4258      0.1457      2.923 0.00346 **
## regione.factorCENTRO              0.4138      0.1757      2.355 0.01850 *
## regione.factorSUD                 0.2990      0.1751      1.707 0.08777 .
## aiutodato.factor0                -0.2493      0.1737     -1.435 0.15120
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
```

```
##      Null deviance: 1416.0  on 1721  degrees of freedom
```

```
## Residual deviance: 1295.6  on 1711  degrees of freedom
```

```
## AIC: 1317.6
```

```
##
```

```
## Number of Fisher Scoring iterations: 6
```

```
fit.5<-glm(aiutoricevutone ~ eta.factor+salute.factor+retel.factor+
  regione1.factor+aiutodato.factor,data=dati[dati$sezzo.factor=="M",],
  family = binomial(link = logit))
summary(fit.5) #AIC 1316
```

```
##
```

```
## Call:
```

```
## glm(formula = aiutoricevutone ~ eta.factor + salute.factor +
```

```
##      retel.factor + regione1.factor + aiutodato.factor, family = binomial(link = logit
```

```
##      data = dati[dati$sezzo.factor == "M", ])
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -1.2655  -0.6111  -0.4675  -0.3001   2.8227
```

```
##
```

```
## Coefficients:
```

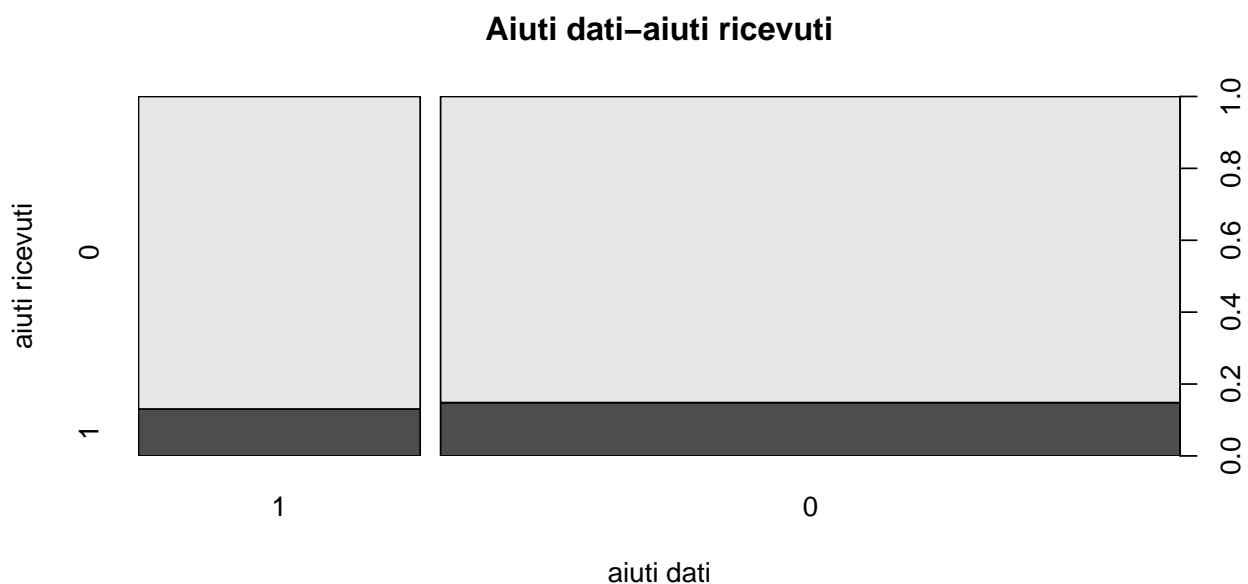
```
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -3.9650      0.4230  -9.374 < 2e-16 ***
## eta.factor70-74      1.1394      0.4124   2.763 0.00573 **
## eta.factor75-79      1.3237      0.4101   3.228 0.00125 **
## eta.factor80-84      1.7146      0.4163   4.118 3.82e-05 ***
## eta.factor85+        2.3004      0.4279   5.376 7.60e-08 ***
## salute.factorNE' BENE, N'E MALE  0.5643      0.1874   3.011 0.00260 **
## salute.factorCATTIVA SALUTE      1.0924      0.2046   5.338 9.39e-08 ***
## retel.factorTANTI      0.4218      0.1455   2.899 0.00374 **
## regione1.factorCENTRO E SUD      0.3551      0.1505   2.360 0.01827 *
## aiutodato.factor0     -0.2524      0.1737  -1.453 0.14610
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1416 on 1721 degrees of freedom
## Residual deviance: 1296 on 1712 degrees of freedom
## AIC: 1316
##
## Number of Fisher Scoring iterations: 6
```

```
plot(dati[dati$sezzo.factor == "M", ]$aiutodato.factor,
     dati[dati$sezzo.factor == "M", ]$aiutoricevuto.factor, main="Aiuti dati-aiuti ricevuti")
```



```
fit.6<-glm(aiutoricevutoneu ~ eta.factor+salute.factor+rete1.factor+
  regione1.factor,data=dati[dati$sezzo.factor=="M",],
  family = binomial(link = logit))
summary(fit.6) #AIC 1316.1
```

```
##
## Call:
## glm(formula = aiutoricevutoneu ~ eta.factor + salute.factor +
## rete1.factor + regione1.factor, family = binomial(link = logit),
## data = dati[dati$sezzo.factor == "M", ])
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -1.1647 -0.6098 -0.4635 -0.3060 2.8701
##
```



```

## Coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -4.1024     0.4134  -9.923  < 2e-16 ***
## eta.factor70-74       1.1282     0.4121   2.738  0.00619 **
## eta.factor75-79       1.2962     0.4094   3.166  0.00155 **
## eta.factor80-84       1.6685     0.4148   4.023 5.76e-05 ***
## eta.factor85+         2.2318     0.4246   5.256 1.47e-07 ***
## salute.factorNE' BENE, N'E MALE  0.5356     0.1862   2.877  0.00401 **
## salute.factorCATTIVA SALUTE      1.0432     0.2012   5.185 2.16e-07 ***
## rete1.factorTANTI         0.4543     0.1436   3.164  0.00156 **
## regione1.factorCENTRO E SUD      0.3431     0.1501   2.286  0.02228 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416.0  on 1721  degrees of freedom
## Residual deviance: 1298.1  on 1713  degrees of freedom
## AIC: 1316.1
##
## Number of Fisher Scoring iterations: 6

```

Nel caso dei maschi la variabile aiuto dato non è significativa (si può vedere anche dal grafico). Per questo il modello finale per i maschi ha come variabili esplicative: età, salute, rete (due livelli) e regione(due livelli).

ANALISI COPPIE

Passiamo ora alla costruzione di un modello logit per la probabilità di ricevere aiuto dalla coppia. Per fare questa analisi abbiamo creato un nuovo dataframe con nuove “variabili di coppia”, in particolare:

REGIONEC: la regione della coppia (1 NORD 2 CENTRO 3 SUD)

ETA'C :

2 OVER 85 - entrambi 85+ anni

1 OVER 85 - uno 85+ anni

0 OVER 85 - entrambi hanno meno di 85 anni

SALUTEC:

2 CATTIVA SALUTE - entrambi in cattiva salute

1 CATTIVA SALUTE - uno in cattiva salute

2 BUONA SALUTE - entrambi in buona salute

NE' BENE NE' MALE - entrambi nè bene nè male

AIUTO DATOC:

2 AIUTI DATI - entrambi hanno dato aiuto

1 AIUTO DATO - uno dei due ha dato aiuto

0 AIUTI DATI - nessuno ha dato aiuto

AIUTO RICEVUTOC:

2 AIUTI RICEVUTI - entrambi hanno ricevuto aiuto

1 AIUTO RICEVUTO - uno dei due ha ricevuto aiuto

0 AIUTI RICEVUTI - nessuno ha ricevuto aiuto

RETEC:

2 TANTI - entrambi hanno famiglia e no famiglia

1 TANTI - uno ha famiglia e no famiglia 0 TANTI - nessuno ha famiglia e no famiglia (ma possono avere ad es. solo famiglia o solo amici)

```
n=length(dati$REGIONE)

#regioneC
regioneC=c(rep (0,n/2))
j=1

i=1
for (j in 1: (n/2)) {
regioneC[j]= dati$regione.factor[i]
i=i+2
}
```

```
table(regioneC)
```

```
## regioneC  
##    1    2    3  
## 761 463 498
```

```
#etaC
```

```
i=1  
etaC=c(rep ("0",n/2))  
for (j in 1:(n/2)) {  
  if(dati$ETA[i]>=85 && dati$ETA[i+1]>=85)  
  
    {etaC[j]="2 over 85"}  
  
  else if((dati$ETA[i]>=85 && dati$ETA[i+1]<85) ||  
          (dati$ETA[i]<85 && dati$ETA[i+1]>=85))  
    {etaC[j]="1 over 85"}  
  else  
    {etaC[j]="0 over 85"}  
  
  i=i+2  
}
```

```
table(etaC)
```

```
## etaC  
## 0 over 85 1 over 85 2 over 85  
##      1544      139      39
```

```
#saluteC
```

```
dati$salute1.factor <- as.factor(dati$SALUTE)  
levels(dati$salute1.factor) <- c(2,2,1, 0, 0)
```

```
#buona salute 2
```

```
#ne bene ne male 1
```

```
#cattiva salute 0
```

```
i=1  
saluteC=c(rep ("0",n/2))  
  
for (j in 1:(n/2)) {  
  if(dati$salute1.factor[i]==2 && dati$salute1.factor[i+1]==2)  
  
    {saluteC[j]="2 BUONA SALUTE"}  
  
  else if((dati$salute1.factor[i]==0 && dati$salute1.factor[i+1]!= 0) ||  
          (dati$salute1.factor[i]!=0 && dati$salute1.factor[i+1]==0))
```

```

{saluteC[j]="1 CATTIVA SALUTE"}

else if ((dati$salute1.factor[i]==0 && dati$salute1.factor[i+1]== 0))

{saluteC[j]="2 CATTIVA SALUTE"}

else {saluteC[j]="NE MALE NE BENE"}

i=i+2
}

```

```
table(saluteC)
```

```
## saluteC
## 1 CATTIVA SALUTE    2 BUONA SALUTE 2 CATTIVA SALUTE  NE MALE NE BENE
##                371                383                166                802
```

```

#aiuto datoC
i=1
AiutoDatoC=c(rep ("0",n/2))

for (j in 1:(n/2)) {
  if(dati$aiutodato.factor[i]==1 && dati$aiutodato.factor[i+1]==1)

  {AiutoDatoC[j]="2 aiutidati"}

  else if ((dati$aiutodato.factor[i]==0 && dati$aiutodato.factor[i+1]== 0))

  {AiutoDatoC[j]="0 aiuti dati"}
  else
  {AiutoDatoC[j]="1 aiutodato"}

  i=i+2
}

```

```
table(AiutoDatoC)
```

```
## AiutoDatoC
## 0 aiuti dati  1 aiutodato  2 aiutidati
##           1084           261           377
```

```

#aiuto ricevuto
AiutoRicevutoC <- rep(0, n / 2)

i = 1

for (j in 1 : (n / 2)) {

  AiutoRicevutoC[j] = dati$aiutoricevutoneu[i]

  i = i + 2
}
#1 sì, 0 no
table(AiutoRicevutoC)

```

```

## AiutoRicevutoC
##      0      1
## 1475  247

```

```

#reteC
i=1
ReteC=c(rep ("0",n/2))

for (j in 1:(n/2)) {
  if(dati$rete1.factor[i]=="TANTI" && dati$rete1.factor[i+1]=="TANTI")

    {ReteC[j]="2 TANTI"}

  else if ((dati$rete1.factor[i]=="POCHI" && dati$rete1.factor[i+1]== "POCHI"))

    {ReteC[j]="0 TANTI"}

  else {ReteC[j]="1 TANTI"}

  i=i+2
}

table(ReteC)

```

```

## ReteC
## 0 TANTI 1 TANTI 2 TANTI
##      817      218      687

```

```
DatiC<-data.frame(etaC,saluteC,ReteC,AiutoDatoC,regioneC,AiutoRicevutoC)
#creato data frame per le coppie quindi
#ha 1722 dati e le variabili sono aggregate
head(DatiC)
```

```
##          etaC          saluteC  ReteC  AiutoDatoC regioneC AiutoRicevutoC
## 1 0 over 85 NE MALE NE BENE 0 TANTI 0 aiuti dati          1          0
## 2 0 over 85 2 BUONA SALUTE 2 TANTI 2 aiutidati          1          0
## 3 0 over 85 2 BUONA SALUTE 2 TANTI 1 aiutodato          1          0
## 4 0 over 85 NE MALE NE BENE 0 TANTI 1 aiutodato          1          0
## 5 0 over 85 NE MALE NE BENE 2 TANTI 1 aiutodato          1          0
## 6 0 over 85 2 BUONA SALUTE 1 TANTI 1 aiutodato          1          0
```

```
# modello con tutte le variabili esplicative
fit.c <- glm(AiutoRicevutoC ~ etaC + regioneC + saluteC + AiutoDatoC + ReteC,
  data = DatiC, family = binomial(link = logit))
summary(fit.c)
```

```
##
## Call:
## glm(formula = AiutoRicevutoC ~ etaC + regioneC + saluteC + AiutoDatoC +
##      ReteC, family = binomial(link = logit), data = DatiC)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5466  -0.5880  -0.4245  -0.2839   2.7329
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -2.0932     0.2420  -8.649  < 2e-16 ***
## etaC1 over 85     0.7200     0.2219   3.245  0.00117 **
## etaC2 over 85     1.4930     0.3571   4.181 2.90e-05 ***
## regioneC         0.1419     0.0870   1.631  0.10298
## saluteC2 BUONA SALUTE -1.7588     0.2846 -6.181 6.38e-10 ***
## saluteC2 CATTIVA SALUTE  0.5030     0.2145   2.345  0.01903 *
## saluteCNE MALE NE BENE -0.7618     0.1719 -4.433 9.31e-06 ***
## AiutoDatoC1 aiutodato  0.5545     0.1983   2.797  0.00516 **
## AiutoDatoC2 aiutidati  0.3515     0.1932   1.819  0.06884 .
## ReteC1 TANTI         0.3193     0.2298   1.390  0.16463
## ReteC2 TANTI         0.5074     0.1586   3.199  0.00138 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##      Null deviance: 1416   on 1721   degrees of freedom
## Residual deviance: 1273   on 1711   degrees of freedom
## AIC: 1295
##
## Number of Fisher Scoring iterations: 5

#senza regione perché risulta non significativa nel primo modello
fit.c1 <- glm(AiutoRicevutoC ~ etaC + saluteC + AiutoDatoC + ReteC,
              data = DatiC, family = binomial(link = logit))
summary(fit.c1) #AIC 1295.7

##
## Call:
## glm(formula = AiutoRicevutoC ~ etaC + saluteC + AiutoDatoC +
##      ReteC, family = binomial(link = logit), data = DatiC)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4838  -0.5654  -0.4454  -0.2976   2.6944
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -1.8100     0.1652 -10.960 < 2e-16 ***
## etaC1 over 85       0.7248     0.2216   3.270  0.00107 **
## etaC2 over 85       1.4530     0.3560   4.082 4.46e-05 ***
## saluteC2 BUONA SALUTE -1.7932     0.2837  -6.320 2.61e-10 ***
## saluteC2 CATTIVA SALUTE  0.5453     0.2128   2.563  0.01039 *
## saluteCNE MALE NE BENE -0.7705     0.1717  -4.488 7.18e-06 ***
## AiutoDatoC1 aiutodato   0.5386     0.1977   2.725  0.00644 **
## AiutoDatoC2 aiutidati   0.3196     0.1920   1.665  0.09595 .
## ReteC1 TANTI           0.3253     0.2290   1.420  0.15547
## ReteC2 TANTI           0.5082     0.1585   3.207  0.00134 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416.0   on 1721   degrees of freedom
## Residual deviance: 1275.7   on 1712   degrees of freedom
## AIC: 1295.7
##
## Number of Fisher Scoring iterations: 5

fit.creg <- glm(AiutoRicevutoC ~ regioneC, data = DatiC,
                family = binomial(link = logit))
summary(fit.creg)
```

```
##
## Call:
## glm(formula = AiutoRicevutoC ~ regioneC, family = binomial(link = logit),
##      data = DatiC)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q        Max
## -0.6301   -0.5622   -0.5007   -0.5007    2.0682
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.26190     0.17406 -12.995  < 2e-16 ***
## regioneC      0.24859     0.08099   3.069  0.00214 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1416.0  on 1721  degrees of freedom
## Residual deviance: 1406.6  on 1720  degrees of freedom
## AIC: 1410.6
##
## Number of Fisher Scoring iterations: 4
# i risultati della regressione semplice con la variabile regione sono in linea
# con quanto visto già in precedenza con le regressioni
# di maschi e femmine: essa è infatti significativa nella regressione semplice,
# ma cessa di esserlo nella regressione multipla
# questo fenomeno di collinearità è dovuto alla lineare dipendenza
# tra la variabile regione e la variabile salute
```

Per le coppie il modello finale ha come variabili esplicative l'età, salute, aiuto dato e rete