1

OUTPUT 1

(MODELLO C.A.P.M. - Capital Asset Pricing Model)

(MODELLO DI REGRESSIONE LINEARE SEMPLICE)

OUTPUT RIEPILOGO

Statistica della regressione	
R multiplo	0,528519838
R al quadrato	0,279333219
R al quadrato corre	0,273225874
Errore standard	0,050552682
Osservazioni	120

ANALISI VARIANZA

	gdl	SQ	MQ	F	Significatività F
Regressione	1	0,116884923	0,116884923	45,73725439	5,48863E-10
Residuo	118	0,301557693	0,002555574		
Totale	119	0,418442616			

	Coefficienti	Errore standard	Stat t	Valore di significatività	Inferiore 95%	Superiore 95%
Intercetta	-0,000489594	0,004640033	-0,10551514	0,916146214	-0,009678122	0,008698934
EXTRAMARKET	0,456820772	0,067547736	6,762932381	5,48863E-10	0,323057866	0,590583678

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

(di solito, nel C.A.P.M., si omette l'intercetta β_0)

Y = EXTRA-RENDIMENTO MENSILE DI IBM (= REND.MENS.IBM − REND. RISK-FREE)

X = EXTRA-RENDIMENTO MENSILE DI MERCATO (= REND.MENS.MERCATO − REND. RISK-FREE)

β_0, β_1 = PARAMETRI INCOGNITI

ε = "ERRORE" (ALEATORIO)

OUTPUT 2

(MODELLO DI REGRESSIONE LINEARE MULTIPLA)

OUTPUT RIEPILOGO

Statistica della regressione	
R multiplo	0,664421825
R al quadrato	0,441456361
R al quadrato corre	0,439084626
Errore standard	12788,69435
Osservazioni	474

ANALISI VARIANZA

	gdl	SQ	MQ	F	Significatività F
Regressione	2	60884114235	30442057118	186,1322301	2,70364E-60
Residuo	471	77032381201	163550703,2		
Totale	473	1,37916E+11			

	Coefficienti	Errore standard	Stat t	Valore di significatività	Inferiore 95%	Superiore 95%
Intercetta	-20978,3036	3087,257655	-6,795125626	3,28056E-11	-27044,80602	-14911,80119
prevexp	12,07128959	5,81044467	2,077515625	0,038295231	0,653688421	23,48889076
educ	4020,343339	210,6498762	19,08542939	1,42045E-60	3606,413526	4434,273151

- $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

Y = retribuzione annua impiegato

X_1 = esperienza lavorativa (in mesi)

X_2 = anni di scolarità

$\beta_0, \beta_1, \beta_2$ = parametri incogniti

ε = errore (aleatorio)

3) OUTPUT 1 — CAMPIONE DI 120 rendimenti mensili

MODELLO STIMATO: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X$, con (arrotond. al 4° dec.)

$\hat{\beta}_0 = -0.0005$ con $S.E._{\hat{\beta}_0}$ (standard error) $= 0.0046$

$\hat{\beta}_1 = 0.4568$ con $SE_{\hat{\beta}_1}$ (standard error) $= 0.0675$

Interpretazione $\hat{\beta}_1 = 0.4568$: ad un incremento unitario dell'extra-rendimento di mercato è associato un incremento pari a 0.4568, in media, dell'extra-rendimento di IBM.

INTERVALLO DI CONFIDENZA PER β_1, AL 95%:

$$\left(\hat{\beta}_1 - 1.96 \cdot SE_{\hat{\beta}_1}, \hat{\beta}_1 + 1.96 \cdot SE_{\hat{\beta}_1}\right) = (0.4568 - 1.96 \cdot 0.0675, 0.4568 + 1.96 \cdot 0.0675) =$$

$$= (0.3231, 0.5906)$$

TEST DI SIGNIFICATIVITÀ (al 5%):

$$\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

Poiché l'intervallo al 95% non contiene il valore 0, si rifiuta H_0, per cui x è significativa per spiegare/prevedere Y.
(Cioè, il modello è significativo).

SCOMPOSIZIONE DELLA DEVIANZA TOTALE e R^2

$$\sum (y_i - \bar{y})^2 = \sum (y_i - \hat{y}_i)^2 + \sum (\hat{y}_i - \bar{y})^2$$

DEV. TOT. = DEV. RESIDUA + DEV. SPIEGATA

$$0.4184 = 0.3015 + 0.1169$$

$$R^2 = \frac{\text{DEV. SPIEGATA}}{\text{DEV. TOTALE}} = \frac{0.1169}{0.4184} = 0.2793$$

Il 27.93% della variabilità dell'extra-rendimento di IBM è spiegato dall'extra-rendimento di mercato.

OUTPUT 2 (2 decimali)

MODELLO STIMATO: $\hat{y} = -20978.30 + 12.07 \cdot X_1 + 4020.34 \cdot X_2$

INTERV. (AL 95%) PER β_1: $(0.6537, 23.4888)$

INTERV. (AL 95%) PER β_2: $(3606.41, 4434.27)$

I due intervalli non contengono 0 per cui, al 5%, sia X_1 che X_2 sono significative.

INTERPRETAZIONE $\hat{\beta}_1$ e $\hat{\beta}_2$:

$\hat{\beta}_1 = 12.07$. Fissato X_2, ad un incremento unitario di X_1 corrisponde ~~un~~ aumento, in media, di 12.07 di Y.

$\hat{\beta}_2 = 4020.34$. Fissato X_1, ad un incremento unitario di X_2 corrisponde un aumento, in media, di 4020.34 di Y.

Scomposizione devianza totale:

$$1.3792 \times 10^{11} = 77032381201 + 60884114235$$

$$R^2 = \frac{60884114235}{1.3792 \times 10^{11}} = \frac{60884114235}{137920000000} = 0.44$$

Il 44% della variabilità delle retribuzioni è spiegato da esperienza lavorativa e anni di scolarità.

PREVISIONE: previsione della retribuzione di impiegati con 100 mesi di esperienza e 10 anni di scolarità:

$$\hat{y} = -20978.30 + 12.07 \cdot 100 + 4020.34 \cdot 10 = -20978.30 + 1207 + 40203.4 =$$
$$= 20432.1$$

5

TEST

1) Per valutare la significatività di X_1 e X_2 (prese singolarmente), ovvero per verificare

$$\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases} \quad \text{oppure} \quad \begin{cases} H_0: \beta_2 = 0 \\ H_1: \beta_2 \neq 0 \end{cases},$$

si può procedere usando gli intervalli di confidenza (vedi pag. 4), oppure usando il p-value (i due p-value sono entrambi minori di 0.05, per cui si rifiuta H_0 a livello 0.05), oppure con la regione di rifiuto $R_{0.05}$: $|T| = |\hat{\beta}/se_{\hat{\beta}}| >$ quantile di ordine 0.975.

Quindi X_1 e X_2 sono entrambe (individualmente) significative a livello 0.05.

2) Per valutare la significatività globale del modello, ovvero per verificare

$$\begin{cases} H_0: \beta_1 = \beta_2 = 0 \\ H_1: \text{almeno uno tra } \beta_1 \text{ e } \beta_2 \text{ diverso da } 0 \end{cases}$$

si può vedere il p-value del test F, pari a $2.7036 \cdot 10^{-60}$, che è minore di ogni α, per cui si rifiuta H_0 ed il modello è, globalmente, significativo.

La statistica F è calcolata come segue:

$$F = \frac{(\text{DEV. TOT.} - \text{DEV. RES.})/2}{\text{DEV. RES.}/471} = \frac{\text{DEV. SPIEG.}/2}{\text{DEV. RES}/471} =$$

$$= \frac{6088416235/2}{7703238201/471} = 186{,}1322$$

6

La regione di rifiuto del test (test F globale) è, a livello 0.05,

$R_{0.05} = F \geq$ [quantile distr. F con 2 e 471 g.d.l.] $= 3$

Poiché $F_{oss} = 186.1322 > 3$, si rifiuta H_0 a livello 0.05.

UNIVERSITÀ CARLO CATTANEO LIUC LIUC