

# Conjoint Analysis

Corso di Statistica per le ricerche di mercato  
Prof.ssa Simona Balbi

Dott.ssa Agnieszka Stawinoga  
[agnieszka.stawinoga@unina.it](mailto:agnieszka.stawinoga@unina.it)

Dott.ssa Maria Spano  
[maria.spano@unina.it](mailto:maria.spano@unina.it)


Dipartimento di Scienze Economiche e Statistiche  
Università degli Studi di Napoli Federico II

---

La **Conjoint Analysis** – o analisi congiunta – è una tecnica di analisi multivariata che consente di misurare l'importanza relativa di una serie di attributi di un *prodotto o servizio*, fornendo anche indicazioni sul gradimento di vari livelli degli attributi stessi.

*La principale caratteristica della conjoint analysis consiste nella richiesta ai rispondenti di espressioni delle loro scelte con modalità simili a quelle adottate nell'ambito dei reali processi d'acquisto, ossia confrontando le diverse caratteristiche di un prodotto e fornendo un giudizio complessivo sulle offerte proposte.*

(Luce, Tukey, (1964) Green P.E., Rao V.R.(1971); Green P.E.; Wind Y. (1975); Green P.E.; Srinivasan V.(1978, 1990)



I fondamenti concettuali della Conjoint Analysis risiedono nelle considerazioni sviluppate da Lancaster in margine alla teoria del consumatore secondo le quali ***"l'utilità d'uso di un bene deriva dalle singole caratteristiche che lo compongono"***.

La Conjoint Analysis consente di decomporre le valutazioni globali in tante utilità separate che traggono origine dalle diverse caratteristiche di quel bene.

Il focus della tecnica è nella misurazione delle preferenze espresse dai consumatori per combinazioni di livelli degli attributi.

# PRESUPPOSTO TEORICO

I consumatori, quando acquistano un prodotto o un servizio ad elevato coinvolgimento psicologico, si comportano in modo rigorosamente razionale e valutano con attenzione le caratteristiche chiave del prodotto/servizio.

## SCOPO

Individuare tra tutte le caratteristiche di un bene, quale influenza maggiormente il consumatore nella scelta d'acquisto, ed in particolare, quale dei vari livelli delle caratteristiche influenza positivamente e quale negativamente la scelta d'acquisto

## RISULTATO

Conoscendo l'importanza di ogni attributo, e conoscendo i livelli degli attributi considerati positivamente o negativamente nella scelta d'acquisto, è possibile definire una sorta di prodotto ideale

# Obiettivi della Conjoint analysis

In base alla “necessità”:

1. Segmentazione del mercato;
2. Decisioni relative al prodotto;
3. Analisi competitiva;
4. Decisioni sul prezzo;
5. Decisioni sulla promozione e sulla distribuzione

# Segmentazione del Mercato

Esistono forme diverse di segmentazione del mercato che si differenziano in relazione alla base utilizzata per suddividere il mercato stesso.

Generalmente si considerano i benefici che il prodotto offre ai clienti e come questi ultimi li percepiscono (approccio diretto ed orientato al consumatore).

La segmentazione a posteriori, creando dei cluster (gruppi) individuati in base alla similarità dei benefici (i benefici percepiti dai clienti sono numericamente rappresentati dai pesi ottenuti con la Conjoint Analysis).

# Decisioni relative al prodotto

Questa metodologia risulta utile anche per ottenere informazioni e suggerimenti di supporto a decisioni del tipo:

- modifiche ad un altro prodotto, come la modifica di alcuni attributi;
- progettazione ottimale di un prodotto, in questo caso vengono presi in considerazione tutti gli attributi;
- riprogettazione della linea;
- valutazione dei concetti di un nuovo prodotto.

## **Analisi Competitiva**

Permette di prevedere i cambiamenti, per esempio, della quota di mercato derivanti da variazioni apportate a prodotti o a linee di prodotto. Questo è possibile tramite dei modelli di simulazione basati sui dati ottenuti tramite la Conjoint Analysis.

## **Decisioni sul prezzo**

Tramite modelli di Conjoint Analysis abbastanza sofisticati è possibile ottenere informazioni e stime dell'elasticità rispetto al prezzo, utili sia dove non vi siano dati storici sia per supportare decisioni di marketing.

# Decisioni sulla promozione e sulla distribuzione

Questi sono scopi che possono essere definiti **indiretti** della Conjoint Analysis.

Infatti, tramite questa metodologia si possono ricavare informazioni utili anche rispetto a queste categorie di decisioni, per fare questo si possono utilizzare come attributi variabili relative ad aspetti promozionali, come pubblicità, modalità di pagamento,... rilevando i giudizi che i distributori danno a certe combinazioni.

Si possono, inoltre, analizzare alcuni problemi relativi alla distribuzione come i tempi di trasporto o la disponibilità del prodotto.

# Caratteristiche della Conjoint Analysis

-**la sua natura "decompositiva"**: da un giudizio su un prodotto intero si ricavano le preferenze per ogni attributo particolare e per ogni suo livello, ossia dei "part-worth" che, combinati tra loro secondo una certa regola (es. additiva), più si avvicinano al giudizio globale del cliente;

-le stime possano essere fatte a **livello individuale**, così da ricavare un modello predittivo per ogni consumatore esaminato; in effetti, generalmente ci sono molte variazioni nelle preferenze tra persone diverse;

-**la flessibilità** nelle relazioni tra variabili dipendenti e indipendenti: la relazione esistente non è definita a priori (in genere negli altri metodi si assume una relazione lineare semplice). Si possono fare stime separate degli effetti per ogni livello della variabile indipendente senza assumere che esista un legame tra livelli. La flessibilità dipende così dalle scelte fatte dall'analista (come dice Louviere, '94), non esiste un approccio unico e la conjoint analysis è "ciò che l'analista stesso fa").

# Terminologia della Conjoint Analysis

-I **fattori** sono le variabili che il ricercatore controlla per misurare l'effetto su un'altra variabile.

Il fattore è la variabile che rappresenta un certo **attributo** o **caratteristica** del prodotto/servizio in esame.

Nella COA i fattori spesso sono variabili non metriche e devono essere espresse mediante due o più valori chiamati livelli.

-Il **livello** è ogni possibile valore che può assumere un certo fattore.

-Il **profilo o combinazione o stimolo** è una specifica combinazione di livelli degli attributi.

-I **pesi parziali** sono i giudizi su un particolare livello di un attributo ricavabili dal peso complessivo attribuito dal singolo rispondente alle alternative di prodotto.

-**La misura della preferenza** rappresenta la variabile di osservazione. Questa misura può essere realizzata tramite due diverse modalità:

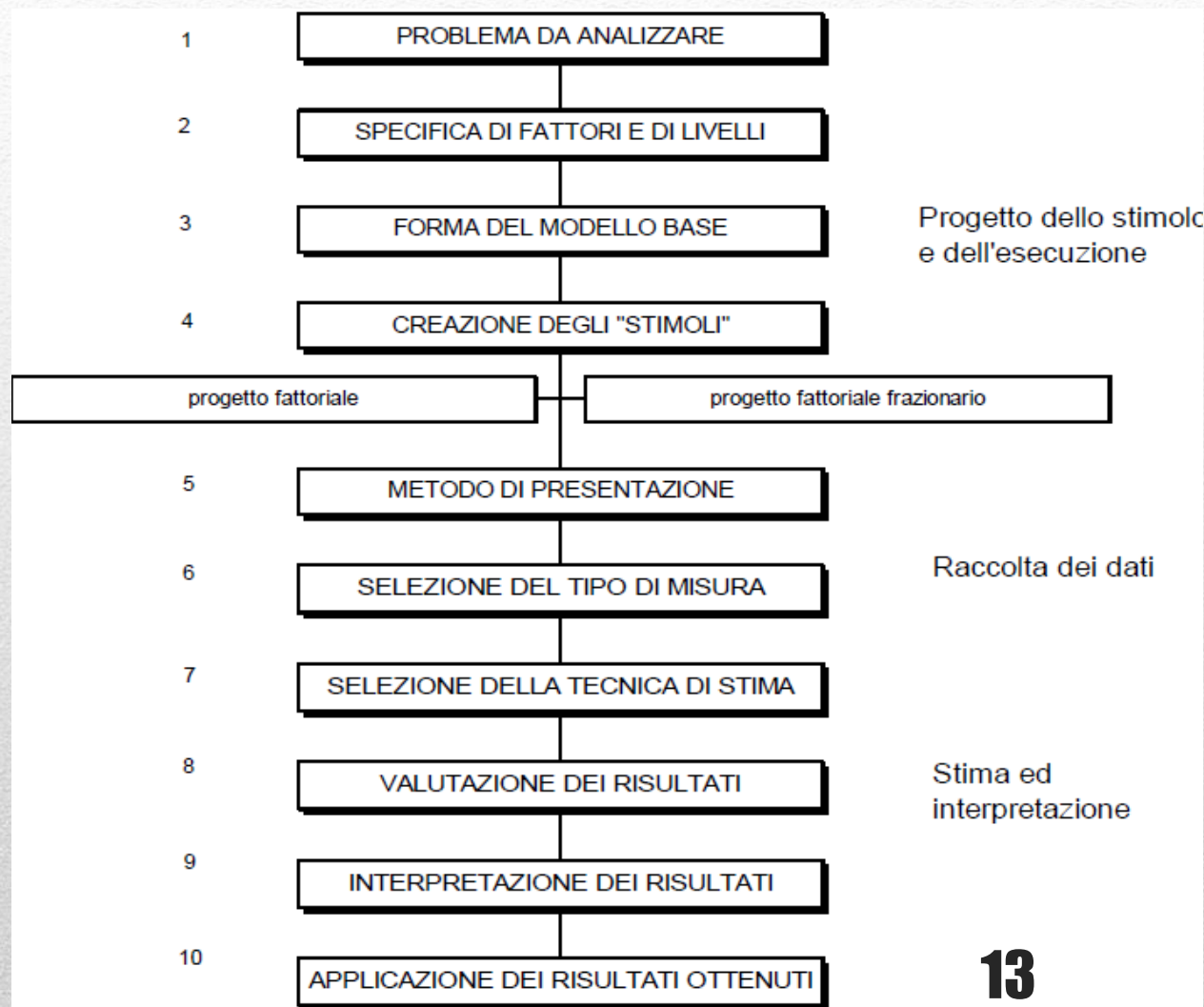
-**Rating** prevede la formulazione da parte del consumatore di un punteggio di utilità o preferenza;

-**Ranking** prevede che il rispondente componga un ordinamento sempre in base alla preferenza o all'utilità percepita.

# Fasi operative della Conjoint Analysis

Consideriamo ora la **sequenza di attività** necessarie per ottenere una Conjoint Analysis, così come suggerito da Hair et al.(1992).

Si tratta di uno schema molto dettagliato, che comprende tutti i passi da compiere per un'analisi completa, suggeriti già da Green e Srinivasan ('78).



# Il problema

Prima di pianificare l'analisi vera e propria è necessario inquadrare il tipo di problematica che si andrà a valutare. Attraverso una ricerca "**esplorativa**" (focus group, interviste telefoniche, interviste "in profondità"), è possibile capire come il consumatore identifichi ed esprima **a parole** certe caratteristiche del prodotto.

Alcuni quesiti possono essere:

- Si possono descrivere tutti gli attributi che danno utilità e valore al prodotto in esame?
- Qual è il criterio di scelta base per questo particolare tipo di prodotto?
- Si deve dare precedenza alla spiegazione del processo di scelta da parte del consumatore (focalizzarsi quindi sull'importanza degli attributi), oppure si cerca di prevedere la quota di mercato per il prodotto?

# Selezione di fattori e livelli (1)

Nella scelta degli attributi rilevanti si devono considerare:

- il **valore globale del prodotto**: vanno inclusi tutti gli attributi che creano potenziale valore nel prodotto e anche quelli che eventualmente ne tolgono, permettendo così una valutazione il più possibile corretta;
- utilizzare attributi e livelli **realistici**, quindi scegliere quelle caratteristiche del prodotto effettivamente realizzabili

Vanno poi selezionati i **livelli** per ogni fattore, tenendo presente che:

- le **caratteristiche selezionate** devono esprimere un concetto ben preciso, il significato dell'una **non** deve "sconfinare" in quello di un'altra;
- il **concetto** deve essere ben espresso per evitare fraintendimenti;
- il **numero di livelli** va quanto più possibile bilanciato per i diversi fattori;

## Selezione di fattori e livelli (2)

- **non** utilizzare livelli che risultino **l'accezione positiva e negativa** di uno stesso attributo, per evitare che l'intervistato si orienti verso la scelta ovvia del livello positivo;
- bisogna evitare la "**collinearità multipla**" tra attributi. In caso di multicollinearità, si può creare un "superattributo" che racchiuda gli aspetti degli attributi correlati, oppure si deve eliminare uno dei fattori.
- Selezionare un numero di attributi e livelli contenuto, in modo da poter condurre l'analisi con un numero di "profili" non troppo elevato, per permettere al consumatore di effettuare una scelta significativa.

# Specificazione del modello di preferenza

La Conjoint Analysis mette in corrispondenza biunivoca le nozioni di **preferenza** e di **utilità**: *quanto più un prodotto incontrerà il gradimento di un consumatore tanto più la sua fruizione fornirà utilità.*

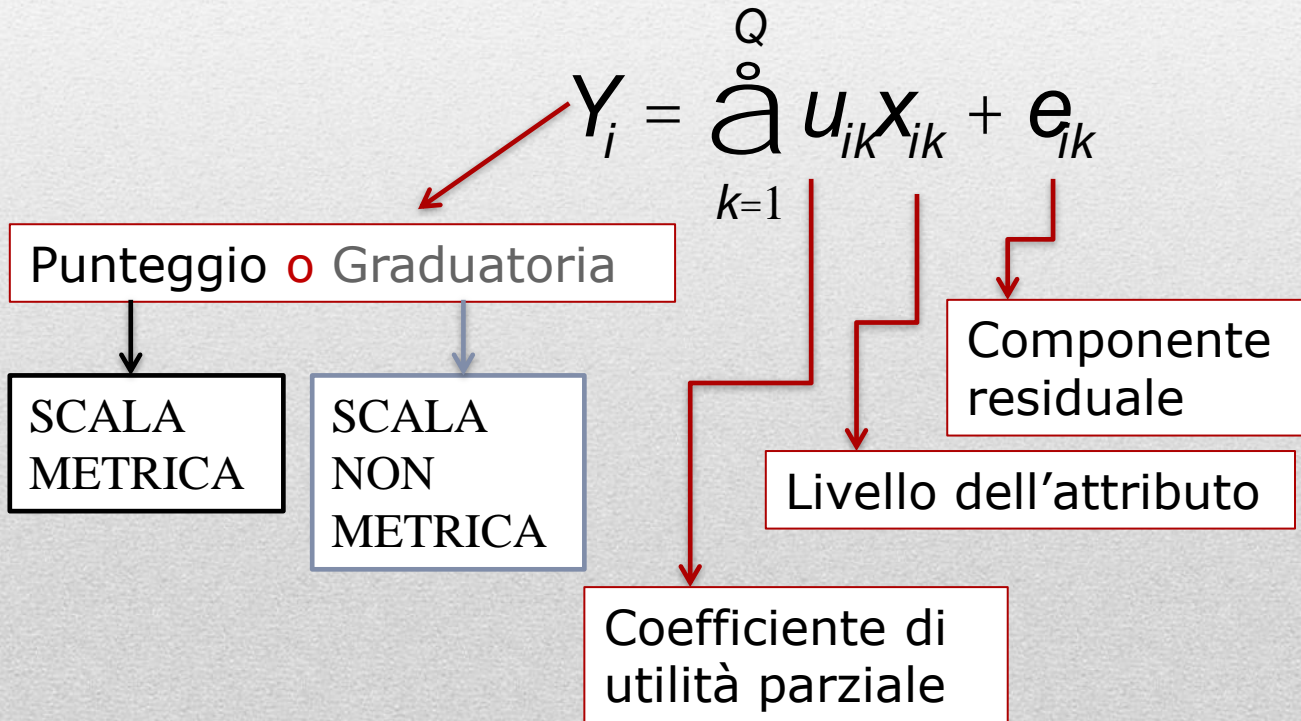
La **preferenza** può essere interpretata, a sua volta, come una funzione delle modalità o dei livelli delle caratteristiche rilevanti del prodotto/servizio.

Dal giudizio del consumatore sul prodotto globale (preferenza) si calcolano le **UTILITA' PARZIALI** ovvero le importanze associate ai livelli degli attributi

**L'UTILITA' TOTALE** è considerata **SOMMA** o **PRODOTTO** delle utilità parziali (modello additivo o moltiplicativo)

# Il modello additivo

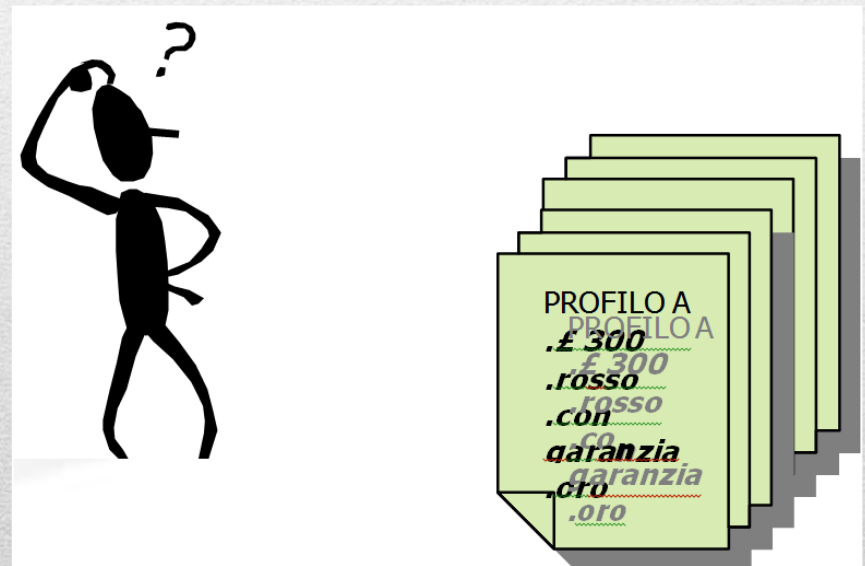
Il modello più utilizzato è quello additivo, nel quale le utilità parziali dei singoli livelli di ogni attributo vengono sommate per ottenere l'utilità complessiva di un prodotto.



## Il metodo di raccolta dati: FULL PROFILE

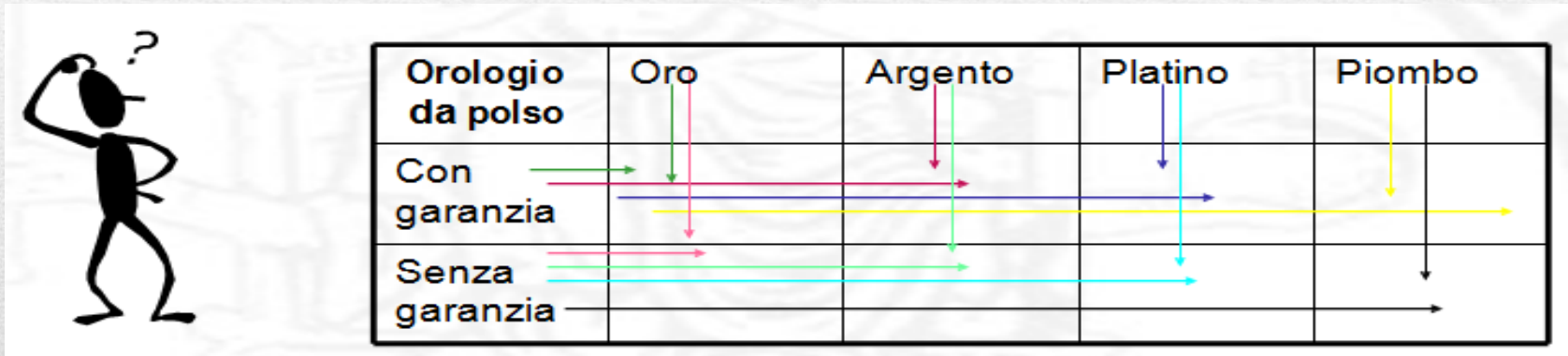
I profili di prodotto da sottoporre a giudizio di preferenza, sono ottenuti come combinazioni di tutti gli attributi e livelli, si utilizzano però "**Disegni fattoriali frazionati**" per ridurre il numero di profili

*Ai consumatori viene chiesto di giudicare contemporaneamente (con una graduatoria o un punteggio) tutti i profili proposti*

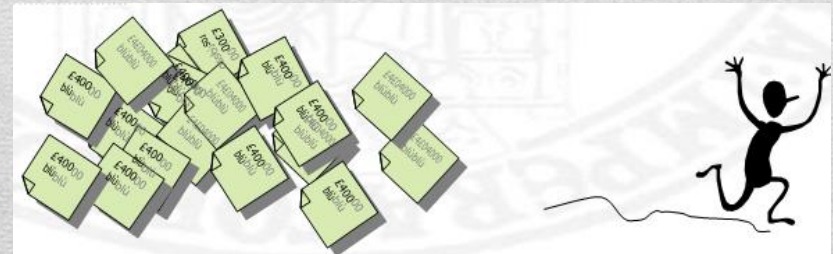


# Il metodo di raccolta dati: TRADE-OFF

Si considerano **due attributi per volta**, e viene chiesto ai rispondenti di esprimere un giudizio di preferenza sui livelli dei due attributi



*Anche con un numero limitato di attributi e livelli, le combinazioni da giudicare sono numerose!!!*



# Scelta del tipo di misura di preferenza

- In generale se i fattori sono pochi (6 o meno) si usa (come suggerito da Green e Srinivasan, '78) il metodo full profile; se i fattori sono in numero da 7 a 10 si ricorre al trade-off.
- Il metodo **trade-off** permette solo un giudizio di tipo numerico (es. su una scala da 1 a 10), mentre con il **full-profile** si possono anche ordinare le combinazioni presentate secondo il giudizio di preferenza (dallo stimolo che piace meno a quello preferito).
- Il metodo **full-profile** è più facile per l'intervistato e più flessibile (si può fare una stima con tipi diversi di regole di composizione), ma l'analisi poi è più difficile da gestire poiché si deve mettere a punto un'intervista personale alla volta, vista la soggettività dell'ordinamento dei cartellini. Nel caso di giudizi numerici, è buona cosa proporre molte categorie di risposta (es. da 1 a 11) se ci sono molti stimoli da valutare.

# Selezione della tecnica di stima

In generale, Green e Srinivasan ('78) suddividono i metodi di stima in tre classi:

- Metodi in cui si assume che la variabile dipendente sia, al più, ridotta in scala di tipo ordinale (*rank-ordering*): si hanno allora la **MONANOVA** (analisi di varianza monotona), il **LINMAP** o il **PREFMAP**;
- Metodi per variabile dipendente in scala ad intervallo (*rating*): in questa classe c'è la **regressione multipla di tipo OLS** (*ordinary least square*, metodo dei minimi quadrati);
- Metodi che relazionano dati da confronti a coppie con modelli di probabilità di scelta: in questa classe ci sono i modelli **LOGIT** e il **PROBIT**.

***N.B. In ambito aziendale, tuttavia, si è imposto l'approccio metrico (regressione OLS), sotto l'ipotesi semplificatrice, per la stima dei parametri del modello, che la variabile risposta sia su scala ad intervalli.***

---

## LA REGRESSIONE MULTIPLA PER LA STIMA DELLE UTILITA' MEDIANTE OLS

Il metodo più utilizzato e conosciuto per stimare le utilità parziali nella Conjoint Analysis, è il metodo di *Regressione Multipla con predittori qualitativi*.

La Funzione di Preferenza più utilizzata è di tipo additivo, quindi sarà fatto riferimento alla *Regressione Lineare Multipla*, e più precisamente dato che il giudizio di preferenza di norma è richiesto non ad un solo individuo ma ad un intero campione, si farà riferimento ad una *Regressione Lineare Multipla Multivariata*.

La funzione espressa in forma matriciale è la seguente:

$$Y = BX + E$$

dove **X** è la matrice del *Disegno Sperimentale*, ovvero la matrice ottenuta da un Disegno Fattoriale Completo o Frazionato.

La matrice **X** è **espressa in codifica disgiuntiva completa**. Sulle righe ci sono i profili-scenari di prodotto, sulle colonne i livelli di attributo.

La matrice è partizionata per colonna ed ogni partizione corrisponde ad un P attributo del Disegno Sperimentale, e contiene tante variabili dummy quanti sono i livelli dell'attributo P.

# Un esempio di matrice X

In ogni colonna la presenza di 1 e 0, indica la presenza (1) o assenza (0), di un livello specifico in uno specifico profilo.

	X <sub>l1</sub> .....X <sub>ek1</sub>	X <sub>j1</sub> .....X <sub>jxj</sub>	X <sub>pi</sub> .....X <sub>ek1</sub>
<b>Prof1</b>	1.....0	0.....1	1.....0
<b>Prof2</b>	0.....1	1.....0	0.....1
<b>Prof3</b>	1.....0	0.....1	0.....1
<b>Prof4</b>	1.....0	1.....0	1.....0
<b>Prof5</b>	0.....1	0.....1	0.....1

## Un esempio di matrice Y

Y è la matrice dei Giudizi di preferenza. In ogni colonna ci sono i giudizi espressi da ogni giudice rispetto agli scenari di prodotto.

	G1	G2	G3	G4	G5	G6	G7
prof1	1	2	7	1	2	6	5
prof2	2	4	6	3	3	7	4
prof3	3	6	5	5	5	1	3
prof4	4	7	4	7	1	2	2
prof5	5	1	3	2	4	3	1
prof6	6	3	2	4	6	4	6
prof7	7	5	1	6	7	5	7

## Stima della matrice **B** dei coefficienti di utilità

Applicando il metodo dei minimi quadrati si minimizza la somma dei quadrati degli scarti tra giudizi forniti (variabili di risposta) e giudizi teorici.

Per tutti i  $Q$  profili del disegno vale il seguente criterio:

$$\min \sum_{i=1}^q e_i^2 = \min \left( \sum_{i=1}^q y_i - \sum_{i=1}^p x_{i1j} b_{i1j} \right)^2$$

La matrice **B dei coefficienti di utilità parziali** di dimensioni **KxG (livelli x giudizi)**, sarà uguale a:

$$B = (X'X)^{-1} X'Y$$

La matrice **X non è di rango pieno**, poiché in ogni partizione i valori sono linearmente dipendenti, dando la possibilità di stimare indipendentemente solo  $K_{j-1}$  coefficienti di utilità. Se la matrice  $X$  non è di rango pieno, la matrice **X'X** sarà una matrice singolare e sarà impossibile in questo modo calcolare la sua inversa

Il sistema di equazioni normali associato alla formula precedente risulta indeterminato, essendo la matrice  $\mathbf{X}$  del piano sperimentale **singolare**. Per risolvere il sistema si deve introdurre una condizione complementare (De Luca, 2000), che può realizzarsi, ad esempio, nel sopprimere, con scelta arbitraria, una colonna in ciascun blocco di **variabili dummy** e nel porre pari a zero il corrispondente coefficiente. Nella matrice  $\mathbf{X}$  viene soppresso, perciò, un numero di colonne pari al numero di fattori. Si giunge così alla matrice del piano sperimentale  $\mathbf{X}_0$ .

$$\begin{array}{c} \mathbf{X} \\ \begin{array}{|c|c|c|c|c|} \hline 0 & 1 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 1 & 0 \\ \hline 1 & 0 & 0 & 1 & 0 \\ \hline \end{array} \end{array} \quad \longrightarrow \quad \begin{array}{c} \mathbf{X}_0 \\ \begin{array}{|c|c|c|} \hline 0 & 1 & 1 \\ \hline 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline 1 & 0 & 1 \\ \hline \end{array} \end{array}$$

In questo caso è stata eliminata l'ultima colonna della prima partizione e della seconda partizione

# Valutazione e interpretazione dei risultati

Poiché si misura una grandezza non fisica (la preferenza) utilizzando un questionario, è importante valutare la bontà dello strumento di misura stesso: questo deve essere **affidabile**, cioè **deve produrre le stesse misure** qualora cambino le condizioni esterne (es. l'intervistatore) e deve essere **valido**, cioè deve misurare proprio quanto si propone di considerare.

L'analisi può essere fatta a livello **aggregato** o **disaggregato**:

- A **livello disaggregato**, per ogni persona intervistata viene costruito un modello, venendo così a conoscenza del comportamento di ogni consumatore.
- A **livello aggregato**, si prepara un modello unico, perdendo informazioni circa i gusti del singolo.

Oltre a descrivere l'impatto di ogni livello, si stima anche l'importanza relativa di ogni fattore (in percentuale) per ogni intervistato.

# Valutazione della bontà del modello

Per verificare, la bontà della ricostruzione delle preferenze, quindi per valutare l'accuratezza delle stime si utilizzano:

<b>Giudizi metrici RATING</b>	<b>Giudizi <u>non</u> metrici RANKING</b>
Indice di determinazione lineare $R^2$	Indice di cograduazione Rho di Spearman
Coefficiente di correlazione di Bravais-Pearson	Tau di Kendall

# Rho di Spearman

Il *rho di Spearman* è un **coefficiente di cograduazione** per misurare la relazione esistente tra due variabili graduate per dimensione o per importanza crescenti.

Nel nostro caso, si considera da un lato l'ordinamento dei concetti così come rilevato nella realtà (secondo il giudizio di preferenza dei clienti), dall'altro l'ordinamento ricavabile dal modello ottenuto.

Detta **D** la differenza tra ranghi di valori corrispondenti delle due osservazioni (reale e stimata) e **N** il numero di coppie di dati, si ha:

$$r_s = 1 - \frac{(6 \times \sum D^2)}{[N \times (N^2 - 1)]}$$

# Tau di Kendall

Il *tau di Kendall* è un metodo per esprimere la corrispondenza tra le preferenze espresse dagli intervistati e i pesi stimati.

Consideriamo un esempio: siano A, B, C, D quattro profili valutati con giudizi "rank order" di cui si siano stimati i pesi (*utilities*)

Profili	Rank	Utility
A	3	2.3
B	4	4.5
C	7	1.7
D	9	-6.1

I pesi stimati dovrebbero rispecchiare l'ordine di preferenza (pesi maggiori ai profili preferiti) e così dovrebbe essere:  
A preferito a B A preferito a C  
A preferito a D B preferito a C  
B preferito a D C preferito a D

In realtà, le differenze tra le "*utilities*" stimate sono:

$$2.3 - 4.5 = -2.2; 2.3 - 1.7 = 0.6; 2.3 + 6.1 = 8.4$$

$$4.5 - 1.7 = 2.8; 4.5 + 6.1 = 10.6; 1.7 + 6.1 = 7.8$$

Si vede così che se una coppia non soddisfa quanto espresso dal cliente compare il segno negativo, quindi tau sarà pari a:

$$t = \frac{n^{\circ} \text{ confronti giusti} - n^{\circ} \text{ confronti sbagliati}}{n^{\circ} \text{ confronti}} = \frac{5 - 1}{6} = 0,666$$

# Le importanze relative percentuali degli attributi

Nell'ambito della metodologia di Conjoint Analysis è consuetudine valutare l'importanza relativa dei fattori.

Questo viene fatto con lo scopo di rendere comparabili le utilità parziali e di pervenire a valori di importanza dei fattori compresi tra 0 e 1 (o ovviamente in percentuale).

$$I_j = \frac{\max_j [u(I_j)] - \min_j [u(I_j)]}{\sum_{j=1}^p \left( \max_j [u(I_j)] - \min_j [u(I_j)] \right)}$$

La percentuale di importanza per il j-esimo attributo si ottiene rapportando la differenza tra il valore massimo e quello minimo di utilità relativo all'attributo stesso, alla somma per tutti gli attributi di tale differenza.

# Applicazione dei risultati ottenuti

I modelli ottenuti (uno per persona) possono servire per:

- **decisioni di segmentazione**: si raggruppano i consumatori con giudizi simili;
- **analisi di redditività relativa** di un certo progetto di prodotto: conoscendo il costo di ogni caratteristica, si può combinare il costo di ogni "prodotto" (combinazione) con la quota di mercato prevista per osservarne la fattibilità;
- **simulazioni**: si può prevedere la quota di mercato che può raggiungere un certo "stimolo" in diversi scenari competitivi; si può stimare cosa succeda se si aggiunge un nuovo prodotto nel settore.

# I rischi della Conjoint Analysis

- ❑ Alcuni prodotti/servizi, soprattutto quelli a **forte contenuto d'immagine**, possono non essere valutati analiticamente dai consumatori nelle loro caratteristiche rilevanti;
- ❑ I risultati ottenuti sono fortemente condizionati dal ricorso ad alcune ipotesi volutamente semplificatrici, in particolare:
  - Lo schema additivo, che collega la preferenza per ogni alternativa da valutare, ai livelli o modalità degli attributi che la caratterizzano;
  - L'assenza di interazioni di qualunque ordine tra tali modalità o livelli.



# Conjoint Analysis in R

---

Installare pacchetto “Conjoint”

Caricare pacchetto “Conjoint” <- `library(conjoint)`

## **Disegno fattoriale:**

Funzione “`expand.grid`” <- creazione di tutti i possibili scenari, a partire da un set di attributi con i relativi livelli.

Esempio:

```
scenari = expand.grid( pagamento = c("tessera_assoc", "spot", "tessera_punti"),  
letterario = c("presentazioni_eventi", "gruppi_discussione"),  
ambientalista = c("usato", "riciclato"),  
passatempo = c("riviste", "giochi_tavolo", "tablet"))
```

Funzione “`caFactorialDesign`” <- creazione del disegno fattoriale completo o frazionato

```
caFactorialDesign(data, type="null", cards=NA)  
Data<- esperimento il cui disegno consiste in 2 o più fattori (attributi), ognugno dei quali con 2 o più livelli  
Type<- tipo del disegno fattoriale (possibilil valori : "full", "fractional", "orthogonal".....)  
Cards<- numero dei scenari (usato solo con disegno fattoriale frazionato )
```

Esempio:

```
design<-caFactorialDesign(data=scenari, type="orthogonal")  
print(design)  
code<-caEncodedDesign(design)  
print(code)
```

Scrivere l’output in Excel/csv

```
write.csv2(design, file="orthogonal_factorial_design.csv", row.names=FALSE)
```

---

Funzione “Conjoint”<- combinazione delle seguenti funzioni caPartutilities, caUtilities and caImportance. Come risultato otteniamo una matrice delle utilità parziali per livelli degli attributi per intervistati, un vettore delle utilità dei livelli degli attributi ed un vettore della importanza degli attributi (in %) con i corrispondenti grafici a barre.

*Conjoint(y, x, z)*

y - matrice dei giudizi

x - matrice degli scenari

z - vettore nomi dei livelli

Funzione “caPartUtilities” <- calcola matrice delle utilità individuali per intervistato

*caPartUtilities(y, x, z)*

Funzione “caUtilities”<- calcola utilità dei livelli degli attributi

*caUtilities(y, x, z)*

Funzione “caTotalUtilities” <- calcola matrice delle utilità totali di n profili (scenari) per ogni intervistato

*caTotalUtilities(y, x)*

Funzione “caImportance” <- calcola l'importanza di tutti gli attributi. La somma dell'importanze relative deve essere uguale a 100.

*calImportance(y, x)*

Funzione “caModel” <- calcola le stime dei parametri della conjoint analysis per ogni intervistato.

```
caModel(y, x)
```

*y* - vettore dei giudizi (vector should be like single profile of preferences)

*x* - matrice dei profili

Per calcolare il tau di Kendall usare la funzione “cor.test”

```
cor.test(x, y, method = "kendall")
```

*x, y* <- vettori numerici della stessa lunghezza

## Esempio pratico:

```
library(conjoint)
data(herbata)
```

*#hprof* - Matrix of profiles (4 attributes and 13 profiles).

```
print(hprof)
```

*#hlevn* - Character vector of names for the attributes' levels.

```
print(hlevn)
```

*#hprefm* - Matrix of preferences (100 respondents and 13 profiles).

```
print(hprefm)
```

*#hsimp* - Matrix of simulation profiles.

```
print(hsimp)
```

*#hpref* - Vector of preferences (length 1300).

```
print(hpref)
```

```
Conjoint(hprefm,hprof,hlevn)
```

```
uslall=caPartUtilities(hpref,hprof,hlevn)
```

```
print(uslall)
```

```
ul=caUtilities(hpref,hprof,hlevn)
```

```
print(ul)
```

```
imp=caImportance(hpref,hprof)
```

```
print("Importance summary: ", quote=FALSE)
```

```
print(imp)
```

```
print(paste("Sum: ", sum(imp)), quote=FALSE)
```

```
Usi=caTotalUtilities(hpref,hprof)
```

```
Print(Usi)
```

```
x<-as.data.frame(hprof)
```

```
y1<-as.data.frame(hpref[1:nrow(x),1])
```

```
model=caModel(y1, x)
```

```
print(model)
```

```
x<-as.numeric(t(hpref))
```

```
y<-matrix(t(Usi),nrow=1300,ncol=1)
```

```
cor.test(x, y, method = "kendall")
```

## Esempio pratico:

```
library(conjoint)
```

```
data(herbata) <- i dati riguardano il giudizio dei consumatori sulle diverse caratteristiche del tè
```

```
#hprof – Matrice dei profili (4 attributi e 13 profili)
```

```
print(hprof)
```

	prezzo	specie	tipo	aroma
	cena	gatunek	rodzaj	aromat
1	3	1	1	1
2	1	2	1	1
3	2	2	2	1
4	2	1	3	1
5	3	3	3	1
6	2	1	1	2
7	3	2	1	2
8	2	3	1	2
9	3	1	2	2
10	1	3	2	2
11	1	1	3	2
12	2	2	3	2
13	3	2	3	2

```
#hlevn – Vettore dei nomi dei livelli degli attributi
```

```
print(hlevn)
```

levels			
1	niska	← basso	P R E Z Z O
2	srednia	← medio	
3	wysoka	← alto	
4	czarna	← nero	S P E C I E
5	zielona	← verde	
6	czerwona	← rosso	
7	ekspresowa	← in bustina	T I P O
8	granulowana	← granulato	
9	lisciasta	← in foglie	
10	tak	← si	A R O M A
11	nie	← no	

```
#hprefm – Matrice delle preferenze (100 rispondenti e 13 profili)
print(hprefm)
```

	profil1	profil2	profil3	profil4	profil5	profil6	profil7	profil8	profil9	profil10	profil11	profil12	profil13
1	8	1	1	3	9	2	7	2	2	2	2	3	4
2	0	10	3	5	1	4	8	6	2	9	7	5	2
3	4	10	3	5	4	1	2	0	0	1	8	9	7
4	6	7	4	9	6	3	7	4	8	5	2	10	9
5	5	1	7	8	6	10	7	10	6	6	6	10	7
6	10	1	1	5	1	0	0	0	0	0	0	1	1
7	8	0	0	0	9	0	0	0	0	0	5	10	8
8	5	2	1	4	3	8	5	9	6	8	3	1	2
9	7	3	3	9	0	5	3	0	5	0	5	10	8
10	8	7	3	10	9	1	2	2	2	2	8	10	8
11	7	6	4	10	8	2	2	3	2	3	9	10	9
12	8	7	3	10	9	2	3	2	2	2	9	10	7
13	0	0	0	2	3	4	0	1	0	1	5	3	0
14	7	0	0	4	4	1	0	1	2	1	3	0	0
15	9	0	0	2	0	8	0	0	6	0	3	0	0
16	9	1	1	2	1	8	1	0	6	0	4	1	1
17	2	6	4	4	2	4	2	8	2	8	4	4	4
18	10	1	1	0	0	0	0	0	0	0	0	1	1
19	5	4	8	1	2	2	4	3	3	1	2	8	6
20	10	0	0	5	0	2	0	0	2	0	5	0	0
21	10	8	0	0	2	0	0	0	0	0	0	6	3
22	10	10	2	6	6	2	2	2	2	2	6	6	6

```
#hpref – Vettore delle preferenze (lunghezza 1300=100*13).
print(hpref)
```

```
#hsimp – Matrice di simulazione dei profili
print(hsimp)
```

## Conjoint(hprefm,hprof,hlevn)

```
> Conjoint(hprefm,hprof,hlevn)
```

```
[1] "Part worths (utilities) of levels (model parameters for whole sample):"
```

	levnms	utls
1	intercept	3,5534
2	niska	0,2402
3	srednia	-0,1431
4	wysoka	-0,0971
5	czarna	0,6149
6	zielona	0,0349
7	czerwona	-0,6498
8	ekspresowa	0,1369
9	granulowana	-0,8898
10	lisciasta	0,7529
11	tak	0,4108
12	nie	-0,4108

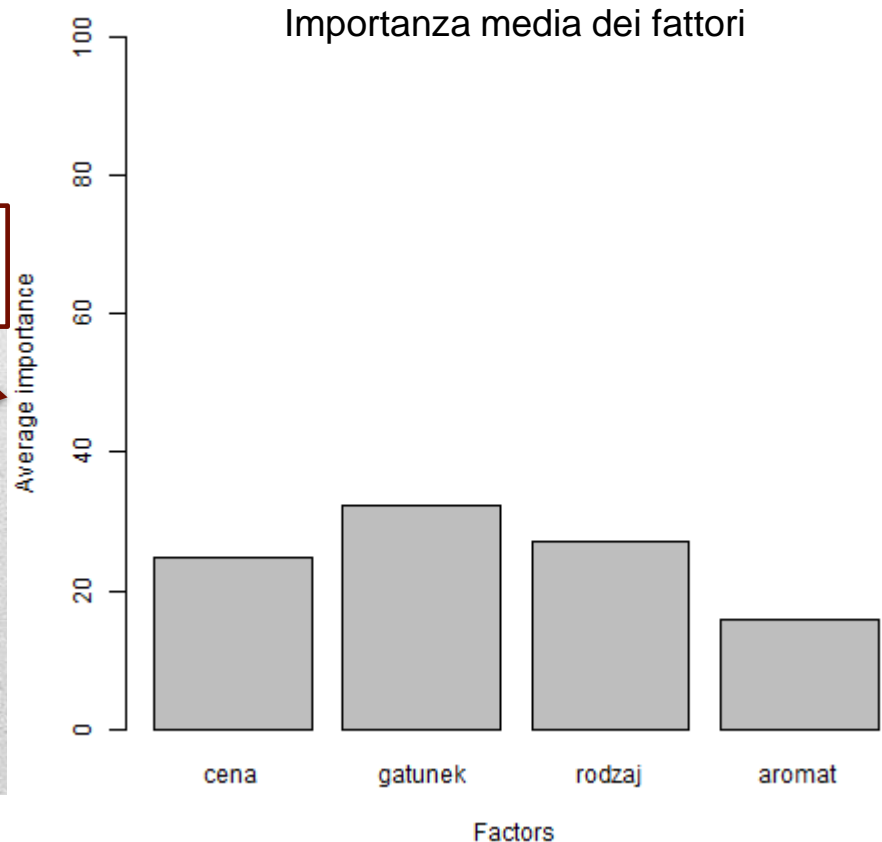
Utilità parziali dei livelli (campione intero)

```
[1] "Average importance of factors (attributes):"
```

```
[1] 24,76 32,22 27,15 15,88
```

```
[1] Sum of average importance: 100,01
```

```
[1] "Chart of average factors importance"
```



```
ul=caUtilities(hpref,hprof,hlevn)  
print(ul)
```

```
> ul=caUtilities(hpref,hprof,hlevn)
```

```
> print(ul)
```

```
[1] 3,55336207 0,24022989 -0,14311494 -0,09711494 0,61488506 0,03488506  
[7] -0,64977011 0,13688506 -0,88977011 0,75288506 0,41077586 -0,41077586
```

```
uslall=caPartUtilities(hpref,hprof,hlevn)
print(uslall)
```

Utilità parziali dei livelli per ogni intervistato

```
> uslall=caPartUtilities(hpref,hprof,hlevn)
> print(uslall)
      intercept  niska srednia wysoka czarna zielona czerwona ekspresowa
[1,]    3,394 -1,517 -1,141  2,659 -0,475 -0,675  1,149  0,659
[2,]    5,049  3,391 -0,695 -2,695 -1,029  0,971  0,057  1,105
[3,]    4,029  2,563 -1,182 -1,382 -0,248  2,352 -2,103 -0,382
[4,]    5,856 -1,149 -0,025  1,175 -0,492  1,308 -0,816 -0,825
[5,]    6,250 -2,333  2,567 -0,233 -0,033 -0,633  0,667 -0,233
[6,]    1,578 -0,713 -0,144  0,856  1,456 -0,744 -0,713  0,656
[7,]    2,635 -0,920 -1,040  1,960 -0,707  0,293  0,414 -1,107
[8,]    4,405 -0,425  0,413  0,013  0,546 -2,454  1,908  1,479
[9,]    3,546 -0,966  0,883  0,083  2,216  1,416 -3,632 -0,917
[10,]   5,460  0,678 -0,639 -0,039  0,228  0,428 -0,655 -1,172
[11,]   5,626  0,678 -0,239 -0,439  0,228  0,428 -0,655 -1,439
[12,]   5,566  0,862 -0,631 -0,231  0,502  0,302 -0,805 -0,898
[13,]   1,319  0,552  0,424 -0,976  0,691 -0,909  0,218 -0,243
[14,]   1,925 -0,264 -0,568  0,832  1,499 -1,901  0,402  0,099
[15,]   1,776 -0,793 -0,103  0,897  3,697 -1,903 -1,793  1,097
[16,]   2,296 -0,632 -0,284  0,916  3,383 -1,417 -1,966  0,983
[17,]   4,506  1,379  0,510 -1,890 -1,090 -0,290  1,379  0,377
[18,]   1,075 -0,402 -0,699  1,101  0,968 -0,232 -0,736  1,101
[19,]   3,486 -1,115  0,757  0,357 -0,976  2,424 -1,448 -0,376
[20,]   1,736  0,218 -0,609  0,391  3,124 -1,676 -1,448  0,591
```

```
x<-as.numeric(t(hpref))
y<-matrix(t(Usi),nrow=1300,ncol=1)
cor.test(x, y,method = "kendall")
```

```
Usi=caTotalUtilities(hpref,hprof)
Print(Usi)
```



Utilità totali dei profili per ogni intervistato

```
x<-as.data.frame(hprof)
y1<-as.data.frame(hpref[1:nrow(x),1])
model=caModel(y1, x)
print(model)
```

```
> x<-as.data.frame(hprof)
> y1<-as.data.frame(hpref[1:nrow(x),1])
> model=caModel(y1, x)
> print(model)
```

```
Call:
lm(formula = frml)
```

```
Residuals:
```

```
    1     2     3     4     5     6     7     8     9    10
1,1345 -1,4897  0,3103 -0,2655  0,3103  0,1931  1,5931 -1,4310 -1,4310  1,1207
    11    12    13
0,3690  1,1931 -1,6069
```

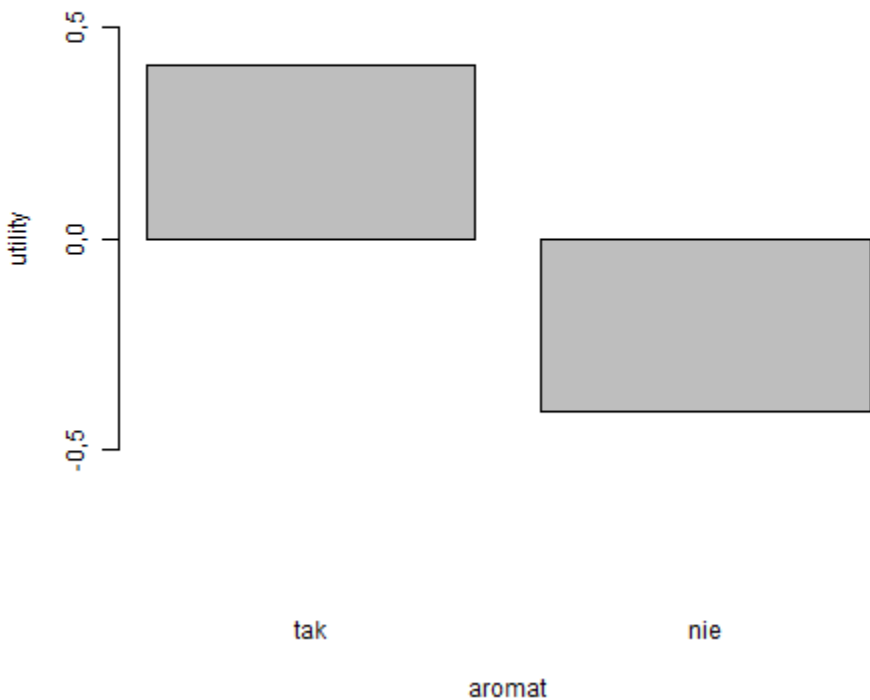
```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	3,3937	0,5439	6,240	0,00155 **
factor(x\$scena) 1	-1,5172	0,7944	-1,910	0,11440
factor(x\$scena) 2	-1,1414	0,6889	-1,657	0,15844
factor(x\$gatunek) 1	-0,4747	0,6889	-0,689	0,52141
factor(x\$gatunek) 2	-0,6747	0,6889	-0,979	0,37234
factor(x\$rodzaj) 1	0,6586	0,6889	0,956	0,38293
factor(x\$rodzaj) 2	-1,5172	0,7944	-1,910	0,11440
factor(x\$aromat) 1	0,6293	0,5093	1,236	0,27150

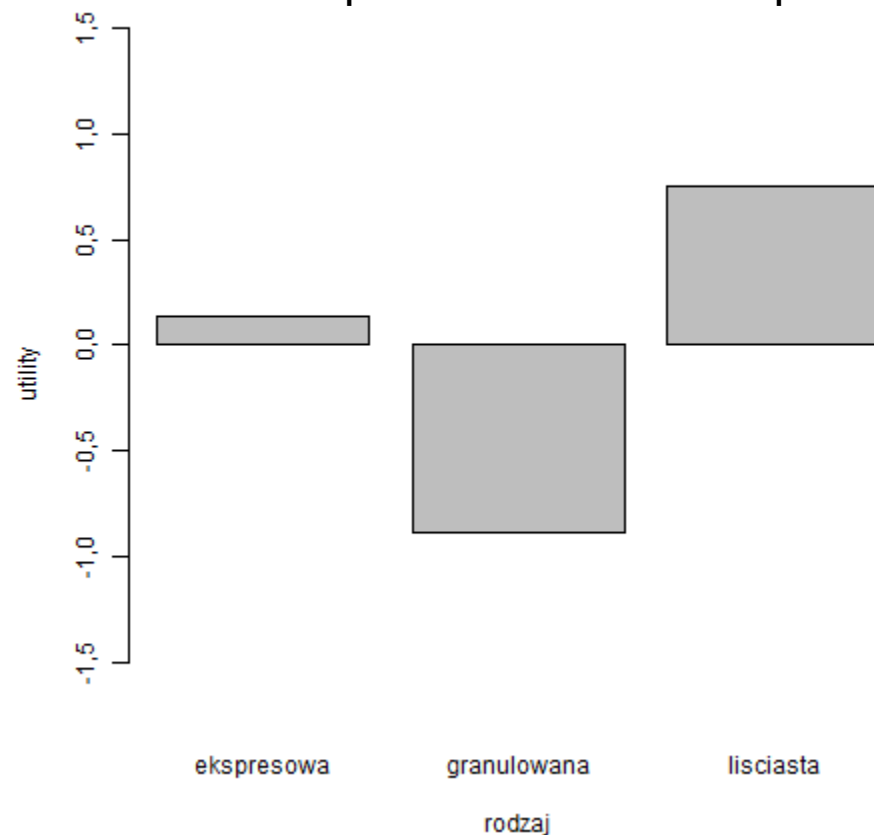
```
---
Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1
```

```
Residual standard error: 1,78 on 5 degrees of freedom
Multiple R-squared:  0.8184,    Adjusted R-squared:  0.5642
F-statistic:  3.22 on 7 and 5 DF,  p-value: 0,1082
```

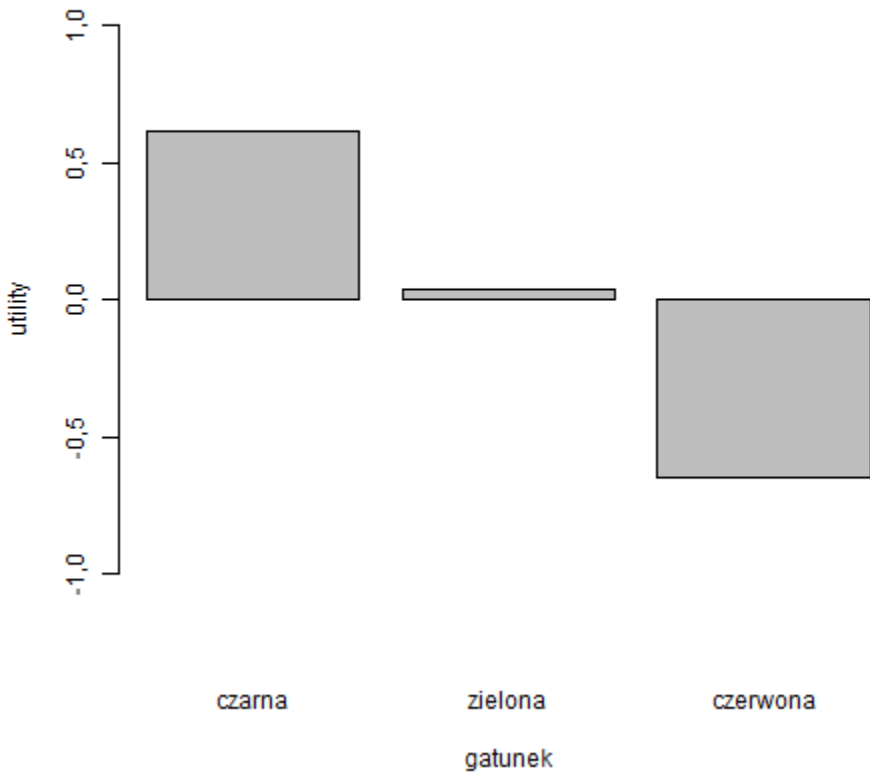
## Utilità parziali del fattore Aroma



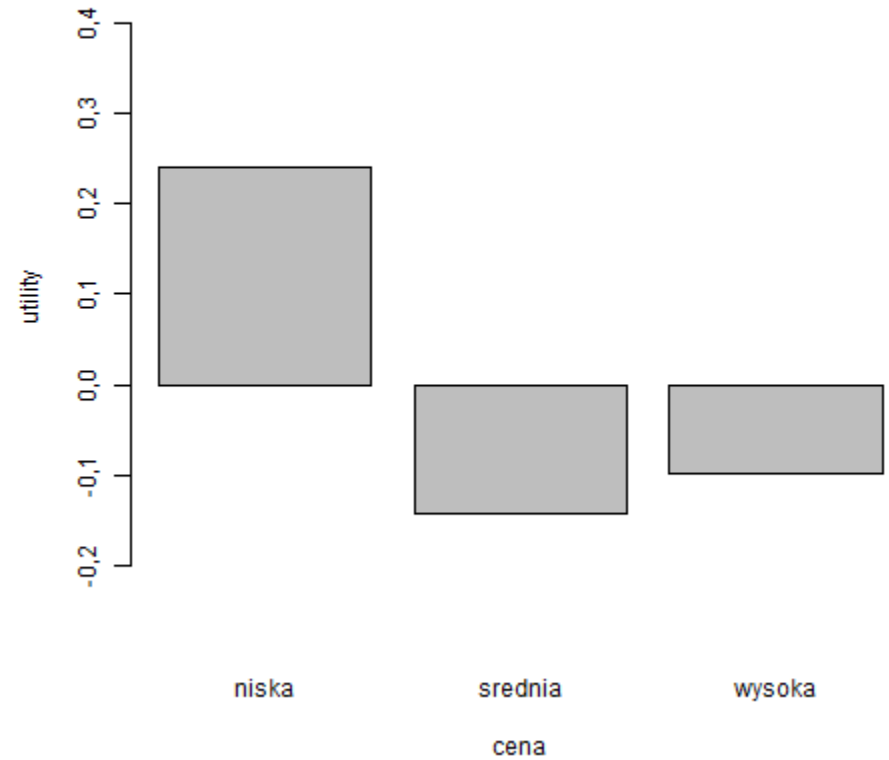
## Utilità parziali del fattore Tipo



## Utilità parziali del fattore Specie



## Utilità parziali del fattore Prezzo



```
> x<-as.numeric(t(hpref))
> y<-matrix(t(Usi),nrow=1300,ncol=1)
> cor.test(x, y,method = "kendall")
```

Kendall's rank correlation tau

```
data: x and y
z = 39,8145, p-value < 2,2e-16
alternative hypothesis: true tau is not equal to 0
sample estimates:
tau
0,7763287
```

← Calcolo del valore di Tau di Kendall