

Determinazione del prezzo delle automobili

Acampora L. Farenga R. Franzo C. Jovicevic M. Scattolon S.

14 dicembre 2025

Sommario

Il presente report analizza i fattori che influenzano il prezzo delle auto usate, utilizzando un dataset di 478 osservazioni contenente informazioni su variabili tecniche e caratteristiche dei veicoli. L'obiettivo principale è quello di identificare le relazioni tra il prezzo delle auto (variabile dipendente) e una serie di variabili indipendenti, tra cui chilometraggio, anzianità della macchina, potenza (in cavalli), tipo di cambio, sistema di alimentazione, presenza di danni non riparati, e marche di auto. Alcune variabili sono state trasformate in variabili dummy per poter essere correttamente utilizzate nel modello di regressione lineare.

Assunzioni principali del modello

- Relazione negativa tra anzianità della macchina e prezzo:** Auto più vecchie perdono di valore col tempo.
- Relazione negativa tra chilometraggio e prezzo:** Maggiore chilometraggio implica una riduzione del prezzo.
- Influenza della potenza:** Maggiore potenza motore è associata a un prezzo più alto.
- Impatto del tipo di carburante e danni:** Carburante e danni influiscono sul valore percepito.

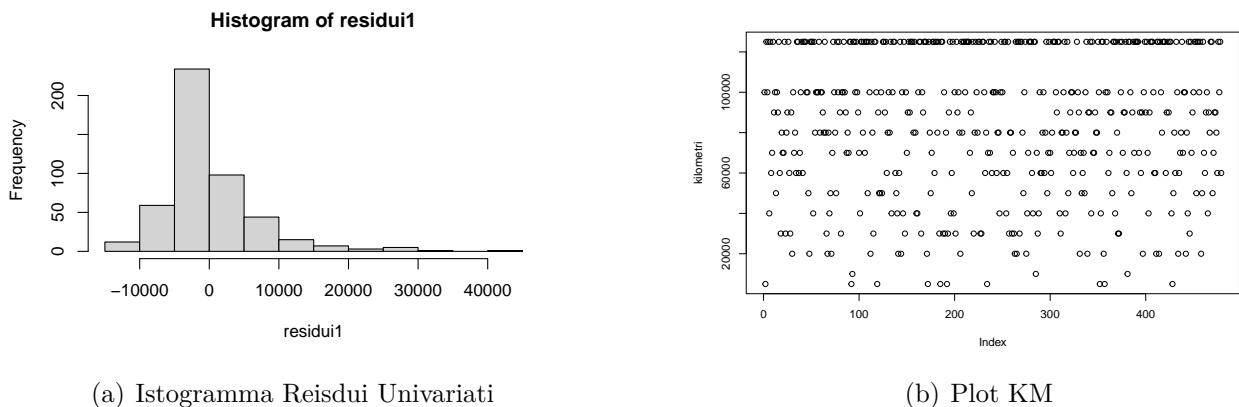
Risultati della Regressione Univariata

$$\text{Prezzo Auto} = \beta_0 + \beta_1 \cdot \text{chilometraggio} + \epsilon \quad (1)$$

La regressione univariata iniziale è stata condotta per analizzare l'effetto del chilometraggio (variabile indipendente) sul prezzo delle auto usate (variabile dipendente). Di seguito sono riportati i principali output della regressione:

Output	Valore	Significato
Intercetta β_0	14.870 ***	il prezzo di un'auto senza considerare nessun regressore.
Chilometraggio β_1	0,08221 ***	Ogni chilometro aggiuntivo riduce il prezzo di € 0,082.
R^2 Adjusted	0,15	Percentuale di varianza spiegata dal modello.
Test di Shapiro (p-value)	< 0,001 ***	Residui non seguono una distribuzione normale.
Test di Breusch-Pagan (BP)	9,3536	Eteroschedasticità presente.
p-value BP	0,002225 **	Conferma di eteroschedasticità significativa.

Nonostante la variabile sia significativa, il modello univariato non è sufficiente a spiegare il fenomeno, come evidenziato dal basso R^2 pari a 0,15.



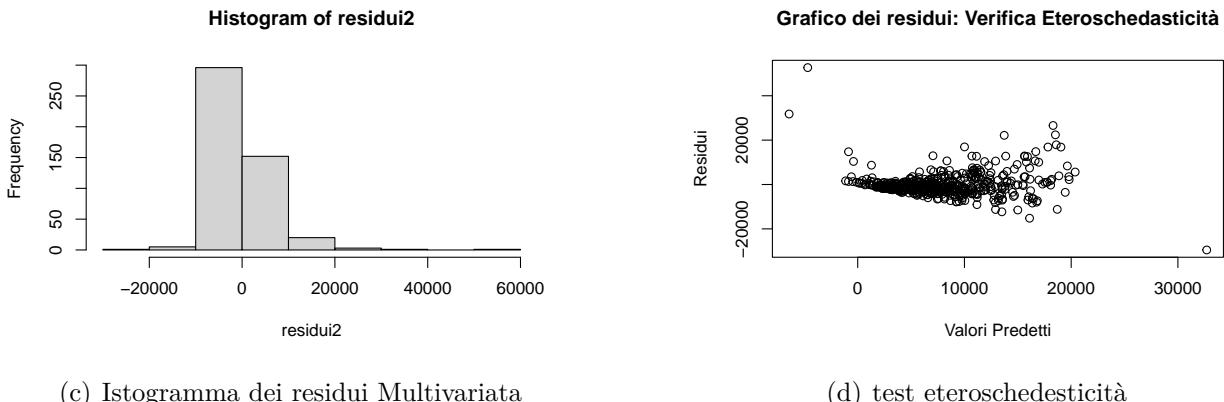
Risultati della Regressione Multivariata¹

$$\text{PrezzoAuto} = \beta_0 + \beta_1 \cdot \text{Anzianità} + \beta_2 \cdot \text{chilometraggio} + \beta_3 \cdot \text{potenza} + \beta_4 \cdot \text{tipo di cambio} + \beta_5 \cdot \text{carburante} + \beta_6 \cdot \text{danni} + \epsilon \quad (2)$$

La seguente tabella riassume i coefficienti stimati per il modello completo, che include tutte le variabili esplicative, e i principali test diagnostici:

¹Inizialmente erano state incluse variabili dummy che indicavano la marca dell'auto: Volkswagen, Peugeot, Renault, Opel, Fiat e Toyota. Tuttavia, non risultando significative, sono state escluse dai risultati principali poiché non contribuiscono in modo rilevante alla spiegazione del modello.

Output	Stima	Significato Economico
Intercetta β_0	11.770 ***	il prezzo di un'auto senza considerare nessun regressore è di €11.770.
Anzianità della macchina β_1	-259,3 ***	Ogni anno che passa la macchina perde valore per €259.
Tipo di cambio (automatico) β_2	2.499 **	Cambio automatico aumenta il prezzo di €2.499 rispetto al manuale.
Potenza (cavalli) β_3	17.73 ***	Ogni cavallo in più aumenta il prezzo di €17.73.
Chilometraggio (km) β_4	-0,0475 ***	Ogni chilometro percorso riduce il prezzo di €0.0475.
Tipo di carburante (diesel) β_5	5.300 ***	Diesel aumenta il prezzo di €5.300 rispetto alla benzina.
Danni non riparati β_6	-2.850 *	Danni riducono il prezzo di €2.850.
$R^2 Adjusted$	0,4141	Percentuale di varianza spiegata dal modello.
Significatività globale	$< 2.2 \times 10^{-16} ***$	Il modello è significativo nel suo complesso.
Test di Shapiro (p-value)	$2.2 \times 10^{-16} ***$	Residui non seguono una distribuzione normale.
Test di Breusch-Pagan (BP)	82,322	Eteroschedasticità presente.
p-value BP	$1,183 \times 10^{-15} ***$	Conferma di eteroschedasticità significativa.
Test di Durbin-Watson (DW)	1,9628	Residui indipendenti.
p-value DW	0,3432	Conferma dell'indipendenza dei residui



Diagnostica del modello

I test diagnostici eseguiti hanno evidenziato alcuni problemi con il modello, tra cui:

- **Eteroschedasticità:** Confermata dal test di Breusch-Pagan (BP = 82,322, p-value = $1,183 \times 10^{-15}$). Per risolvere il problema si opterà per l'utilizzo del metodo Weight Least Squared (WLS).
- **Residui non normali:** Il test di Shapiro ha confermato la non normalità dei residui. Nonostante ciò, la significatività delle variabili principali rimane valida grazie al metodo dei WLS.
- **Indipendenza dei residui:** Confermata dal test di Durbin-Watson (DW = 1.9628, p-value = 0,3432).

Regressione WLS e Correzione dell'Eteroschedasticità

Per affrontare il problema di eteroschedasticità rilevato nel modello OLS, è stata utilizzata una regressione pesata (Weighted Least Squares, WLS). La varianza condizionale è stata modellata come funzione delle variabili esplicative, stimata, e successivamente invertita per calcolare i pesi da utilizzare nel modello WLS.

$$\begin{aligned} PrezzoAuto = \beta_0 + \beta_1 \cdot \text{Anzianità} + \beta_2 \cdot \text{chilometraggio} + \beta_3 \cdot \text{potenza} \\ + \beta_4 \cdot \text{tipo di cambio} + \beta_5 \cdot \text{carburante} + \beta_6 \cdot \text{danni} + \epsilon, \end{aligned} \quad (3)$$

dove i pesi sono definiti come $w_i = \frac{1}{\hat{\sigma}_{\epsilon_i}^2}$.

Di seguito sono riportati i risultati della regressione WLS.

Variabile	Coefficiente	Std. Error	Pr(> t)
Intercetta β_0	6.778	758,5	***
Anzianità della macchina	-162,6	33,11	***
Tipo di cambio (automatico)	190,6	791	0,80965
Potenza (cavalli)	58,69	4,567	***
Chilometraggio (km percorsi)	-0,04919	0,006021	***
Tipo di carburante (diesel)	3.995	641,1	***
Presenza di danni non riparati	-1.575	490,9	**

Diagnostica			
$R^2 Adjusted$	0,5328	Percentuale di varianza spiegata dal modello.	
Residual Standard Error	2,102	Valore standard dei residui.	
Test di Breusch-Pagan (BP)	$1,3616 \times 10^{-6}$	***	p -value = 1 (eteroschedasticità corretta).

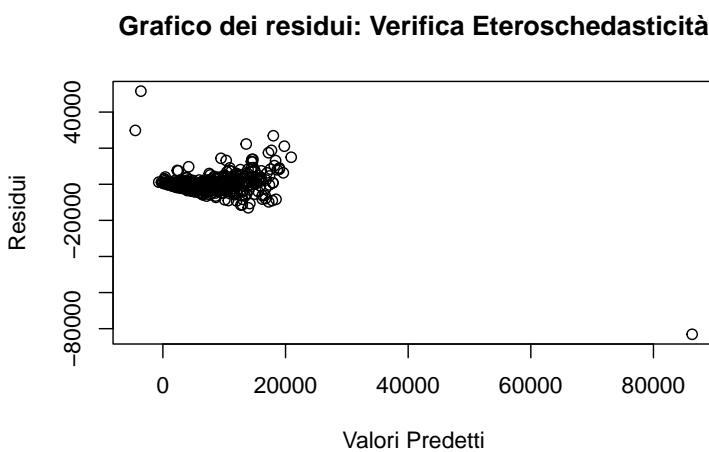


Figura 1: Eteroschedasticità WLS

Conclusioni

Il modello WLS ha corretto efficacemente il problema di eteroschedasticità, come confermato dal test di Breusch-Pagan (p -value = 1). Le variabili anzianità, potenza, chilometraggio, tipo di carburante, e danni sono risultate significative. Tuttavia, il tipo di cambio non ha un impatto statisticamente significativo sul prezzo delle auto usate.

Alla luce del modello, abbiamo confermato le assunzioni principali definite precedentemente.

Applichiamo il modello in un caso reale ipotizzando un veicolo che abbia le seguenti caratteristiche:

1. **Anzianità della macchina:** 6 anni
2. **Tipo di Cambio:** Automatico
3. **Power PS:** 150 cv
4. **Chilometraggio:** 80.000 km
5. **Tipo di carburante:** Benzina
6. **Presenza di Danni non riparati:** SI

$$\begin{aligned} PrezzoAuto = & 6.778 - 162,6 * (6) + 190,6 * (1) + 58,69 * (150) \\ & - 0,04919 * (80.000) + 3.995 * (0) - 1.575 * (1) \end{aligned} \tag{4}$$

Risultati

Una macchina che presenta le caratteristiche assunte, risulta avere un valore stimato pari a circa €9.286,3