



**Università della Valle d'Aosta**

**Université de la Vallée d'Aoste**

**Corso di laurea triennale in Economia e Management**

## **Livello di stress durante la sessione di esami**

**Analisi quantitativa delle determinanti del livello di stress percepito  
durante la sessione d'esami: Modello di Regressione Lineare**

**Juglair Kristel Claire**

**METODI QUANTITATIVI – ECONOMIA E MARKETING**

**A.A. 2024-2025**

## Keywords

- Questionario
  - Livelli di stress in sessione
  - Variabili dipendenti, indipendenti
  - Variabili *dummy* e continue
  - Modello di Regressione lineare
  - Significatività statistica
- 

## Abstract

Questo breve elaborato si propone di commentare i risultati di un modello di regressione lineare semplice, applicato a variabili ricavate da un questionario sulla soddisfazione universitaria degli studenti. Il modello indica che il genere e la quantità degli esami svolti al semestre sono variabili significativamente correlate con il livello di stress percepito durante la sessione di esami. Il modello mostra inoltre una capacità esplicativa tutto sommato migliorabile, ma offre comunque degli spunti di riflessione.

---

## Indice

### Premessa

Metodologia di ricerca

Strumenti di elaborazione

### 1. Costruzione del modello

### 2. Analisi del modello

Test T

Analisi dei parametri

Analisi dei residui

Test F

Valutazione della bontà del modello

### Conclusione

### Fonti

---

## Premessa

L'obiettivo di questa ricerca è quello di commentare i risultati forniti dal modello di regressione lineare, sviluppato con *Excel*, su alcuni dati selezionati dal sondaggio sullo [Stile di vita dello Studente universitario](#) (Daglio, Juglair, Massocco, Raffa, Serravalle, 2024), e di formulare in seguito ipotesi e interpretazioni che possano dare spiegazioni sul perché il modello abbia conseguito certi risultati.

Si precisa però che i valori delle variabili selezionate presentano numerose distorsioni che non saranno oggetto dell'elaborato. Pertanto, i risultati conseguiti non sono utilizzabili per eventuali ricerche su come perseguire effettivamente una riduzione dello stress in ambito accademico.

## Metodologia di ricerca

Prima di proseguire con l'analisi, è necessario precisare che il collettivo statistico preso in esame, costituito da 29 individui, è composto principalmente dalle matricole iscritte all'Università della Valle D'Aosta durante l'anno accademico 2023-2024.

Il questionario, elaborato utilizzando la tecnica a clessidra<sup>1</sup>, è stato diviso nelle sezioni descritte come segue:

1. Introduzione – In questa sezione l'utente viene accolto ed informato degli obiettivi di tale questionario e del rispetto della privacy tramite clausola legale (D.L. del 30 giugno 2003, n. 196, in vigore dal 1° gennaio 2004);
2. Metodo di studio ed esami – Questa sezione riguarda l'organizzazione dello studio (tra lezioni e non), delle preferenze di orario e delle modalità, della preparazione agli esami e gestione dello stress ad essi connesso;
3. Sonno – Questa breve sezione ha il fine di rilevare le abitudini legate al sonno per indagare eventuali collegamenti con la progressione del percorso universitario, in particolare sull'influenza della qualità del sonno e sul rendimento Universitario;
4. Abitudini alimentari – Data l'importanza dell'alimentazione e dei suoi effetti sulla salute fisica, si è deciso di prendere in esame le abitudini alimentari in generale degli utenti;
5. Tempo libero – Al fine di proseguire l'analisi sulla salute fisica e mentale degli studenti, questa sezione approfondisce in che modo lo stress viene gestito tramite hobby e impiego del tempo libero;
6. Informazioni generali – Nel rispetto della struttura a clessidra prescelta per il sondaggio, quest'ultima sezione di raccolta dei dati rilevanti dona un'*overview* sull'utente, indagandone le informazioni anagrafiche e la tipologia di studente (frequentante, fuori sede o altro);

---

<sup>1</sup> Struttura che prende in considerazione i livelli medi di attenzione dell'utente, e che per questo pone le domande più "faticose" all'inizio, ovvero quelle la cui risposta richiede maggiore attenzione, per poi terminare con domande che richiedono un minor livello di attenzione.

7. Feedback – Infine qui l'utente ha potuto inviare le proprie risposte includendo un commento personale riguardo eventuali miglioramenti del questionario o altre note da rimarcare.

I parametri utilizzati nel modello che sarà analizzato sono stati selezionati dalla Sezione 2 (*Metodo di studio ed esami*) e dalla Sezione 6 (*Informazioni generali*).

### Strumenti di elaborazione

La realizzazione di tabelle di output di riepilogo del Modello di Regressione Lineare (**MRL**) è stata effettuata tramite il software *Excel*, mentre i grafici di dispersione e di linee di regressione con il linguaggio di programmazione *Python 3.13.3* sul software *VSCode* traendo i dati dalla versione *.csv* del file *Excel*.

## 1. Costruzione del modello

Supponendo di condurre una ricerca sul grado di stress durante il periodo degli esami da parte degli studenti, di voler comprenderne le determinanti e come poter ridurre tale valutazione, come variabile dipendente **y**, è stato scelto il *livello di stress durante la sessione degli esami* indicato dagli utenti alla fine del questionario.

Come variabili indipendenti è stato selezionato come parametro con valori quantitativi (**X**) il *numero di esami dati al semestre*, in questo caso considerato variabile continua. Con la scelta di tale parametro, si cerca di comprendere in che misura la quantità di esami a cui uno studente si prepara influenzi i livelli di stress.

Ai fini del modello si considera anche un parametro qualitativo, ovvero il *genere*, convertito a variabile dicotomica (*dummy*, **D**). In particolare, si assegna valore 0 per i Maschi e valore 1 per le Femmine (*Tabella 1*).

Tabella 1 – Criterio di conversione della variabile qualitativa **Genere** a *dummy*

Variabile D	0	1
Genere	Maschio	Femmina

La combinazione di y, X e D nel Modello di Regressione Lineare genera i dati che saranno oggetto del presente elaborato (*Figura 1*).

Figura 1 – Output di riepilogo del MRL

Statistica della regressione						
R multiplo	0,6278215					
R al quadrato	0,394159835					
R al quadrato corretto	0,347556746					
Errore standard	1,995130736					
Osservazioni	29					
ANALISI VARIANZA						
	<i>gdl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>Significatività F</i>	
Regressione	2	67,33337324	33,66668662	8,45780481	0,00148134	
Residuo	26	103,494213	3,980546652			
Totale	28	170,8275862				
	<i>Coefficienti</i>	<i>Errore standard</i>	<i>Stat t</i>	<i>Valore di significatività</i>	<i>Inferiore 95%</i>	<i>Superiore 95%</i>
Intercetta	10,1400463	2,08213699	4,870018806	4,73587E-05	5,860152418	14,41994017
esami a sem	-1,211805556	0,455324212	-2,661412513	0,013161874	-2,147737878	-0,275873233
Genere F	2,349537037	0,803116421	2,925524838	0,007046377	0,698707591	4,000366483

## 2. Analisi del modello

### Test T

Prima di scomporre il modello nei suoi elementi è necessario fare un breve approfondimento sul test  $t$  e il suo ruolo nella determinazione del livello di significatività.

Il **test  $t$** , che ha lo scopo di valutare se ogni singolo coefficiente stimato nel modello sia significativamente diverso da zero, è un indicatore a stretto contatto con il  $p$ -value.

In particolare, si testa l'ipotesi nulla (*Formula 1*) per cui il coefficiente non ha effetto su  $y$ , contro l'ipotesi alternativa secondo cui il coefficiente ha un effetto significativo.

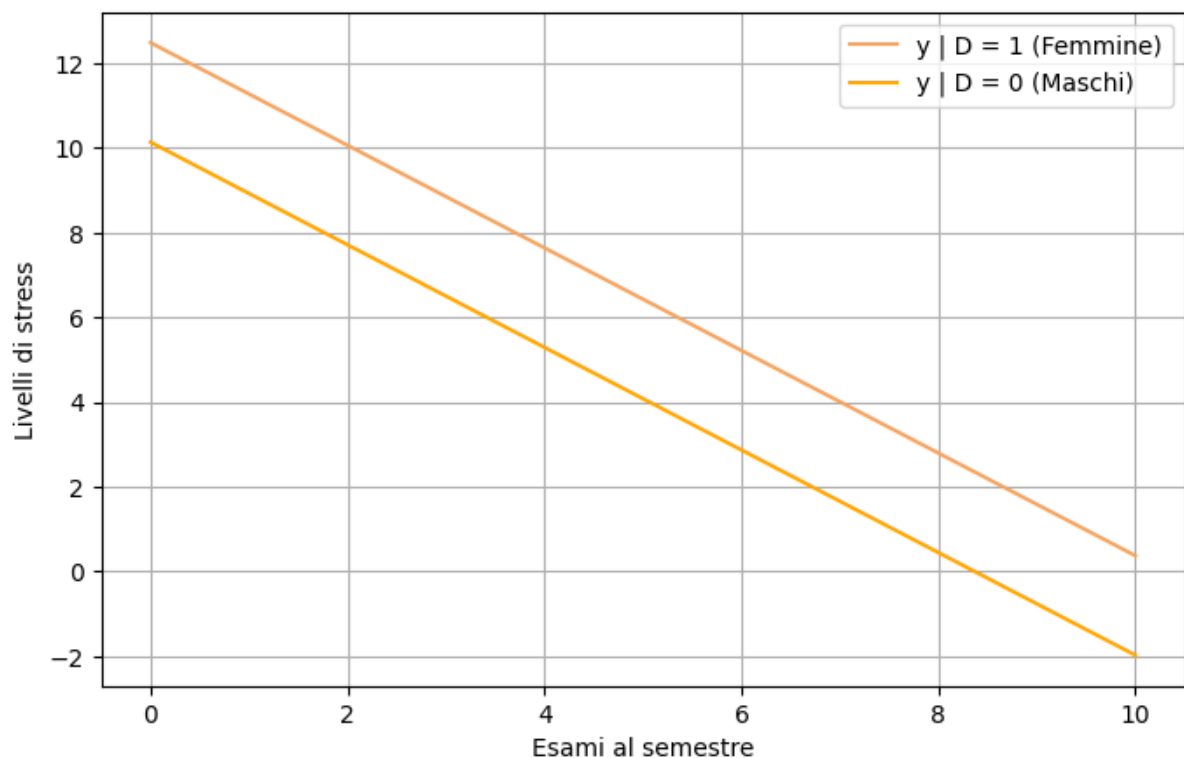
Se il valore  $p$  associato è inferiore al livello di significatività scelto (ad esempio 0,05), si rifiuta l'ipotesi nulla: ciò indica che la variabile  $X_j$  ha un'influenza statisticamente significativa sulla variabile dipendente  $y$ .

[Formula 1]  $H_0: \beta_j = 0$  vs  $H_1: \beta_j \neq 0$

### Analisi dei parametri

Dall'osservazione dei parametri selezionati, è possibile valutarne il livello di significatività statistica (*Tabella 2*), e utilizzare i coefficienti per tracciare un grafico lineare esplicativo e interpretarne il significato (*Grafico 1*).

Grafico 1 – Rette di regressione dei livelli di stress all'aumentare degli esami sostenuti in un semestre.



Il livello di stress percepito durante la sessione degli esami è statisticamente significativo poiché il *p-value* assume valori inferiori all'1%.

Questo parametro rappresenta l'**intercetta** delle rette di regressione, e assume valore  $\beta_0$  di 10.14, ovvero presumibilmente il livello di stress per gli uomini quando non si sostengono esami nel periodo della sessione.

Gli **esami al semestre** sono la variabile continua X, il cui valore medio  $\beta_1$  determina la pendenza delle rette corrispondente a -1.21. Il segno negativo indica un rapporto inverso rispetto allo stress. Ciò significa che più esami vengono dati durante il periodo di sessione, meno l'individuo è sottoposto a condizioni di stress, rapporto controintuitivo di primo acchito, ma su cui si possono formulare differenti ipotesi. Ad esempio, si può pensare che lo stress diminuisca all'aumentare degli esami sostenuti in un semestre grazie ad eventuali doti organizzative e di studio, oppure grazie alla distribuzione degli esami nelle sessioni primaverili ed estive che diminuisce il carico di studio e conseguentemente la pressione accademica.

Anche questa variabile è statisticamente significativa ma solo per  $\alpha < 5\%$  e quindi anche  $\alpha < 10\%$ .

Per quanto riguarda la variabile dicotomica D del **genere**, il coefficiente ha valore positivo e altamente significativo poiché il *p-value* conferma la significatività statistica del parametro (in quanto  $\alpha < 1\%$ ). Il valore di tale coefficiente  $\beta_2$ , corrispondente a circa 2.35 e che esprime l'effetto incrementale della categoria delle Femmine rispetto alla categoria di riferimento (Maschi), suggerisce che le studentesse mediamente percepiscono maggiori livelli di stress di 2.34 punti rispetto ai loro colleghi maschi a parità delle altre condizioni.

Tabella 2 – Confronto dei livelli di significatività dei parametri y, X e D

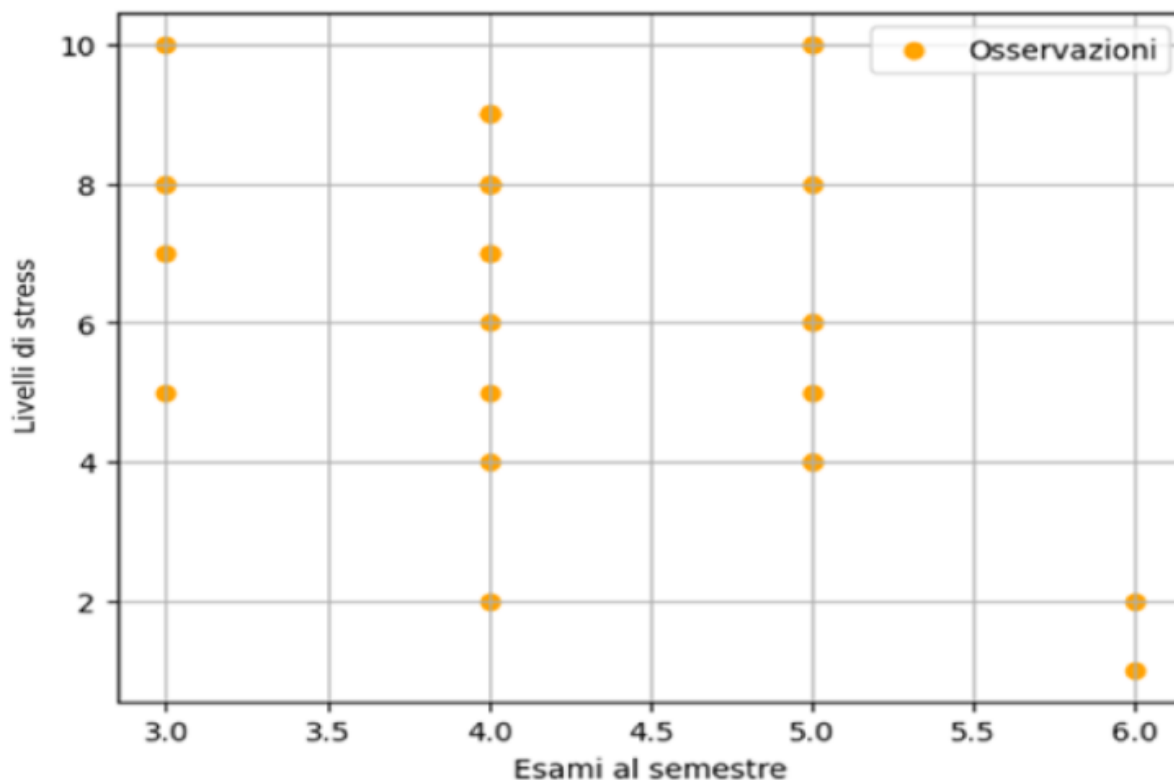
		Intercetta (y)	Esami al semestre (X)	Genere (D)
<b>p-value</b>		0.0000473587	0.013161874	0.007046377
<b><math>\alpha</math></b>	0.10	Stat significativo	Stat significativo	Stat significativo
	0.05			
	0.01		Stat non significativo	

## Analisi dei residui

Eseguendo una rapida analisi visiva delle osservazioni del parametro X (*Grafico 2*), di elevata importanza per l'ipotesi dell'assunzione di **omoschedasticità**, è possibile accertarsi che i residui siano aleatori, senza pattern sistematici.

In effetti possiamo notare dalla distribuzione casuale<sup>2</sup> delle osservazioni che l'assunzione di omoschedasticità è plausibile e che pertanto i risultati conseguiti dal modello non sono "sporcati" da schemi sistematici.

Grafico 2 – Distribuzione delle osservazioni degli esami dati al semestre in relazione ai livelli di stress



## Test F

Prima di proseguire con la valutazione della bontà del modello, è necessario soffermarsi sulla **statistica F**, che testa l'ipotesi nulla (*Formula 2*) per cui tutti i coefficienti siano uguali a 0, contro l'ipotesi alternativa per cui almeno un coefficiente risulti statisticamente significativo.

[Formula 2]  $H_0: \beta_1 = 0, \beta_2 = 0, \dots, \beta_k = 0$  vs  $H_1: \text{almeno un } \beta_j \neq 0$

Nel presente MRL il valore F indica una buona variabilità, poiché assume valori positivi e maggiori di 1. Osservando il  $p\text{-value} = 0.00148$  di F vi è un'ulteriore conferma che il modello è globalmente significativo per  $\alpha < 1\%$ , e di conseguenza anche per  $\alpha < 5\%$  e del 10%.

<sup>2</sup> Si noti che dal numero di osservazioni  $n$  non sia possibile stabilire con certezza che il fenomeno in osservazione abbia distribuzione Normale. Ai fini del commento del presente MRL si presume che la distribuzione sia *Normale* e che quindi l'errore standard sia  $\varepsilon \sim IID(0; \sigma^2)$



## Valutazione della bontà del modello

Per concludere l'analisi del modello, non resta altro da fare che determinare se sia un buon modello. Tale valutazione può essere fatta attraverso l'interpretazione del valore di  $R^2$  ed  $R^2$  corretto.

Il primo, detto anche **coefficiente di determinazione**, è dato dal rapporto tra la regressione della somma dei quadrati delle differenze tra i punti predetti  $\hat{y}_i$  e la media della variabile  $y$  (**ESS** - *Explained Sum of Squares*), e la somma dei quadrati totali delle differenze tra i singoli punti dati  $y_i$  e la media della variabile di risposta (**TSS** - *Total Sum of Squares*), e indica in che misura la variabilità di  $y$  è spiegata dal modello.

Il secondo invece è il valore corretto per il numero di parametri inseriti nel modello ( $k = 3$ ).

In questo caso i valori 0.3941 e 0.3475 rispettivamente per  $R^2$  e  $R^2$  corretto indicano che solo il 39.4% della variabilità del livello di stress percepito è spiegata dal modello e che, aggiustando tale valore, si giunge a un livello di spiegabilità ulteriormente ridotto a circa il 35%.

Questo dato poco promettente indica la necessità di costruzione di un modello predittivo più esaustivo che utilizzi una combinazione di variabili più significativa, ovvero che il modello complessivamente è buono, ma migliorabile.

---

## Conclusione

Dall'analisi del modello di regressione lineare, prendendo come variabile di studio il livello di stress durante la sessione, e le possibili determinanti – genere e numero di esami sostenuti al semestre – è emerso che le ragazze si sentono più sotto pressione nel periodo degli esami e che i soggetti che sostengono più esami si sentono meno stressati, forse per maggiori capacità organizzative, o di alleviamento dello stress in seguito alla graduale diminuzione di esami da sostenere in futuro. Ad ogni modo, il modello si è mostrato essere poco esaustivo e sembra che le variabili selezionate spieghino solo in parte la variabilità dei livelli di stress, proprio perché ci sono altri fattori che possono influenzare la percezione dello stress in sessione.

---

## Fonti

- [Sondaggio: Stile di vita dello Studente universitario \(Daglio, Juglair, Massocco, Raffa, Serravalle, 2024\)](#)

