

## Prova 7

### PARTE A

I dati MichelinNY.txt si riferiscono a un campione di 164 locali (ristoranti, bar, ...) di New York. Le variabili presenti nel dataset sono le seguenti:

- InMichelin: il locale è presente sulla guida Michelin
- Food: qualità del cibo, espressa su una scala 1-30
- Decor: qualità dell'ambiente, espressa su una scala 1-30
- Service: qualità del servizio, espressa su una scala 1-30
- Price: Prezzo medio a pasto (in US Dollars)

1. Ricorrendo alla notazione matriciale si stimino i coefficienti di regressione con i rispettivi errori standard e p-value per il modello di regressione lineare che spiega la variabile Price in funzione di InMichelin, Food, Decor, Service. Si ripeta la stima usando la funzione lm().
2. Si commentino sia la stima sia la significatività statistica di ciascun coefficiente di regressione stimato al punto precedente.
3. Si svolga un'analisi grafica dei residui, commentando opportunamente i grafici ottenuti, al fine di verificare l'eventuale violazione delle assunzioni di base del modello di regressione lineare stimato.

---

### PARTE B

Il dataset efige.txt (oppure efige.xls) contiene i dati relativi ad un campione di aziende italiane operanti in diversi settori di attività economica. Per ogni azienda vengono fornite una serie di variabili, la cui legenda è disponibile sul file efige.xls. Si selezioni un modello di regressione logit per la variabile risposta binaria deloc\_fdi (nel 2008 l'impresa ha effettuato investimenti diretti esteri), arrivando ad un insieme di variabili esplicative statisticamente significative e fornendo un'interpretazione dell'effetto di ciascuna sulla variabile risposta.

---