

LE OPINIONI DEI CONSUMATORI CIRCA LE NUOVE TIPOLOGIE DI ALIMENTAZIONE PER AUTOMOBILI



– GRUPPO –

Codari Stefano (19492)

De Agostini Karin (19507)

Ferrigno Chiara (19853)

Giraldo Andrea (19723)

Analisi Bivariate

Analisi di correlazione lineare

La seguente analisi serve per verificare se esiste correlazione lineare tra le seguenti variabili:

X: Grado di propensione verso le nuove tecnologie e innovazioni

Y: Livello di informazione sulle seguenti tecnologie d'alimentazione: GPL, Metano ed Elettrico

Analizzando la tabella seguente, possiamo osservare i coefficienti di correlazione lineare tra X e le tre diverse Y (una per ogni tecnologia); facendo ciò notiamo che $\rho > 0$, quindi possiamo dire che tra le due variabili c'è una relazione lineare positiva (per tutte e tre le tipologie d'alimentazione).

| Coefficienti di correlazione di Pearson, N = 188 Prob > r con H0: Rho=0 | | | | |
|--|-------------------|-------------------|-------------------|-------------------|
| | PROP_TECNOL | INFORM_GPL | INFORM_METANO | INFORM_ELETT |
| PROP_TECNOL PROP_TECNOL | 1.00000 | 0.27965 0.0001 | 0.34430 <.0001 | 0.40429 <.0001 |
| INFORM_GPL INFORM_GPL | 0.27965 0.0001 | 1.00000 | 0.57790 <.0001 | 0.53379 <.0001 |
| INFORM_METANO INFORM_METANO | 0.34430 <.0001 | 0.57790 <.0001 | 1.00000 | 0.69053 <.0001 |
| INFORM_ELETT INFORM_ELETT | 0.40429 <.0001 | 0.53379 <.0001 | 0.69053 <.0001 | 1.00000 |

Verifichiamo ora se esiste questa correlazione attraverso il

TEST PER L'INDIPENDENZA LINEARE:

H_0 : X e Y sono linearmente indipendenti ($\rho=0$)

H_1 : X e Y sono linearmente dipendenti ($\rho\neq 0$)

Per effettuarlo, utilizziamo la statistica test t di Student:

Dalla tabella vediamo che il p-value è pari a 0,0001, quindi, sia fissando $\alpha = 0,05$ che $\alpha = 0,01$, esso sarà sempre minore; questo confronto tra i due valori ci porta alla seguente conclusione: RIGETTO H_0 , X e Y sono linearmente dipendenti: il test conferma la correlazione lineare positiva indicata dai coefficienti.

La correlazione lineare tra le due variabili ci porta ad affermare che più una persona è propensa all'innovazione e alle nuove tecnologie e più, di conseguenza, è informato sulle nuove tipologie d'alimentazione per automobili.

Questa relazione però, non è molto forte, per nessuna delle tre tipologie d'alimentazione considerate: è vero che i coefficienti lineari trovati sono diversi da 0 ma non presentano valori molto elevati, che si assestano fra 0,3 e 0,4; tra le tre comunque possiamo notare che la più forte si ha con l'informazione sull'elettrico: questo potrebbe essere spiegato dal fatto che si tratta della tecnologia più rivoluzionaria e pertanto potrebbe attirare maggiormente l'attenzione degli appassionati di tecnologia.

Analisi di indipendenza in media

L'analisi seguente ha l'obiettivo di verificare se c'è dipendenza in media tra le variabili X e Y:

X: Livello di informazione sulle seguenti tecnologie d'alimentazione: GPL, Metano ed Elettrico

Y: Sesso

Possiamo osservare dalle tabelle seguenti che i valori di eta-quadro sono tutti e tre di poco superiori a 0 e questo potrebbe indicare l'assenza di dipendenza in media tra X e Y; dovremo affidarci al test d'ipotesi ma in caso di dipendenza gli indici dimostrano che la relazione tra le due variabili sarebbe debole, per tutte e tre le tipologie d'alimentazione, ma tra di esse la meno debole sarebbe quella tra sesso e informazione sull'alimentazione GPL.

La media tra i gruppi essendo sempre più alta per i consumatori di sesso maschile, sembrerebbe invece indicare l'esistenza di una relazione di dipendenza, anche di moderata intensità.

| R-quadro | Var coeff | Radice MSE | Media di INFORM_GPL |
|----------|-----------|------------|---------------------|
| 0.116298 | 49.50689 | 2.849279 | 5.755319 |

| Livello di SESSO | N | INFORM_GPL | |
|------------------|-----|------------|------------|
| | | Media | Dev std |
| F | 75 | 4.49333333 | 2.82543148 |
| M | 113 | 6.59292035 | 2.86492698 |

| R-quadro | Var coeff | Radice MSE | Media di INFORM_METANO |
|----------|-----------|------------|------------------------|
| 0.064225 | 56.13171 | 2.809571 | 5.005319 |

| Livello di SESSO | N | INFORM_METANO | |
|------------------|-----|---------------|------------|
| | | Media | Dev std |
| F | 75 | 4.10666667 | 2.76841907 |
| M | 113 | 5.60176991 | 2.83643332 |

| R-quadro | Var coeff | Radice MSE | Media di INFORM_ELETT |
|----------|-----------|------------|-----------------------|
| 0.089071 | 58.52788 | 2.733377 | 4.670213 |

| Livello di SESSO | N | INFORM_ELETT | |
|------------------|-----|--------------|------------|
| | | Media | Dev std |
| F | 75 | 3.62666667 | 2.51360263 |
| M | 113 | 5.36283186 | 2.86936388 |

TEST PER L'INDIPENDENZA IN MEDIA

Ho: X e Y sono indipendenti in media

H1: X e Y sono dipendenti in media

Utilizziamo la statistica F di Fisher:

per tutte e tre le tipologie d'alimentazione (GPL, Metano, Elettrico) il p-value è minore di α (sia fissando $\alpha = 0,05$ che $\alpha = 0,01$)

quindi RIGETTO Ho: X e Y sono dipendenti in media.

Il test appena effettuato conferma che esiste una relazione di dipendenza in media tra le due variabili; osservando le medie infatti notiamo come le persone di sesso maschile siano notevolmente più informate rispetto a quelle di sesso femminile, su tutte e tre le tipologie d'alimentazione analizzate; questo fatto è sicuramente spiegato dal maggiore interesse verso l'argomento da parte degli intervistati di sesso maschile.

| Origine | DF | Somma dei quadrati | Media quadratica | Valore F | Pr > F |
|-----------------|-----|--------------------|------------------|----------|--------|
| Modello | 1 | 198.723678 | 198.723678 | 24.48 | <.0001 |
| Errore | 186 | 1510.021003 | 8.118392 | | |
| Totale corretto | 187 | 1708.744681 | | | |

| Origine | DF | Somma dei quadrati | Media quadratica | Valore F | Pr > F |
|-----------------|-----|--------------------|------------------|----------|--------|
| Modello | 1 | 100.768368 | 100.768368 | 12.77 | 0.0004 |
| Errore | 186 | 1468.226313 | 7.893690 | | |
| Totale corretto | 187 | 1568.994681 | | | |

| Origine | DF | Somma dei quadrati | Media quadratica | Valore F | Pr > F |
|-----------------|-----|--------------------|------------------|----------|--------|
| Modello | 1 | 135.882631 | 135.882631 | 18.19 | <.0001 |
| Errore | 186 | 1389.670560 | 7.471347 | | |
| Totale corretto | 187 | 1525.553191 | | | |

Analisi di correlazione lineare

La seguente analisi ha lo scopo di indagare se c'è correlazione lineare tra le seguenti variabili X e Y:

X: Willness to pay per un auto dotata di alimentazione a basso impatto

Y: Età

Quest'analisi è stata effettuata (solamente per quanto riguarda la variabile Willness to pay), su un dataset che è già stato ripulito dei valori considerati outliers.

Il coefficiente di correlazione lineare tra X e Y mostrato in tabella è pari a $-0,25748$; esso indica una relazione lineare negativa tra le due variabili analizzate, anche se di debole intensità.

| Coefficienti di correlazione di Pearson Prob > r con H0: Rho=0 Numero osservazioni | | |
|--|---------------------------|---------------------------|
| | WTP_BASSO_IMPATTO | ETA |
| WTP_BASSO_IMPATTO WTP_BASSO_IMPATTO | 1.00000 135 | -0.25748 0.0026 135 |
| ETA ETA | -0.25748 0.0026 135 | 1.00000 159 |

Verifichiamo l'esistenza della relazione lineare negativa attraverso il TEST PER L'INDIPENDENZA LINEARE

Ho: X e Y sono linearmente indipendenti ($\rho=0$)

H1: X e Y sono linearmente dipendenti ($\rho \neq 0$)

Utilizziamo il test t di Student e vediamo che:

Confrontando il p-value pari a 0,0026 con α (sia essa = 0,05 o =0,01) esso sarà minore, quindi RIGETTO Ho: X e Y sono linearmente dipendenti.

Il test conferma che esiste una relazione tra le due variabili considerate.

Dato che la relazione lineare tra X e Y è negativa, man mano che l'età aumenta, la disponibilità a spendere per un'automobile a basso impatto ambientale, diminuisce.

Questa conclusione dimostra che le persone più giovani sono maggiormente interessate e più sensibili all'aspetto ambientale.



Analisi di indipendenza in media

L'analisi seguente verifica se c'è dipendenza in media tra le variabili X e Y:

X: Willness to pay per un auto dotata di alimentazione a basso impatto

Y: Sesso

Quest'analisi è stata effettuata (solamente per quanto riguarda la variabile X), su un totale di risposte che è già stato ripulito dei valori considerati outliers.

Possiamo osservare dalla tabella seguente che il valore di eta-quadro è pari a 0,061083: questo sembrerebbe indicare indipendenza in media tra X e Y; la relazione è molto scarsa in quanto l'indice supera di poco lo 0. La media della willness to pay tra i due gruppi è simile anche se si evidenzia una maggiore propensione da parte delle acquirenti di sesso femminile.

| R-quadro | Var coeff | Radice MSE | Media di WTP_BASSO_IMPATTO |
|----------|-----------|------------|----------------------------|
| 0.061083 | 78.89776 | 9.861051 | 12.49852 |

| Livello di SESSO | N | WTP_BASSO_IMPATTO | |
|------------------|----|-------------------|------------|
| | | Media | Dev std |
| F | 53 | 15.6037736 | 10.0696414 |
| M | 82 | 10.4914634 | 9.7247832 |

Verifichiamo l'esistenza di questa relazione attraverso il seguente test:

TEST PER L'INDIPENDENZA IN MEDIA

Ho: X e Y sono indipendenti in media

H1: X e Y sono dipendenti in media

Utilizziamo la statistica F di Fisher:

Il p-value è pari a 0,0039 quindi sia fissando $\alpha = 0,05$ che $\alpha = 0,01$ esso risulta minore di α : quindi RIGETTO Ho:
X e Y sono dipendenti in media.

| Origine | DF | Somma dei quadrati | Media quadratica | Valore F | Pr > F |
|-----------------|-----|--------------------|------------------|----------|--------|
| Modello | 1 | 841.37643 | 841.37643 | 8.65 | 0.0039 |
| Errore | 133 | 12932.96327 | 97.24033 | | |
| Totale corretto | 134 | 13774.33970 | | | |

Nonostante il valore di eta-quadro molto vicino a 0 il test stabilisce che c'è una relazione di dipendenza in media tra la disponibilità a spendere per un'automobile a basso impatto ambientale e il sesso. La relazione tuttavia è piuttosto debole, osservando le medie si nota infatti che la willingness to pay delle donne è solo di pochi punti percentuali più alta.

Analisi di indipendenza statistica

Attraverso quest'analisi si vuole valutare se c'è indipendenza statistica tra le seguenti variabili:

X: Tecnologia scelta al momento dell'acquisto

Y: Sesso

La V di Cramer, come mostrato nella tabella, è pari a 0,1666: essendo molto vicino a 0 potrebbe indicare indipendenza statistica tra X e Y.

L'indice Chi-Quadro assume invece un valore pari a 5,2189: essendo abbastanza basso potrebbe indicare indipendenza statistica tra le due variabili, tuttavia essendo influenzato dal numero di osservazioni non ci permette di esprimere un giudizio attendibile.

Anche l'osservazione della tabella di contingenza non sembra indicare dipendenza statistica.

| Statistica | DF | Valore | Prob |
|-----------------------------|----|--------|--------|
| Chi-quadrato | 4 | 5.2189 | 0.2656 |
| Chi-quadrato rapp verosim | 4 | 5.2113 | 0.2663 |
| Chi-quadrato MH | 1 | 0.0426 | 0.8366 |
| Coefficiente Phi | | 0.1666 | |
| Coefficiente di contingenza | | 0.1643 | |
| V di Cramer | | 0.1666 | |

| Tabella di SESSO per TECNOL_SCELTA | | | | | | |
|------------------------------------|------------------------------|-------------------------------|------------------------------|-------------------------------|------------------------------|---------------|
| SESSO(SESSO) | TECNOL_SCELTA(TECNOL_SCELTA) | | | | | Totale |
| | BENZINA | DIESEL | ELETTRICO | GPL | METANO | |
| F | 11 5.85 14.67 45.83 | 23 12.23 30.67 34.85 | 14 7.45 18.67 51.85 | 14 7.45 18.67 31.11 | 13 6.91 17.33 50.00 | 75 39.89 |
| M | 13 6.91 11.50 54.17 | 43 22.87 38.05 65.15 | 13 6.91 11.50 48.15 | 31 16.49 27.43 68.89 | 13 6.91 11.50 50.00 | 113 60.11 |
| Totale | 24 12.77 | 66 35.11 | 27 14.36 | 45 23.94 | 26 13.83 | 188 100.00 |
| Frequenza mancanti = 1 | | | | | | |

Andiamo ora a verificare se il test conferma l'indipendenza statistica ipotizzata:

TEST PER L'INDIPENDENZA STATISTICA

Ho: X e Y sono statisticamente indipendenti

H1: X e Y sono statisticamente dipendenti

Avendo un p-value pari a 0,2656, sia fissando $\alpha = 0,05$ che $\alpha = 0,01$, il p-value sarà maggiore di α per cui
ACCETTO Ho: X e Y sono statisticamente indipendenti: la statistica Chi-Quadro che abbiamo utilizzato per questo test ci conferma l'indipendenza statistica ipotizzata in precedenza.

Possiamo quindi concludere affermando che non c'è alcuna associazione tra le due variabili considerate; pertanto la tecnologia d'alimentazione per l'automobile scelta non è influenzata dal sesso.

Riepilogo

In queste analisi abbiamo visto la willingness to pay in relazione ad altre variabili. È emerso che i consumatori giovani e quelli di sesso femminile sono i più propensi a spendere, pertanto suggeriamo che le campagne di marketing vengano sviluppate tenendo conto di questo aspetto.

Nonostante sia ben noto che il pubblico maschile sia maggiormente interessato all'argomento le analisi hanno evidenziato che è meno propenso a spendere, inoltre non c'è differenza fra la tipologia che sceglierebbe in media un uomo e quella che sceglierebbe una donna. Comunque gli uomini si confermano più informati sull'argomento rispetto alle donne.

Sempre a riguardo del livello d'informazione esso si conferma più alto nei soggetti che già presentano una personale propensione verso le tecnologie e le innovazioni.

Nelle successive analisi vedremo come proprio il livello di conoscenza dei consumatori emergerà come fattore di notevole rilevanza.