



# Heterogeneity in Household Consumptions: The case of the Marche Region

Chiara Gigliarano

Università Politecnica delle Marche, Ancona

con F. Chelli and E. Mattioli



# INDICE

- Problematica e obiettivo del lavoro
- Dati
- Metodologia
- Risultati dall'analisi empirica
- Osservazioni conclusive



# PROBLEMATICA

- Le famiglie differiscono in termini di comportamento di consumo
- Per spiegare tale eterogeneità, le famiglie vengono solitamente aggregate in gruppi in base a caratteristiche socio-demografiche.

## Esempi:

- ISTAT (2007a): num. figli, num. componenti, età PR
  - ISTAT (2007b): famiglie in affitto, di pensionati, con basso reddito
  - Bono, Cuffaro e Giaimo (2004): regione, classe di reddito
- Partizioni arbitrarie e a priori!



# OBIETTIVO DEL LAVORO

Ridurre l'ARBITRARIETÀ nella scelta delle partizioni entro cui classificare le famiglie secondo i comportamenti di spesa:

- Ricercare la PARTIZIONE OTTIMA in cui i gruppi siano internamente il più OMOGENEI possibile in termini di abitudini di spesa per consumo
- Identificare le CARATTERISTICHE FAMILIARI che meglio SPIEGANO LA VARIABILITÀ nei comportamenti di consumo delle famiglie



# DATI

- Indagine sui Consumi delle Famiglie, condotta dall'ISTAT per l'anno 2005
- Unità di analisi: 1,026 famiglie private rappresentative delle 585,821 famiglie residenti nelle Marche
- Variabili d'interesse: 279 variabili specifiche di consumo e caratteristiche socio-demografiche
- Scala d'equivalenza Carbonaro per tener conto di economie di scala all'interno della famiglia

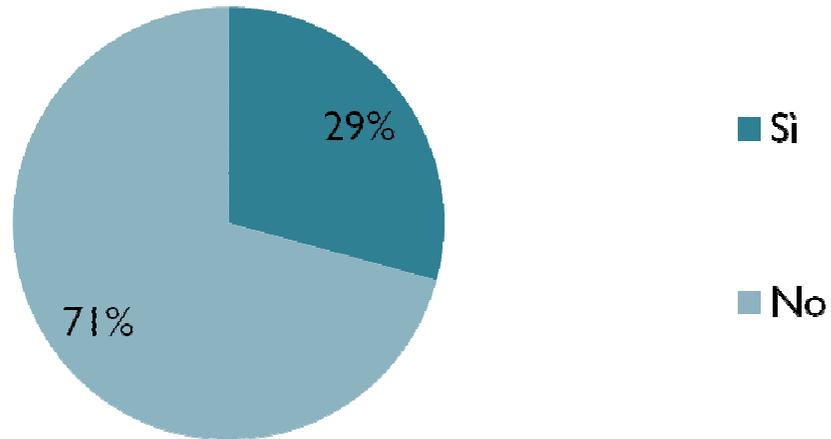


# METODOLOGIA

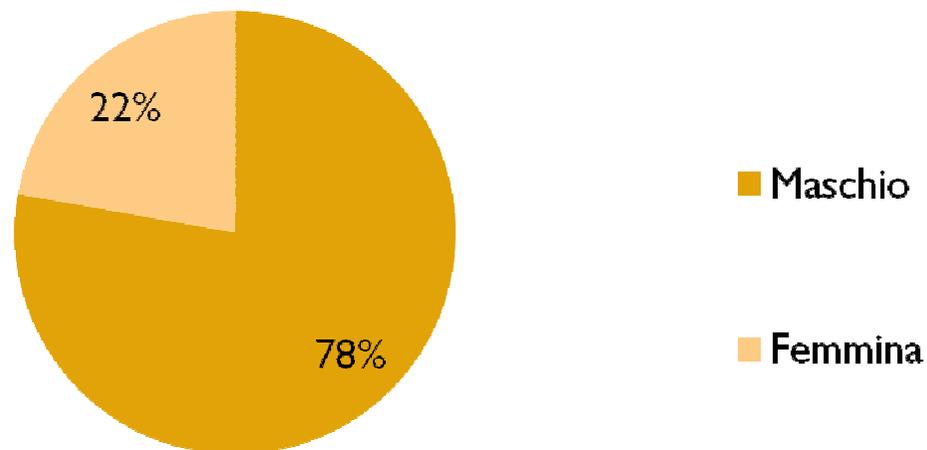
- **Analisi discriminante:**
  - variabili quantitative : 279 voci di spesa
  - variabile di classificazione: ognuna delle 9 partizioni considerate
- **Utilizzare l'informazione di ciascuna voce di spesa specifica, anziché usare la variabile del consumo totale**

# Partizioni considerate

## Presenza di minori

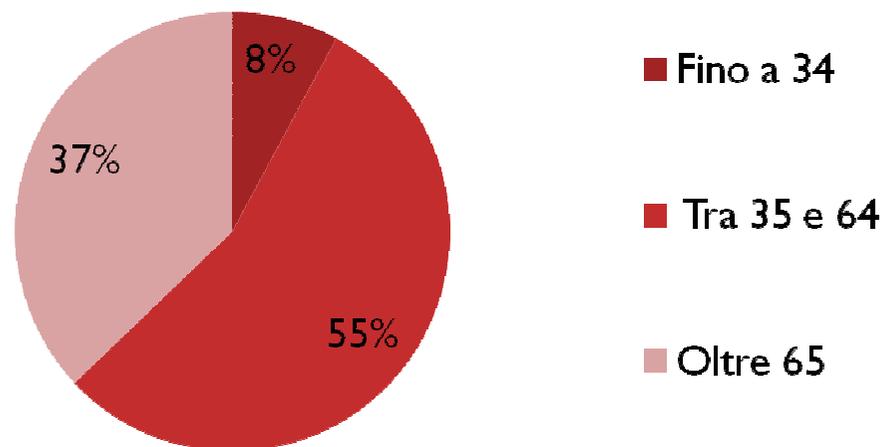


## Genere della PR

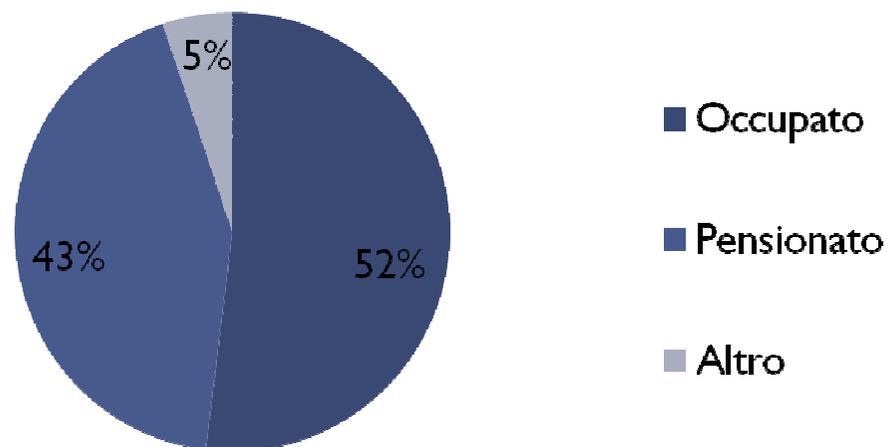


# Partizioni considerate (2)

## Età in anni della PR

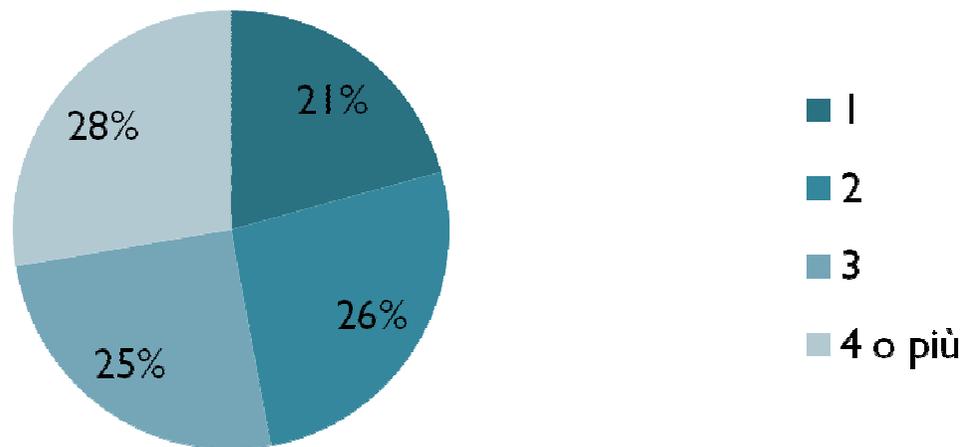


## Professione PR

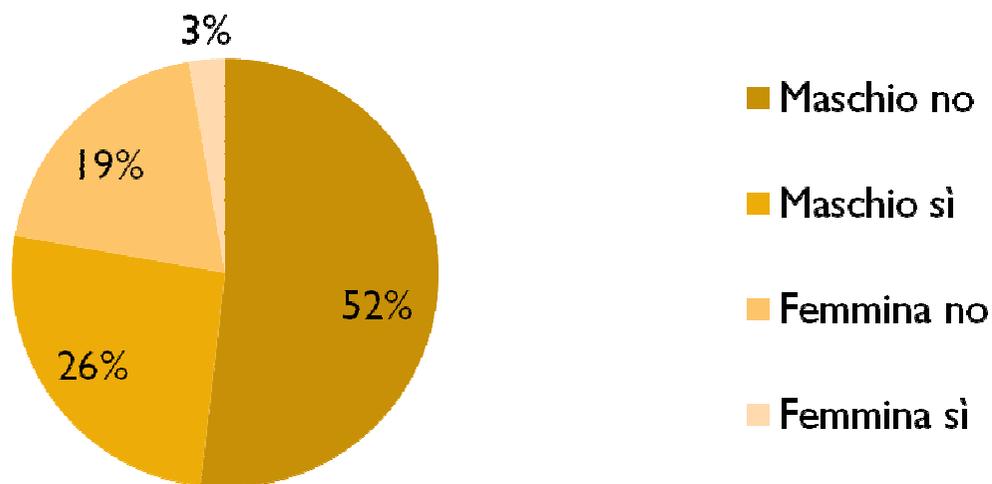


# Partizioni considerate (3)

## Numero componenti

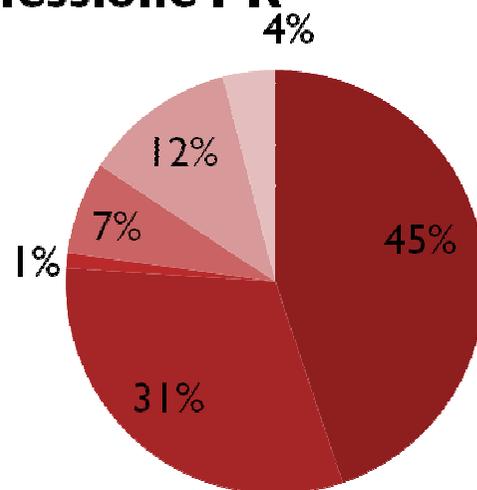


## Genere PR e presenza minori



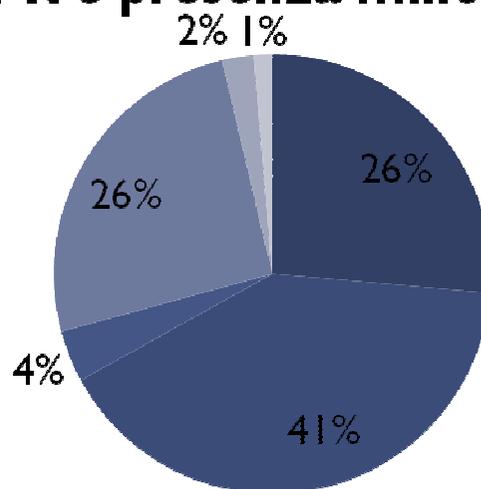
# Partizioni considerate (4)

## Genere e professione PR



- Maschio occupato
- Maschio pensionato
- Maschio altro
- Femmina occupata
- Femmina pensionata
- Femmina altro

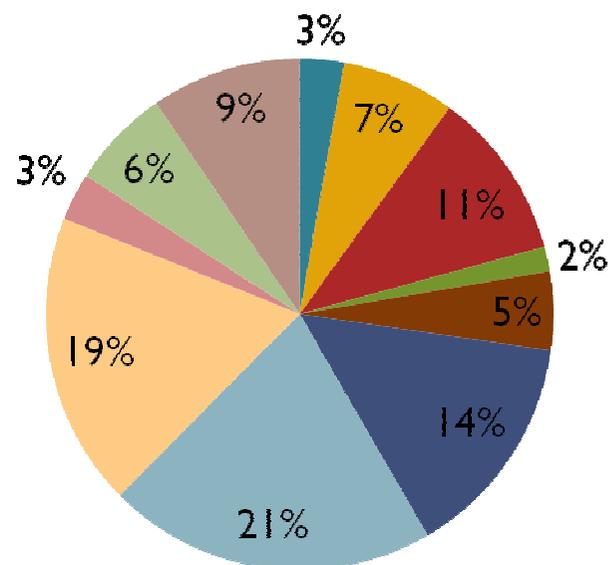
## Professione PR e presenza minori



- Occupato no
- Pensionato no
- Altro no
- Occupato sì
- Pensionato sì
- Altro sì

# Partizioni considerate (5)

## Partizione ISTAT



■ Single <35 anni

■ Single [35;64] anni

■ Single >=65 anni

■ Coppia senza figli, PR <35 anni

■ Coppia senza figli, PR [35;64] anni

■ Coppia senza figli, PR >=65 anni

■ Coppia con 1 figlio

■ Coppia con 2 figli

■ Coppia con 3 figli

■ Famiglie monogenitori

■ Altri tipi famiglie

# METODOLOGIA (2)

- Ogni famiglia è classificata nel gruppo a cui corrisponde la più alta probabilità a posteriori  $p(x/h) \approx \pi_h f_h(x)$
- Probabilità a priori  $\pi_h$  data dalle proporzioni campionarie di ogni gruppo
- Stima non parametrica della densità secondo il metodo del *k-esimo nearest neighbor* (Fix e Hodges, 1951). Per  $k$  fissato,

$$f_h(x) = \frac{1}{n_h v_h} \sum 1(y \in h : d^2(x, y) \leq r_k^2(x))$$

# METODOLOGIA (3)

- Scelta di  $k$ :  $k$  con minimo tasso d'errore;  $k \approx n^{2/8}$ ;  $k \approx n^{3/8}$  (Enas e Choi, 1986)

- Valutazione di ogni partizione secondo la probabilità di errata classificazione:

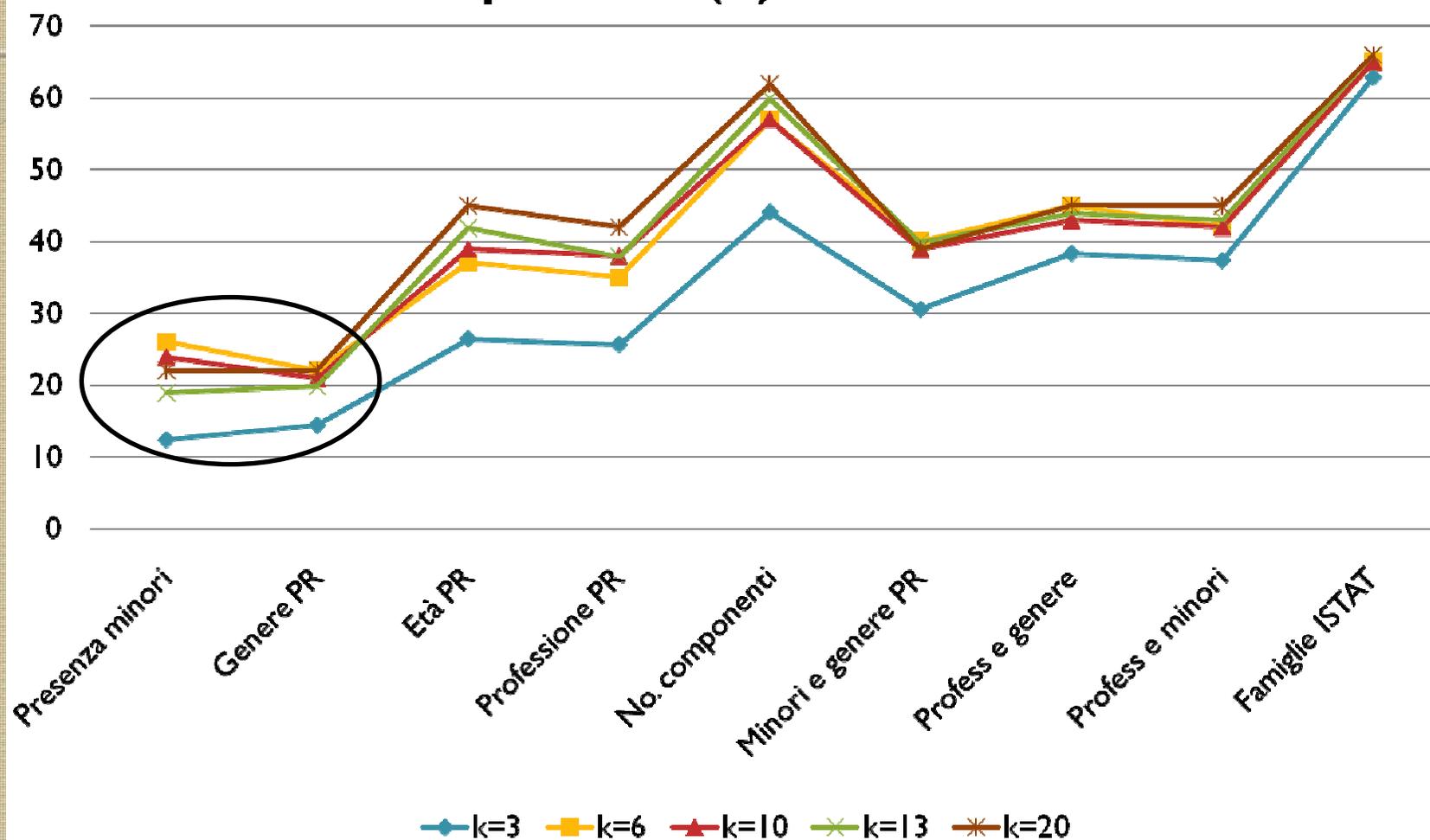
$$e = \sum_h \pi_h \left( 1 - \frac{1}{n\pi_h} \sum_{R_h} p(h|x) \right)$$

dove  $R_h$  è l'insieme delle  $x$  per cui  $p(x/h)$  più alta

- Partizione migliore: quella con più bassa probabilità di errata classificazione

# Risultati da analisi empirica

## Probabilità a posteriori (%) di errata classificazione



Indici di concordanza di Kendall e di Spearman compresi tra 0.91 e 0.99



# Risultati da analisi empirica: descrizione delle partizioni ottimali

- Analisi discriminante permette di individuare le voci di consumo con media **SIGNIFICATIVAMENTE DIVERSA TRA I GRUPPI** individuati
- Il numero di variabili specifiche di consumo si riduce notevolmente: da 279
  - a 62, nel caso di presenza di minori
  - a 47, nel caso di genere della PR

## Variabili significative di consumo mensile (€) tra famiglie senza (A) o con minori (B)

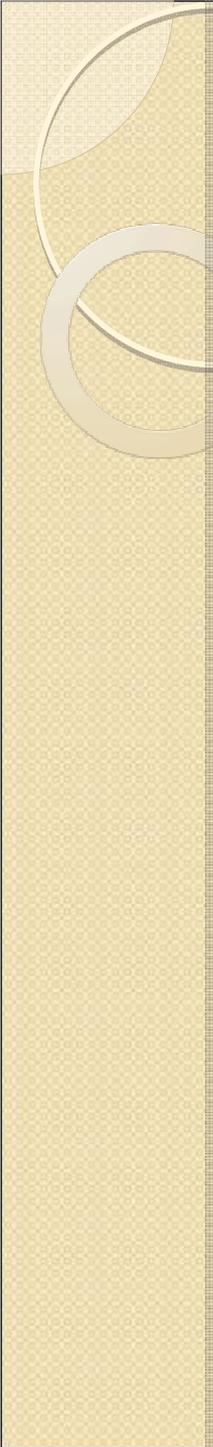
Variabile di consumo	A	B	(A)-(B)	Variabile di consumo	A	B	(A)-(B)
	Media	Media			Media	Media	
Mutuo per acquisto abitazioni	35.59	75.88	-40.29	Legumi secchi o conservati	1.50	0.84	0.65
Abbigliamento bambini, neonati	4.85	44.86	-40.00	Conserve	2.24	1.41	0.83
Pensione completa, in Italia	8.04	26.89	-18.84	Farina	2.31	1.33	0.99
Mobili singoli	4.95	23.60	-18.65	Zucchero	3.52	2.31	1.21
Assicurazione vita	17.62	28.00	-10.38	Pere	3.86	2.58	1.28
Calzature da bimbo, neonato	0.52	9.81	-9.29	Riso	3.51	1.88	1.63
Affitto abitazioni (secondarie)	0.52	9.25	-8.73	Uova	5.03	3.38	1.65
Giocattoli, giochi e videogiochi	4.33	12.64	-8.32	Bottoni, fili, passamaneria	3.08	1.15	1.93
Asili nido, baby sitter, etc.	0.00	5.56	-5.56	Pomodori freschi	8.80	5.75	3.05
Libri scolastici	0.69	6.18	-5.48	Spese per animali domestici	9.14	5.35	3.79
Sport: piscine, palestre	2.51	7.59	-5.08	Maiale	14.49	10.69	3.79
Fazzoletti carta, pannolini neonati	9.32	14.22	-4.90	Agrumi	12.09	8.11	3.98
Pasti fuori casa	19.31	23.91	-4.60	Detersivi, cere, insetticidi	18.47	14.41	4.06
Cereali e prodotti preparati	5.06	8.28	-3.22	Formaggi	25.37	21.22	4.15
Mense scolastiche	0.14	3.21	-3.07	Abbonamento TV, internet	9.65	5.45	4.20
acquisto telefono cellulare	0.94	3.98	-3.05	Manzo	12.69	8.09	4.60
Cerimonie religiose	1.27	4.17	-2.90	Mele	11.13	6.42	4.71
Rette, corsi di lingue e computer	1.00	3.60	-2.60	Caffè	10.80	6.07	4.73
Gelati	5.73	8.05	-2.32	Acqua	15.57	10.81	4.76
Lezioni private e ripetizioni	0.16	2.37	-2.21	Vino	14.20	9.12	5.08
Servizi ausiliari sanitari	0.84	2.91	-2.07	Pollame	15.08	9.96	5.13
Pasticceria e dolciumi	9.67	11.59	-1.92	Ortaggi freschi o surgelati	19.14	13.57	5.57
Succhi di frutta, etc.	3.74	5.61	-1.87	Energia elettrica	34.93	29.29	5.63
Pecora e agnello	2.21	4.03	-1.83	Carbone e legna	7.48	1.74	5.74
Quaderni, cancelleria, fotocopie	1.57	3.35	-1.79	Pane grissini e cracker	27.79	21.75	6.04
Yogurt	6.13	7.61	-1.47	Telefono	40.81	34.15	6.66
Scuolabus	0.05	1.50	-1.46	Olio di oliva	16.90	9.61	7.29
Mobili per esterno	0.00	1.05	-1.05	Pesce fresco o surgelato	40.82	30.07	10.75
Altri prodotti derivati dal latte	1.87	2.45	-0.59	Medicinali	44.10	29.56	14.54
Carbone e legna (case secondarie)	0.00	0.20	-0.20	Gas da rete	66.02	49.58	16.43
Olio di semi e di sansa	1.43	0.79	0.64	Affitto figurativo mensile	471.89	316.27	155.62

Fonte: Propria elaborazione dai dati ISTAT 2005

## Variabili sig. di consumo mensile (€) tra famiglie con PR maschio (C) e femmina (D)

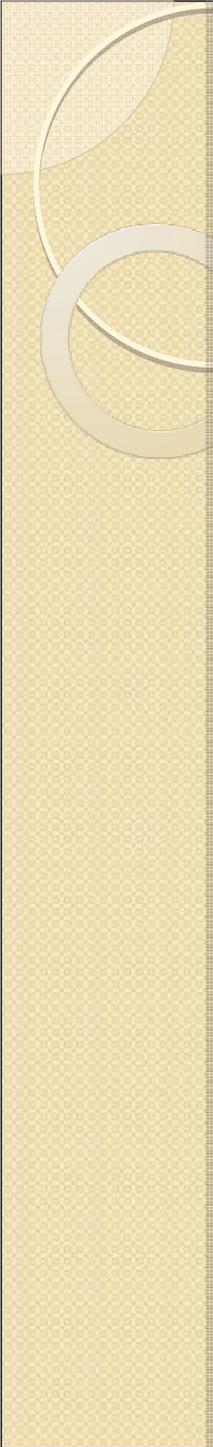
Variabile di consumo	C	D	(C)-(D)	Variabile di consumo	C	D	(C)-(D)
	Media	Media			Media	Media	
Affitto figurativo mensile	395.47	526.56	-131.09				
Canone mensile d'affitto	25.38	54.66	-29.28	Zucchero	2.67	4.77	-2.10
Gas da rete	55.73	78.92	-23.19	Yogurt	6.11	8.03	-1.93
Parrucchiere, ist. bellezza	28.56	47.05	-18.49	Tè, orzo, cacao	2.39	4.25	-1.86
Abbigliamento per donna	33.61	49.21	-15.59	Quaderni, cancelleria	1.76	3.14	-1.39
Vitello e vitellone	29.97	41.62	-11.65	Riso	2.71	4.08	-1.37
Agrumi	9.68	14.93	-5.25	Burro	1.72	2.99	-1.27
Pollame	12.36	17.54	-5.18	Farina	1.76	2.87	-1.11
Riscaldamento	2.13	7.04	-4.91	Sale, spezie, salse	2.07	3.13	-1.06
Telefono	37.75	42.46	-4.71	Margarina e altri grassi	0.19	0.43	-0.24
Abbonamento TV, Internet	7.32	11.99	-4.67	Pesce conservato, secco	2.54	0.95	1.59
Energia elettrica	32.19	36.80	-4.61	Olio e altri lubrificanti	3.20	1.48	1.71
Assistenza disabili e anziani	2.83	6.92	-4.09	Acquisto telefono cellulare	2.25	0.45	1.80
Olio di oliva	13.82	17.85	-4.03	Cereali e prodotti preparati	6.47	4.47	2.00
Acqua	13.26	17.18	-3.92	Giornali, riviste fumetti	11.50	8.32	3.18
Detersivi, cere, insetticidi	16.42	20.10	-3.68	Vino	13.62	9.80	3.82
Bottoni, fili, passamaneria	1.70	5.13	-3.43	Salumi	26.00	21.82	4.17
Biscotti, fette biscottate, etc	8.33	11.69	-3.36	Onorari commercialisti	6.29	1.76	4.53
Occhiali da vista, etc.	3.42	6.45	-3.03	Assicurazione vita	22.79	13.73	9.06
Mele	9.17	11.67	-2.50	Calzature da uomo	13.79	4.48	9.31
Caffè	8.86	11.22	-2.36	Pezzi di ricambi e accessori	19.77	5.10	14.66
Piatti, bicchieri di carta	4.42	6.63	-2.21	Combustibili per auto e moto	21.61	4.09	17.52
Patate o patatine fritte	3.92	6.12	-2.20	Assicurazione veicoli	66.06	44.21	21.85
Pomodori freschi	7.41	9.52	-2.12	Abbigliamento per uomo	43.94	18.70	25.24

Fonte: Propria elaborazione da dati ISTAT 2005



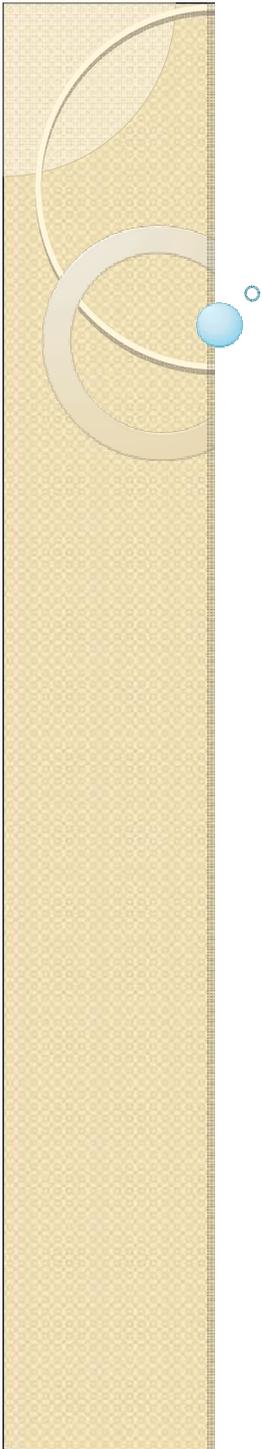
# Osservazioni conclusive

- Proposta di un metodo non arbitrario per identificare le caratteristiche che meglio spiegano la variabilità nei comportamenti di spesa delle famiglie
- Possibili applicazioni di tale metodo:
  - Indici di prezzi al consumo per sottogruppi ottimali di famiglie
  - Analisi della disuguaglianza e della polarizzazione nei consumi per sottogruppi ottimali di famiglie



# Alcuni riferimenti bibliografici

- Bono, F., Cuffaro, M. and Giaimo, R. (2004). Consumption behaviour across Regions: a multilevel approach. *Proceedings of the XLII Scientific Meeting SIS, Cleup*, 117–120.
- Enas, G. G. and Choi, S.C. (1986). Choice of the smoothing parameter and efficiency of  $k$ -nearest neighbor classification. *Computers and Mathematics with Applications A*, 12, 235–244.
- Fix, E. and Hodges, J. (1951). *Discriminatory Analysis, nonparametric discrimination: Consistency properties*. Technical Report, U.S. Air Force, School of Aviation Medicine.
- ISTAT (2007a). *I consumi delle famiglie. Anno 2005*. ISTAT Annuari, Roma.
- ISTAT (2007b). *Indicatori della dinamica dei prezzi al consumo per alcune tipologie di famiglie*. ISTAT Approfondimenti, Roma.



**Grazie per l'attenzione!**