

Broom

Pacchetto per gestire l'output dei modelli, ed avere tabelle organizzate come dataframe.

Con queste funzioni, abbiamo dataframe che possono anche essere meglio utilizzati per altre elaborazioni (inserire variabili nel dataset originario, fare grafici ecc.), e per essere inseriti nei report riproducibili (Rmd → Word, PDF, ecc.).

Principali funzioni	
<code>glance()</code>	tabella di sintesi del modello
<code>tidy()</code>	tabella dell'output
<code>augment()</code>	tabella con i dati del modello, caso per caso

Vedi: [Vignettes](#); [RDocumentation](#)

Esempi

glance

`glance()` fornisce le informazioni necessarie alla valutazione del modello. Vengono presentate anche statistiche che nell'output normale (con `summary()`) non sono di default.

```
library(broom)
lm.res <- lm(data=cars, dist ~ speed)

glance(lm.res)
```

```
## # A tibble: 1 x 11
##   r.squared adj.r.squared sigma statistic  p.value    df logLik
##   <dbl>      <dbl> <dbl>    <dbl>    <dbl> <int>  <dbl>
## 1  0.651      0.644   15.4     89.6 1.49e-12     2 -207.419
## # ... with 2 more variables: deviance <dbl>, df.residual <int>
```

tidy

Restituisce l'output di `summary()`, semplificato e organizzato in tabella:

```
tidy(lm.res)
```

```
## # A tibble: 2 x 5
```

```
##      term          estimate std.error statistic  p.value
##      <chr>          <dbl>      <dbl>      <dbl>    <dbl>
## 1 (Intercept)    -17.6         6.76       -2.60 1.23e- 2
## 2 speed           3.93         0.416       9.46 1.49e-12
```

augment

```
augment(lm.res)
```

```
## # A tibble: 50 x 9
##      dist speed .fitted .se.fit .resid   .hat .sigma .cooksd
##      <dbl> <dbl>   <dbl>   <dbl> <dbl>   <dbl> <dbl>   <dbl>
## 1      2     4   -1.85    5.21   3.85  0.115   15.5  0.00459
## 2     10     4   -1.85    5.21  11.8  0.115   15.4  0.0435
## 3      4     7    9.95    4.11  -5.95  0.0715   15.5  0.00620
## 4     22     7    9.95    4.11  12.1  0.0715   15.4  0.0255
## 5     16     8   13.9    3.77   2.12  0.0600   15.5  0.000645
## 6     10     9   17.8    3.44  -7.81  0.0499   15.5  0.00713
## 7     18    10   21.7    3.12  -3.74  0.0413   15.5  0.00133
## 8     26    10   21.7    3.12   4.26  0.0413   15.5  0.00172
## 9     34    10   21.7    3.12  12.3  0.0413   15.4  0.0143
## 10    17    11   25.7    2.84  -8.68  0.0341   15.5  0.00582
## # ... with 40 more rows
```

Con gli altri pacchetti del tidyverse

Selezioniamo i casi con i residui standard $> |1,5|$ (per studiare gli *outliers*):

```
library(tidyverse)
library(knitr)
```

```
augment(lm.res) %>%
  select(dist, speed, .fitted, .resid, .std.resid) %>% # seleziona
le colonne
  filter(.std.resid > abs(1.5)) %>% # seleziona
i casi
  arrange(desc(.std.resid)) %>% # ordina le
righe
  kable() # formatta
la tabella
```

dist	speed	.fitted	.resid	.std.resid
120	24	76.79872	43.20128	2.919060
80	14	37.47463	42.52537	2.795166
84	18	53.20426	30.79574	2.027818
76	18	53.20426	22.79574	1.501039

Questa è una tabella formattata, in formato markdown (quindi può essere esportata in formato html, word, ecc.).

[Comandi](#), [Funzioni](#), [Tidyverse](#)

From:

<https://www.agnesevardanega.eu/wiki/> - **Ricerca Sociale con R**

Permanent link:

https://www.agnesevardanega.eu/wiki/r/tidyverse/pacchetto_broom?rev=1670001142

Last update: **02/12/2022 17:12**

