

# **Corso di Modelli per l'analisi statistica**

Prof. G. Scandurra  
a.a. 2020-2021

# La rappresentazione dei dati

Il manager è consapevole che presentare le informazioni raccolte in forma di matrice dei dati non ha senso

È utile invece rappresentarle in forma organizzata e sintetica allo scopo di:

- evidenziarne le caratteristiche principali
- facilitarne la lettura e l'interpretazione

➔ Rappresentazione tabellare

➔ Rappresentazione grafica

# Rappresentazioni tabellari

Distribuzione di frequenze

Distribuzione di quantità

Serie storica

Serie territoriale

# Distribuzione di frequenze

Organizzazione dei dati mediante una tabella risultante dalle operazioni di:

- Classificazione
- Conteggio

Ad ogni modalità di un carattere (qualitativo o quantitativo) si fa corrispondere il numero di volte che esso si presenta nel collettivo (la sua frequenza assoluta)

# Distribuzione semplice di frequenze

X	Freq.
$x_1$	$n_1$
$x_2$	$n_2$
...	...
$x_j$	$n_j$
...	...
$x_k$	$n_k$
<b>Totale</b>	<b>n</b>

$x_1, x_2, \dots, x_k$   
sono le modalità  
distinte che  
assume il  
carattere X nel  
collettivo di n  
unità esaminato

$n_1, n_2, \dots, n_k$   
sono le freq.  
assolute  
associate a  
ciascuna  
modalità

$n_1$  indica  
quante unità  
presentano la  
modalità  $x_1$  del  
carattere X

La somma delle frequenze  
assolute è uguale al numero  
totale di unità del collettivo

$$\sum_{j=1}^k n_j = n$$

# Distribuzione di frequenze

Punti vendita	Addetti
1	6
2	6
3	10
4	10
5	7
6	3
7	3
8	6
9	4

Addetti (valori distinti)	Numero punti vendita (frequenze)
3	2
4	1
6	3
7	1
10	2

- Quanti sono i punti vendita con 3 addetti? 2  
Quanti sono i punti vendita con 4 addetti? 1  
Quanti sono i punti vendita con 6 addetti? 3  
Quanti sono i punti vendita con 7 addetti? 1  
Quanti sono i punti vendita con 10 addetti? 2

# Frequenze relative e frequenze relative percentuali

La frequenza relativa è data dal rapporto tra frequenza assoluta e numero totale di unità del collettivo

per la  $j$ -esima modalità  $f_j = \frac{n_j}{n}$

Vale che  $\sum_{j=1}^K f_j = 1$

La frequenza relativa percentuale altro non è che la frequenza relativa moltiplicata per 100

$$p_j = f_j \cdot 100 = \frac{n_j}{n} \cdot 100 \quad \sum_{j=1}^K p_j = 100$$

# Calcolo delle frequenze relative e percentuali

<b>Addetti</b>	<b>Frequenze assolute</b>	<b>Frequenze relative</b>	<b>Frequenze rel. perc.</b>
<b>3</b>	<b>2</b>	<b><math>2/9=0,22</math></b>	<b>22,2</b>
<b>4</b>	<b>1</b>	<b><math>1/9=0,11</math></b>	<b>11,1</b>
<b>6</b>	<b>3</b>	<b><math>3/9=0,33</math></b>	<b>33,3</b>
<b>7</b>	<b>1</b>	<b><math>1/9=0,11</math></b>	<b>11,1</b>
<b>10</b>	<b>2</b>	<b><math>2/9=0,22</math></b>	<b>22,2</b>
<b>Tot</b>	<b>n=9</b>	<b>1,00</b>	<b>100,0</b>

La somma delle freq. rel. è pari a 1 e la somma delle freq. rel. perc. è pari a 100 (in questo caso sono state arrotondate perché risultavano pari rispettivamente a 0,99 e a 99,9)

I punti vendita con 3 addetti sono 2 (freq. ass.)  
Rappresentano il 22% del totale dei punti vendita



# Perché si calcolano le frequenze relative e percentuali?

Le frequenze assolute dipendono da  $n$   
Quindi non possono essere utilizzate per effettuare confronti tra collettivi con diversa numerosità

Al contrario, le frequenze relative e quelle percentuali sono numeri puri  
Si utilizzano per confrontare distribuzioni di frequenza riferite a collettivi di diversa numerosità

# Confronto tra distrib. di frequenze

## Campania

Addetti	Freq. ass.	Freq. rel. perc.
3	2	22,2
4	1	11,1
6	3	33,3
7	1	11,1
10	2	22,2
<b>Tot</b>	<b>n=9</b>	<b>100</b>

## Lazio

Addetti	Freq. ass.	Freq. rel. perc.
3	4	8,3
4	4	8,3
5	11	22,9
6	15	31,3
8	8	16,7
10	6	12,5
<b>Tot</b>	<b>n=48</b>	<b>100,0</b>

In termini di freq. rel. perc., i 2 p.v. con 3 addetti costituiscono il 22,2 % del totale dei p.v. in Campania e solo l'8,3% del totale dei p.v. nel Lazio

L'incidenza dei p.v. con pochi addetti è maggiore in Campania, come si supponeva prima di conoscere i dati

# Creazione di una tabella pivot con Excel (1/2)

Da una matrice dei dati su un foglio di lavoro Excel, è possibile creare una tabella (distr. di freq. o di quant.) selezionando dal menu "Dati" l'opzione "Rapporto Tabella Pivot"

Seguono vari passaggi:

Indico dove si trovano i dati da utilizzare (area del foglio Excel)

Indico dove voglio che sia prodotta la tabella  
Dalla finestra di dialogo "Layout" specifico la variabile di classificazione (ad es. VAR1) trascinando il corrispondente pulsante nell'area "RIGA"

# Creazione di una tabella pivot con Excel (2/2)

- Se voglio la distribuzione di frequenza di VAR1 trascino il pulsante di VAR1 nell'area "DATI"

Ci faccio doppio clic e nel riquadro "Riepiloga Per" seleziono "Conteggio"

- Se voglio la distribuzione di quantità (somma) di VAR2 trascino il pulsante di VAR2 nell'area "DATI"

Ci faccio doppio clic e nel riquadro "Riepiloga Per" seleziono "Somma"

# Serie storica

Tabella che ad ogni riferimento temporale (ad esempio, l'anno, il mese, il giorno) associa l'ammontare del carattere X in esame

Evidenzia la dinamica di un certo fenomeno nel tempo

Esempi:

il valore aggiunto di un'azienda negli ultimi cinque anni

l'indice S&P/Mib alla Borsa di Milano nell'ultima settimana

# Serie storica (esempio)

Facendo riferimento al nostro esempio base, la banca può richiedere il R.O. (risultato operativo) di ogni punto vendita degli ultimi quattro anni

Per ogni punto vendita si ha una serie storica del tipo:

anni	R.O. (migliaia euro)
2004	85
2005	120
2006	215
2007	161

# Serie territoriale

Tabella che ad ogni unità territoriale (ad esempio paese, regione, distretto industriale)

fa corrispondere l'ammontare del carattere X in esame

Mostra la distribuzione del fenomeno in rapporto al territorio

Esempi:

il tasso di inflazione nei paesi UE

le emissioni di CO<sub>2</sub> nei capoluoghi di regione italiani

# Serie territoriale (esempio)

Valori del PIL pro-capite in alcuni Paesi  
(Dati del Fondo Monetario Internazionale 2007)

Paese	PIL nominale (in dollari USA)
Italia	31.802
Spagna	27.951
Regno Unito	39.681
Svezia	43.190



# Esercizi di riepilogo

## **Esercizio 1.**

Supponete di disporre dei seguenti dati del fatturato in migliaia di euro di un'azienda

120 123 221 135 146 123 167 123 123 121  
135 136 136 221 222 223 167 135 135 121

Costruire la tabella di frequenza in classi.

- Decidete di costruire tre classi

# Distribuzione in classi del fatturato

Classi di fatturato	$n_j$	$f_j$	$p_j$
(110-130]	7	0.35	35%
(130-200]	9	0.45	45%
(200+	4	0.20	20%