

# Esercizio Regressione logistica

*Metodi Quantitativi per Economia,  
Finanza e Management*

*Esercitazione n°12*

# Consegna Lavoro di gruppo

- Scadenza per la consegna del lavoro di gruppo è fissata inderogabilmente per il giorno:

**Giovedì 10 Gennaio 2019**

- La consegna va effettuata entro le **ore 12** alla Sig.ra **Enrica Luezza** (Segreteria 4° Piano)
- Il materiale da consegnare consiste in:
  - stampa cartacea della presentazione in Power Point;
  - Chiavetta USB contenente:
    - questionario;
    - base dati in formato Excel;
    - Script di R;
    - presentazione Power Point.

**N.B.** Il supporto elettronico (chiavetta USB) non sarà restituito

# Metodi Quantitativi per Economia, Finanza e Management

Tipo di analisi	Strumenti	Lavoro di gruppo
ANALISI UNIVARIATA	<ul style="list-style-type: none"> <li>- DISTRIBUZIONI DI FREQUENZA</li> <li>- INDICI DI POSIZIONE (MISURE DI TENDENZA CENTRALE E MISURE DI TENDENZA NON CENTRALE)</li> <li>- INDICI DI DISPERSIONE</li> <li>- MISURE DI FORMA DELLA DISTRIBUZIONE</li> </ul>	<b>OBBLIGATORIO</b>
<b>ANALISI BIVARIATA E TEST STATISTICI PER LO STUDIO DELL'ASSOCIAZIONE TRA VARIABILI</b>	<p><b><u>Due variabili qualitative o quantitative discrete:</u></b>            TABELLA DI CONTINGENZA E INDICI CHI QUADRO E V DI CRAMER            TEST CHI QUADRO PER L'INDIPENDENZA STATISTICA</p> <p><b><u>Due variabili quantitative continue:</u></b>            INDICE DI CORRELAZIONE DI PEARSON (<math>\rho</math>) E COVARIANZA            TEST t PER L'INDIPENDENZA LINEARE</p> <p><b><u>Una variabile qualitativa e una quantitativa continua:</u></b>            INDICE <math>\eta^2</math>            TEST F PER L'INDIPENDENZA IN MEDIA</p>	<b>OBBLIGATORIO</b>
ANALISI MULTIVARIATA	<ul style="list-style-type: none"> <li>- ANALISI FATTORIALE</li> <li>- REGRESSIONE LINEARE</li> <li>- REGRESSIONE LOGISTICA</li> </ul>	<b>OBBLIGATORIO</b>

# Lavoro di gruppo – Schema di valutazione

## Topics

### 1. Introduzione

- 1.1. Definizione Obiettivi di Ricerca
- 1.2. Descrizione del Contesto
- 1.3. Definizione della Popolazione
- 1.4. Disegno del Campione
- 1.5. Fieldwork

### 2. Analisi Preliminari

- 2.1. Controllo Rappresentatività del Campione
- 2.2. Analisi Univariate
- 2.3. Analisi Connessione
- 2.6. Analisi Correlazione
- 2.7. ANOVA

### 3. Analisi Fattoriale

- 3.1. Scelta Numero dei fattori
- 3.2. Interpretazione dei fattori

### 4. Regressione Lineare

- 4.1. Definizione obiettivo di analisi
- 4.2. Scelta variabili di input
- 4.3. Valutazione bontà del modello
- 4.4. Analisi Multicollinearità
- 4.5. Interpretazione del modello

### 5. Regressione Logistica

- 5.1. Definizione obiettivo di analisi
- 5.2. Scelta variabili di input
- 5.3. Valutazione bontà del modello
- 5.4. Analisi Multicollinearità
- 5.5. Interpretazione del modello

### 6. Conclusioni

### 7. Layout

# Esercizio: Regressione Logistica

Il dataset **ct\_pp\_propensity** contiene 23 variabili e 145.401 clienti di una banca.

Di seguito la descrizione di ogni variabile:

#	Variable	Descrizione Variabili
1	ID CLIENTE	Codice identificativo cliente
2	ACQUISTO_PP	Acquisto un prestito personale nel periodo considerato (1=acquistato, 0=non acquistato)
3	FLG_REF_CC	Referente di del conto corrente analizzato
4	FLG_DOMICILIAZIONE_UTENZE	Domiciliazione utenze
5	FLG_POSSESSO_DEBITO	Possesso di Carte di Debito
6	FLG_ACCREDITO_PENSIONE	Accredito dello stipendio e/o della pensione
7	FLG_POSSESSO_BT	Possesso di Finanziamenti BT (a breve termine)
8	FLG_RID_12M	Attivazione di un RID verso un concorrente negli ultimi 12 mesi
9	FLG_CONTO_CORRENTE	Possesso di c/c
10	FLG_POSSESSO_CREDITO	Possesso Carte di Credito
11	ETA	Età in classi
12	N_COINT_POSS_CC	Numero di cointestari del Conto Corrente
13	N_NDG_TOTALI	Numero di contratti attivi per il cliente
15	CAR_NUM_PREL	Numero di prelievi con carta di credito
16	AFI_IMP_TOT_M6	Afi (Attività Fruttifere d'Interessi) medio degli ultimi 6 mesi
17	CS_NUM_CS	Numero di classi di famiglie di prodotti
18	MAR_PCT_IMP_FINNZ_IMPGH_BT_M6	Percentuale media Margini Finanziari Impieghi BT sul totale margini nell'ultimo semestre
19	OPR_IMP_PREL_TOT_M6	Importo medio prelevato nell'ultimo semestre
20	OPR_IMP_SPE_CAR_D_TOT_M6	Importo medio speso mediante l'utilizzo del bancomat nell'ultimo semestre
21	OPR_IMP_USC_TOT_M6	Importo medio delle uscite nell'ultimo semestre
23	OPR_PCT_IMP_USC_INT_M6	Percentuale delle uscite sul canale internet

# Esercizio: Regressione Logistica

1. Cambiare la directory di lavoro in modo che punti alla cartella in cui si è salvato il dataset.
2. Effettuare una regressione logistica utilizzando come variabile dipendente la variabile 'ACQUISTO\_PP' e come potenziali regressori le seguenti variabili:

*FLG\_DOMICILIAZIONE\_UTENZE*

*FLG\_ACCREDITO\_PENSIONE*

*FLG\_POSSESSO\_BT*

*FLG\_RID\_12M*

*FLG\_CONTO\_CORRENTE*

*MAR\_PCT\_IMP\_FINNZ\_IMPGH\_BT\_M6*

*FLG\_POSSESSO\_CREDITO*

*N\_COINT\_POSS\_CC*

*N\_NDG\_TOTALI*

*CAR\_NUM\_PREL*

*AFI\_IMP\_TOT\_M6*

*OPR\_IMP\_PREL\_TOT\_M6*

*OPR\_IMP\_USC\_TOT\_M6*

*OPR\_PCT\_IMP\_USC\_INT\_M6*

# Esercizio: Regressione Logistica

- Qual è la percentuale di clienti che ha acquistato un Prestito Personale?
- Stimare il modello di regressione logistica e valutarne la bontà
- Esiste un problema di multicollinearità? Se sì, risolverlo.
- Quali regressori risultano significativi?
- Quale regressore è più importante nella spiegazione della variabile risposta?
- Quali regressori hanno un impatto positivo e negativo sulla variabile risposta? Qual è la loro interpretazione?