

Analisi dei dati per il marketing (compito 09 gennaio 2025)

Osservazione: occorre avere installato FSDA versione 8.7.7.0

Caricare in memoria il dataset SportHealth tramite l'istruzione

```
load SportHealth.mat
```

Questo dataset una tabella di contingenza relativa alla frequenza con cui gli intervistati hanno dichiarato di fare sport (sulle righe, fenomeno X) e all'autovalutazione dello stato di salute (sulle colonne fenomeno Y).

Calcolare e mostrare nella Command Window l'indice di Somers $d_{Y|X}$ (punti 3)

```
out=corrOrdinal(SportHealth);
```

Test of H₀: independence between rows and columns
The standard errors are computed under H₀

	Coeff	se	zscore	pval
gamma	0.59385	0.053088	11.186	0
taua	0.33958	0.03852	8.8157	0
taub	0.45635	0.040796	11.186	0
tauc	0.45128	0.040343	11.186	0
dyx	0.4525	0.040451	11.186	0

Indexes and 95% confidence limits
The standard error are computed under H₁

	Value	StandardError	ConflimL	ConflimU
gamma	0.59385	0.04837	0.49905	0.68866
taua	0.33958	0.011291	0.31745	0.36171
taub	0.45635	0.040331	0.3773	0.5354
tauc	0.45128	0.040343	0.37221	0.53036
dyx	0.4525	0.040106	0.37389	0.53111

```
disp(['Indice di Somers=' num2str(out.ConfLimtable{'dyx','Value'})])
```

Indice di Somers=0.4525

% oppure

```
disp(out.ConfLimtable('dyx','Value'))
```

```
Value
-----
dyx  0.4525
```

Commentare il valore dell'indice di Somers (punti 3)

```

% v. p. 279
% La differenza tra le proporzioni di coppie concordanti e la proporzione
% di coppie discordanti calcolata con riferimento alle sole coppie di
% osservazioni che non sono a pari merito sulla variabile X è pari a 0.45
% In altri termini l'indice segnala che c'è una relazione diretta e
% crescente
% tra la valutazione dello stato di salute e l'ammontare di sport praticato
% e che questa relazione è pari al 45 per cento del valore massimo possibile

```

Calcolare il numero di coppie concordanti con quelle che presentano la modalità Sometimes e Good Health (punti 3)

```

% La modalità Sometimes e Good Health si trova all'incrocio della riga tre e
% colonna tre
% CoppieConcordantiCon33=numero di coppie concordanti con quelle della riga
% 3 colonna 3
numCella33=SportHealth{3,3};
% 41*39=1599 (v. pp. 275-276)
CoppieConcordantiCon33=numCella33*sum(SportHealth{4:end,4:end}, 'all');

```

Calcolare il numero di coppie discordanti con quelle che presentano la modalità Sometimes e Good Health (punti 3)

```

% 41 * 9 = 369 (v. pp. 275-276)
CoppieDiscordantiCon33=numCella33*sum(SportHealth{4:end,1:2}, 'all');

```

Mostrare nella command Window l'intervallo di confidenza dell'indice $d_{Y|X}$ (punti 3)

```

disp(out.ConfLimtable('dyx', ["ConflimL" "ConflimU"]))

```

	ConflimL	ConflimU
	-----	-----
dyx	0.37389	0.53111

Commentare l'intervallo di confidenza dell'indice $d_{Y|X}$ (punti 3)

```

% L'intervallo non contiene il valore 0 di conseguenza la relazione è
% significativa

```

Effettuare l'analisi delle corrispondenze (punti 3)

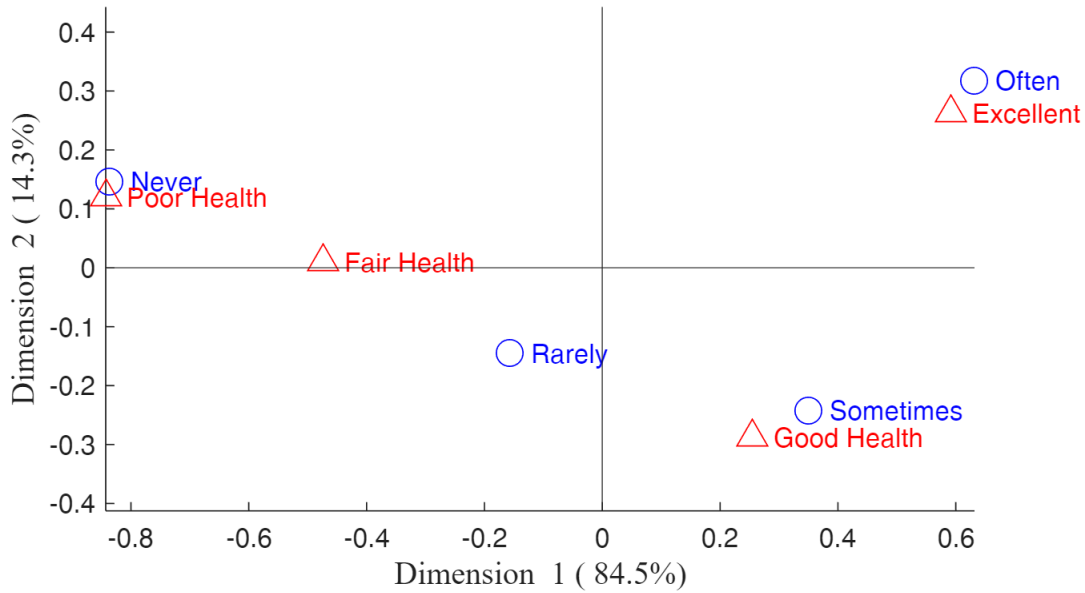
```

outCA=CorAna(SportHealth);

```

French symmetrical model: rows and cols in principal coordinates.

Plot of $X = D_r^{-1/2}U\Gamma$ and $Y = D_c^{-1/2}V\Gamma$



Summary

	<u>Singular_value</u>	<u>Inertia</u>	<u>Accounted_for</u>	<u>Cumulative</u>
dim_1	0.53961	0.29118	0.84484	0.84484
dim_2	0.22205	0.049305	0.14306	0.9879
dim_3	0.064582	0.0041708	0.012102	1

ROW POINTS

Results for dimension: 1

	<u>Scores</u>	<u>CntrbPnt2In</u>	<u>CntrbDim2In</u>
Never	-0.83708	0.55595	0.96747
Rarely	-0.15662	0.023076	0.45076
Sometimes	0.34919	0.11333	0.65833
Often	0.63178	0.30764	0.79769

Results for dimension: 2

	<u>Scores</u>	<u>CntrbPnt2In</u>	<u>CntrbDim2In</u>
Never	0.14676	0.10093	0.029739
Rarely	-0.14477	0.11643	0.38512
Sometimes	-0.24241	0.32254	0.31725
Often	0.31794	0.46011	0.20201

COLUMN POINTS

Results for dimension: 1

	<u>Scores</u>	<u>CntrbPnt2In</u>	<u>CntrbDim2In</u>
Poor Health	-0.84279	0.42669	0.96904
Fair Health	-0.4737	0.16532	0.94866
Good Health	0.25474	0.074286	0.43801
Excellent Health	0.59202	0.3337	0.83537

Results for dimension: 2

	<u>Scores</u>	<u>CntrbPnt2In</u>	<u>CntrbDim2In</u>
--	---------------	--------------------	--------------------

Poor Health	0.12005	0.05113	0.019663
Fair Health	0.009764	0.0004148	0.00040305
Good Health	-0.28784	0.56014	0.55926
Excellent Health	0.26279	0.38831	0.1646

 Overview ROW POINTS

	Mass	Score_1	Score_2	Inertia	CntrbPnt2In_1	CntrbPnt2In_2	CntrbDim
Never	0.23102	-0.83708	0.14676	0.16732	0.55595	0.10093	0.967
Rarely	0.27393	-0.15662	-0.14477	0.014906	0.023076	0.11643	0.450
Sometimes	0.27063	0.34919	-0.24241	0.050126	0.11333	0.32254	0.658
Often	0.22442	0.63178	0.31794	0.1123	0.30764	0.46011	0.797

Overview COLUMN POINTS

	Mass	Score_1	Score_2	Inertia	CntrbPnt2In_1	CntrbPnt2In_2	C
Poor Health	0.17492	-0.84279	0.12005	0.12821	0.42669	0.05113	
Fair Health	0.21452	-0.4737	0.009764	0.050743	0.16532	0.0004148	
Good Health	0.33333	0.25474	-0.28784	0.049382	0.074286	0.56014	
Excellent Health	0.27723	0.59202	0.26279	0.11631	0.3337	0.38831	

 Legend

Row scores in principal coordinates

Column scores in principal coordinates

CntrbPnt2In = relative contribution of points to explain total Inertia of the latent dimension
 The sum of the numbers in a column is equal to 1

CntrbDim2In = relative contribution of latent dimension to explain total Inertia of a point
 $CntrbDim2In_1 + CntrbDim2In_2 + \dots + CntrbDim2In_K = 1$

Interpretare la prima dimensione latente (punti 3)

```
% La prima dimensione latente è sia il livello di sport, sia lo stato di
% salute. Andando da sx verso destra aumenta sia il livello di sport sia la
% condizione di salute
```

Interpretare la posizione della coppia Often ExcellentHealth (punti 3)

```
% Chi pratica molto sport generalmente presenta una condizione
% di salute molto buona
```

Calcolare l'inerzia totale della tabella di contingenza $in(I)=in(J)$, tramite l'equazione 13.14 oppure tramite la somma dei quadrati delle distanze di ogni punto riga dal centroide moltiplicata per la rispettiva massa (equazione 13.5) (punti 5)

```
% Inerzia totale tramite l'equazione (13.14)
In1=outCA.Chi2stat/sum(SportHealth{:, :}, 'all');
disp(In1)
```

0.3447

```
% Inerzia totale tramite equazione (13.4)
```

```
% Vedi il facsimile dell'Esercizio di riepilogo 13.4 nel sito di  
% Giappichelli  
In1chk=sum(outCA.OverviewRows.Inertia)
```

```
In1chk =  
0.3447
```