

Esame di Probabilità, distribuzioni e regressione multipla / Statistica II
Prova parziale di Regressione Multipla **31.01.08**

COGNOME _____ **NOME** _____ **Matr.** _____

Docente: Prof.ssa Cazzaro Prof.ssa Greselin Prof.ssa Pollastri Prof.ssa Zenga

Attenzione: lo studente deve fornire i diversi passaggi dei calcoli eseguiti e i commenti richiesti. Il presente foglio deve essere compilato e riconsegnato. E' vietato l'uso di calcolatrici programmabili o con funzione di agenda elettronica.

ATTENZIONE: Approssimare i calcoli alla quarta cifra decimale.

Un concessionario che tratta anche l'usato, ha rilevato su un collettivo di 9 autoveicoli le seguenti variabili:

- X_1 : fatturato relativo alla vendita del bene (espresso in migliaia di Euro);
- X_2 : valore all'acquisto (espresso in migliaia di Euro);
- X_3 : chilometri percorsi (espressi in migliaia);
- X_4 : anni dalla prima immatricolazione.

Sono disponibili la matrice di varianze-covarianze:

	X_1	X_2	X_3	X_4
X_1	66.2190			
X_2	89.6364	141.2727		
X_3	-143.7810	-67.2727	1 735.5372	
X_4	-16.8512	-16.0000	97.60 3	8.2314

e le medie aritmetiche delle singole variabili:

	X_1	X_2	X_3	X_4
Media	11.5909	23.0000	44.009	3.3636

1) Si calcolino i parametri delle due rette interpolanti a minimi quadrati:

$$I. \hat{X}_1 = a + \alpha_{13} X_3 \qquad II. \hat{X}_1 = b + \alpha_{14} X_4$$

e si individui, tramite opportuni indici, la retta che presenta il migliore adattamento ai dati.

- 2) Si determinino e si interpretino i parametri del piano a minimi quadrati che spiega il fatturato X_1 in funzione del valore all'acquisto X_2 e della variabile esplicativa ritenuta migliore al punto 1).
- 3) Si calcoli la varianza residua del piano determinato al punto 2) e si fornisca una misura della sua bontà di adattamento.
- 4) Si valuti il grado di miglioramento in termini di varianza totale e residua che si ottiene passando dalla retta $I.$ al piano determinato al punto 2).
- 5) Si calcoli la matrice di correlazione. Si valuti poi il coefficiente di correlazione parziale $r_{24,3}$ e lo si confronti con il corrispondente coefficiente grezzo r_{24} , commentando il risultato.
- 6) Sapendo che il coefficiente di correlazione parziale di ordine due, $r_{14,23}$, assume il valore -0.2228, si calcoli l'indice di determinazione multiplo $I^2_{1,234}$ riferito all'iperpiano:

$$\hat{X}_1 = a + \alpha_{12,34} X_2 + \alpha_{13,24} X_3 + \alpha_{14,23} X_4$$

e si commenti il risultato ottenuto.

Esame di Probabilità, distribuzioni e regressione multipla / Statistica II
Prova parziale di Regressione Multipla **31.01.08**

COGNOME _____ **NOME** _____ **Matr.** _____

Docente: Prof.ssa Cazzaro Prof.ssa Greselin Prof.ssa Pollastri Prof.ssa Zenga

Attenzione: lo studente deve fornire i diversi passaggi dei calcoli eseguiti e i commenti richiesti. Il presente foglio deve essere compilato e riconsegnato. E' vietato l'uso di calcolatrici programmabili o con funzione di agenda elettronica.

ATTENZIONE: Approssimare i calcoli alla quarta cifra decimale.

Su un collettivo di 25 aziende manifatturiere vengono rilevati i seguenti caratteri:

- X_1 = fatturato (centinaia di migliaia di Euro)
- X_2 = costo dei dipendenti (centinaia di migliaia di Euro)
- X_3 = costo delle materie prime (centinaia di migliaia di Euro)

Vengono riportate la matrice di correlazione e alcune sintesi dei caratteri:

	X_1	X_2	X_3
X_1	1		
X_2	06814	1	
X_3	01321	01211	1

	Media	Scarto quad. medio
X_1	7076	897
X_2	2544	526
X_3	1025	188

- a) Si determinino i parametri della retta a minimi quadrati $\hat{X}_1 = a + \alpha_{12}X_2$ e se ne forniscano le rispettive interpretazioni. Si calcoli inoltre la bontà di adattamento della retta.
- b) Si determinino i parametri del piano a minimi quadrati $\hat{X}_1 = b + \alpha_{12,3}X_2 + \alpha_{13,2}X_3$ e se forniscano le relative interpretazioni. Si commenti inoltre la differenza tra α_{12} e tra $\alpha_{12,3}$.
- c) Si preveda il fatturato di un'azienda con costo dei dipendenti pari a 25 centinaia di migliaia di Euro ed un costo delle materie prime di 9 centinaia di migliaia di Euro.
- d) Si determini la varianza residua del piano determinato al punto b) e da questa si deduca un indice per valutare la bontà di adattamento del piano.
- e) Si calcoli il coefficiente di correlazione multipla $R_{1,23}^2$ e si interpreti adeguatamente il risultato.
- f) Si valuti il grado di miglioramento, in termini di varianza spiegata e residua, che si ottiene passando dalla retta di cui al punto a) al piano di cui al punto b). Si commentino i risultati ottenuti.
- g) Si calcoli il coefficiente di correlazione $r_{12,3}$ e lo si confronti con r_{12} .

Esame di Probabilità, distribuzioni e regressione multipla / Statistica II
Prova parziale di Regressione Multipla **31.01.08**

COGNOME _____ **NOME** _____ **Matr.** _____

Docente: Prof.ssa Cazzaro Prof.ssa Greselin Prof.ssa Pollastri Prof.ssa Zenga

Attenzione: lo studente deve fornire i diversi passaggi dei calcoli eseguiti e i commenti richiesti. Il presente foglio deve essere compilato e riconsegnato. E' vietato l'uso di calcolatrici programmabili o con funzione di agenda elettronica.

ATTENZIONE: Approssimare i calcoli alla quarta cifra decimale.

- 1) Nei giorni feriali il direttore di un supermercato è di cattivo umore con probabilità pari a 0.6. E' noto che se è di cattivo umore litiga con la cassiera Paola con probabilità pari a 0.9, mentre se non è di cattivo umore la probabilità che litighi con Paola è solo pari a 0.5.
- a) Si stabilisca, motivando la risposta, se gli eventi "Il direttore non è di cattivo umore" e "Il direttore non litiga con Paola" sono indipendenti.
 - b) Si calcoli la probabilità che il direttore sia di cattivo umore sapendo che non ha litigato con Paola.
 - c) Sia X la variabile casuale "Numero di giorni di una settimana di lavoro (6 giorni lavorativi) nei quali il direttore non litiga con Paola". Supponendo che gli eventi "Il direttore non litiga con Paola in un dato giorno della settimana" sono indipendenti ed identicamente distribuiti, si calcoli $P(X > 3)$.
- 2) Si supponga che la perdita X delle aziende di una provincia lombarda appartenente ad un dato settore segua la legge di Pareto con media pari a 25000€ e perdita minima 1000€.
- a) Determinare il primo quartile.
 - b) Determinare la varianza della perdita delle aziende.
 - c) Determinare la percentuale di aziende che presenta una perdita maggiore di 40000€.
 - d) Se un'azienda ha una perdita superiore a 2000€, qual è la probabilità che abbia perso meno di 5000€?

3) Si consideri la seguente funzione

$$f(x) = \begin{cases} 1/3 & 3 < x \leq a, \\ 0 & \text{altrove.} \end{cases}$$

- a) Determinare il valore di a tale per cui $f(x)$ sia una funzione di densità.
- b) Si determini la funzione di ripartizione e il terzo quartile della v.c. X che ha come funzione di densità $f(x)$.
- c) Si considerino le v.c. X_1 e X_2 identicamente distribuite con funzione di densità $f(x)$ e sia $Y = X_1 + X_2$. Sapendo che il coefficiente di correlazione lineare tra X_1 e X_2 è 0.23, si ricavi la media e la varianza di Y .