GROUP 3

Thành viên

Nguyễn Thị Thùy Trang 1752048

Nguyễn Minh Nguyệt 1752038

Trương Thanh Trúc 1752051

Nguyễn Anh Khôi 1752024

Đặng Hữu Phước Vinh 1752052

Nguyễn Hoàng Long 1752028

Ex3.r

Admin

2020-05-25

```
EX3<-read.csv(file.choose())


reg<-lm(EX3$y_i~., data=EX3)
step(reg)
```

```
## Start:  AIC=25.03
## EX3$y_i ~ x_i.1 + x_i.2 + x_i.3 + x_i.4 + x_i.5 + x_i.6 + x_i.7 +
##     x_i.8 + x_i.9 + x_i.10 + x_i.11 + x_i.12 + x_i.13
##
##          Df Sum of Sq    RSS    AIC
## - x_i.10  1    0.0000 36.138 23.028
## - x_i.6   1    0.0183 36.157 23.048
## - x_i.2   1    0.0372 36.175 23.068
## - x_i.5   1    0.0578 36.196 23.090
## - x_i.11  1    0.1223 36.261 23.160
## - x_i.7   1    0.1561 36.294 23.196
## - x_i.13  1    0.4744 36.613 23.537
## - x_i.12  1    1.0769 37.215 24.173
## - x_i.3   1    1.3751 37.513 24.484
## <none>                36.138 25.028
## - x_i.8   1    2.6392 38.777 25.777
## - x_i.9   1    3.3887 39.527 26.523
## - x_i.4   1    3.4710 39.609 26.605
## - x_i.1   1    5.0696 41.208 28.148
##
## Step:  AIC=23.03
## EX3$y_i ~ x_i.1 + x_i.2 + x_i.3 + x_i.4 + x_i.5 + x_i.6 + x_i.7 +
##     x_i.8 + x_i.9 + x_i.11 + x_i.12 + x_i.13
##
##          Df Sum of Sq    RSS    AIC
## - x_i.6   1    0.0183 36.157 21.048
## - x_i.2   1    0.0496 36.188 21.081
## - x_i.5   1    0.0585 36.197 21.091
## - x_i.11  1    0.1258 36.264 21.163
## - x_i.7   1    0.1638 36.302 21.204
## - x_i.13  1    0.4767 36.615 21.539
## - x_i.12  1    1.0905 37.229 22.187
## - x_i.3   1    1.4005 37.539 22.511
## <none>                36.138 23.028
```

```
## - x_i.8   1    2.7176 38.856 23.856
## - x_i.9   1    3.3927 39.531 24.527
## - x_i.4   1    3.5201 39.658 24.653
## - x_i.1   1    5.1050 41.243 26.181
##
## Step:  AIC=21.05
## EX3$y_i ~ x_i.1 + x_i.2 + x_i.3 + x_i.4 + x_i.5 + x_i.7 + x_i.8 +
##     x_i.9 + x_i.11 + x_i.12 + x_i.13
##
##          Df Sum of Sq    RSS    AIC
## - x_i.2   1    0.0387 36.195 19.089
## - x_i.11  1    0.1467 36.303 19.206
## - x_i.5   1    0.1665 36.323 19.227
## - x_i.7   1    0.1749 36.332 19.236
## - x_i.13  1    0.5166 36.673 19.601
## - x_i.12  1    1.2401 37.397 20.363
## - x_i.3   1    1.7711 37.928 20.913
## <none>                36.157 21.048
## - x_i.8   1    3.1461 39.303 22.302
## - x_i.9   1    5.0785 41.235 24.174
## - x_i.1   1    5.3107 41.467 24.392
## - x_i.4   1    8.0463 44.203 26.884
##
## Step:  AIC=19.09
## EX3$y_i ~ x_i.1 + x_i.3 + x_i.4 + x_i.5 + x_i.7 + x_i.8 + x_i.9 +
##     x_i.11 + x_i.12 + x_i.13
##
##          Df Sum of Sq    RSS    AIC
## - x_i.11  1    0.1088 36.304 17.206
## - x_i.5   1    0.1586 36.354 17.260
## - x_i.7   1    0.4233 36.619 17.543
## - x_i.13  1    0.5178 36.713 17.643
## - x_i.12  1    1.3982 37.594 18.567
## - x_i.3   1    1.7346 37.930 18.915
## <none>                36.195 19.089
## - x_i.8   1    3.1586 39.354 20.352
## - x_i.9   1    5.1173 41.313 22.247
## - x_i.1   1    5.2923 41.488 22.412
```

```
## - x_i.4   1    8.0863 44.282 24.953
##
## Step:  AIC=17.21
## EX3$y_i ~ x_i.1 + x_i.3 + x_i.4 + x_i.5 + x_i.7 + x_i.8 + x_i.9 +
##     x_i.12 + x_i.13
##
##          Df Sum of Sq    RSS    AIC
## - x_i.5   1    0.1650 36.469 15.383
## - x_i.7   1    0.3509 36.655 15.582
## - x_i.13  1    0.9729 37.277 16.238
## - x_i.3   1    1.7593 38.063 17.052
## <none>                36.304 17.206
## - x_i.8   1    3.0867 39.391 18.389
## - x_i.12  1    3.4995 39.804 18.796
## - x_i.9   1    5.0977 41.402 20.331
## - x_i.1   1    5.4046 41.709 20.619
## - x_i.4   1    8.0335 44.338 23.003
##
## Step:  AIC=15.38
## EX3$y_i ~ x_i.1 + x_i.3 + x_i.4 + x_i.7 + x_i.8 + x_i.9 + x_i.12 +
##     x_i.13
##
##          Df Sum of Sq    RSS    AIC
## - x_i.7   1    0.3371 36.806 13.742
## - x_i.13  1    0.8809 37.350 14.314
## - x_i.3   1    1.6675 38.137 15.127
## <none>                36.469 15.383
## - x_i.8   1    3.3212 39.790 16.782
## - x_i.12  1    3.5882 40.057 17.043
## - x_i.9   1    4.9612 41.430 18.358
## - x_i.1   1    5.2504 41.719 18.629
## - x_i.4   1    8.2684 44.737 21.353
##
## Step:  AIC=13.74
## EX3$y_i ~ x_i.1 + x_i.3 + x_i.4 + x_i.8 + x_i.9 + x_i.12 + x_i.13
##
##          Df Sum of Sq    RSS    AIC
## - x_i.13  1    0.5567 37.363 12.328
```

```
## - x_i.3   1    1.7695 38.576 13.573
## <none>                36.806 13.742
## - x_i.8   1    3.0218 39.828 14.819
## - x_i.12  1    3.5512 40.357 15.334
## - x_i.1   1    4.9177 41.724 16.633
## - x_i.9   1    5.6111 42.417 17.276
## - x_i.4   1    8.2662 45.072 19.644
##
## Step:  AIC=12.33
## EX3$y_i ~ x_i.1 + x_i.3 + x_i.4 + x_i.8 + x_i.9 + x_i.12
##
##          Df Sum of Sq    RSS    AIC
## - x_i.3   1    1.5471 38.910 11.910
## <none>                37.363 12.328
## - x_i.12  1    3.3449 40.708 13.671
## - x_i.8   1    3.6354 40.998 13.949
## - x_i.9   1    5.0559 42.419 15.277
## - x_i.1   1    6.2339 43.597 16.346
## - x_i.4   1    7.7383 45.101 17.669
##
## Step:  AIC=11.91
## EX3$y_i ~ x_i.1 + x_i.4 + x_i.8 + x_i.9 + x_i.12
##
##          Df Sum of Sq    RSS    AIC
## <none>                38.910 11.910
## - x_i.12  1    3.6128 42.523 13.373
## - x_i.8   1    5.3552 44.265 14.939
## - x_i.9   1    5.4273 44.337 15.002
## - x_i.4   1    8.2199 47.130 17.384
## - x_i.1   1    9.8731 48.783 18.729
##
## Call:
## lm(formula = EX3$y_i ~ x_i.1 + x_i.4 + x_i.8 + x_i.9 + x_i.12,
##     data = EX3)
##
## Coefficients:
## (Intercept)        x_i.1        x_i.4        x_i.8        x_i.9       x_i.12
##    10.10349     -0.06871     -0.11023      0.79395      0.06545     -0.72254
```

À l’aide de méthode AIC :

On utilise fonction step() pour trouver le meilleur modèle par critère du AIC

On voit le résultat :

Start : le plus faible c’est : 23.028 (rejette X_i.10)

Step() : le plus faible c’est : 21.048 (rejette X_i.6)

Step() : le plus faible c’est : 19.089 (rejette X_i.2)

Step() : le plus faible c’est : 17.206 (rejette X_i.11)

Step() : le plus faible c’est : 15.383 (rejette X_i.5)

Step() : le plus faible c’est : 13.742 (rejette X_i.7)

Step() : le plus faible c’est : 12.328 (rejette X_i.13)

Step() : le plus faible c’est : 11.910 (rejette X_i.3)

Step() : le plus faible c’est : 11.910 (rejette none)

=> Le meilleur modele :

Y_chapeau = 10.10349 - 0.06871*X1_chapeau - 0.11023*X4_chapeau + 0.79395*X8_chapeau + 0.06545*X9_chapeau - 0.72254*X12_chapeau