

# STATISTICA 1 ESERCITAZIONE 1

*Dott. Giuseppe Pandolfo*

*30 Settembre 2013*

**Popolazione statistica:** insieme degli elementi oggetto dell'indagine statistica.

**Unità statistica:** ogni elemento della popolazione statistica.

**Variabile casuale:** il fenomeno collettivo oggetto dell'indagine statistica.

**Modalità:** il modo in cui si presenta la variabile casuale nelle diverse unità statistiche.

**Osservazione:** Il valore assunto dalla variabile casuale in una data unità statistica.

## CLASSIFICAZIONE DELLE VARIABILI CASUALI

- **Variabile quantitativa:** assume valori numerici
  - **Continua:** assume valori continui in un intervallo (altezza, peso, ecc.)
  - **Discreta:** assume valori discreti (numero di pazienti, numero di figli, ecc.)
  
- **Variabile qualitativa:** assume valori non numerici
  - **Ordinale:** i dati sono in ordine (titolo di studio)
  - **Nominale:** uomo/donna, basso/medio/alto, ecc.

### Esercizio 1

Indicare per ognuna delle seguenti variabili se sono quantitative (discrete o continue) o qualitative (nominali o ordinali).

1. Un negozio vende 3 tipi di bevande: bibite, tè e caffè
2. Il tempo di volo da Roma a Milano
3. Numero di telefoni per famiglia
4. Numero di telefonate per mese
5. Sesso (Uomo/Donna)
6. Possesso di un computer (SI/NO)
7. Titolo di studio
8. Grado di soddisfazione

## Soluzione

1. Un negozio vende 3 tipi di bevande: bibite, tè e caffè → Variabile qualitativa nominale
2. Il tempo di volo da Roma a Milano → Variabile quantitativa continua
3. Numero di telefoni per famiglia → Variabile quantitativa discreta
4. Numero di telefonate per mese → Variabile quantitativa discreta
5. Sesso (Uomo/Donna) → Variabile qualitativa nominale
6. Possesso di un computer (SI/NO) → Variabile qualitativa nominale
7. Titolo di studio → Variabile qualitativa ordinale
8. Grado di soddisfazione → Variabile qualitativa ordinale

## **MATRICI DEI DATI**

I dati rilevati su  $n$  unità statistiche con riferimento a  $p$  variabili vengono raccolti in una tabella chiamata *matrice dei dati*.

**Esempio:** la seguente tabella riporta i dati relativi a tre unità statistiche per tre variabili casuali.

	<b>Sesso</b>	<b>Titolo di studio</b>	<b>Età</b>
Marco	M	Laurea	34
Laura	F	Diploma	24
Stefania	F	Laurea	63

Unità statistica: singola persona (3 unità statistiche).

Variabili casuali: Sesso (qualitativa nominale), titolo di studio (qualitativa ordinale), età (quantitativa continuo).

## **Esercizio 2**

Costruire le distribuzioni di frequenze relative e percentuali per ogni variabile considerata.

<b>Gruppo sanguigno</b>	<b><math>n_i</math></b>
0	9
AB	3
A	3
B	5
Totale	20

<b>Altezza (cm)</b>	<b><math>n_i</math></b>
150	4
160	9
170	15
180	12
Totale	40

## Soluzione

Gruppo sanguigno	$n_i$	$f_i$	$p_i$
0	9	0,45	45%
AB	3	0,15	15%
A	3	0,15	15%
B	5	0,25	25%
Totale	20	1,00	100%

Altezza (cm)	$n_i$	$f_i$	$p_i$
150	4	0,1	10%
160	9	0,225	22,5%
170	15	0,375	37,5%
180	12	0,3	30%
Totale	40	1,00	100%

### TABELLE A DOPPIA ENTRATA

Una tabella a doppia entrata contiene le frequenze congiunte  $n_{ij}$  delle diverse coppie di modalità ( $i$ -esima per la prima variabile e  $j$ -esima per la seconda variabile) di 2 variabili.

Cosa contiene una tabella a doppia entrata:

- Frequenze congiunte
- Frequenze marginali
- Frequenze condizionate

#### Esercizio 3

La seguente tabella a doppia entrata riporta i dati relativi a 200 unità per le variabili *Fumo* e *BPCO* (affezione cronica polmonare).

Fumo	BPCO (affezione cronica polmonare)		Totale
	Assente	Presente	
<i>Fumatore</i>	119	2	112
<i>Ex fumatore</i>	16	12	28
<i>Non fumatore</i>	24	36	60
Totale	150	50	200

Indicare le frequenze congiunte:

119	2
16	12
24	36
	200

Indicare le frequenze marginali della variabile *Fumo*:

<b>Fumo</b>	
Non fumatore	112
Ex fumatore	28
Fumatore	60
Totale	200

Indicare le frequenze marginali della variabile *BPCO*:

<b>BPCO</b>	
Assente	150
Presente	50
Totale	200

Indicare le frequenze della variabile *Fumo* condizionata alla modalità *assente* della variabile *BPCO*:

<b>Fumo</b>	<b>BPCO = Assente</b>
Non fumatore	110
Ex fumatore	16
Fumatore	24
Totale	150

Indicare le frequenze della variabile *BPCO* condizionata alla modalità *Ex fumatore* della variabile *Fumo*:

<b>BPCO</b>	<b>Fumo = Ex fumatore</b>
Assente	16
Presente	12
Totale	28

## LE DISTRIBUZIONI IN CLASSI

Il numero delle classi dipende dal numero delle unità statistiche. Generalmente, un maggiore (minore) numero di unità statistiche richiede un maggiore (minore) numero di classi.

### Esercizio 4

Per la variabile *Altezza* suddividiamo la distribuzione in 6 classi equiampie riportando le frequenze assolute e le frequenze relative.

<b>Altezza (cm)</b>	<b><math>n_i</math></b>
20	1
21	1
23	1
24	2
27	2
29	2
30	3
33	8

35	1
38	3

Soluzione

L'ampiezza delle classi si ottiene

$$\text{Ampiezza} = \frac{\text{range(Variabile)}}{\text{numero di classi}}$$

Per la variabile *Altezza*

$$\text{Ampiezza} = \frac{\text{range(Altezza)}}{6} = \frac{38 - 20}{6} = 3$$

<b>Altezza (cm)</b>	<b><math>n_i</math></b>	<b><math>f_i</math></b>	<b><math>p_i</math></b>
[20-23]	2	0,08	8%
(23-26]	3	0,13	13%
(26-29]	2	0,08	8%
(29-32]	5	0,21	21%
(32-35]	8	0,33	33%
(35-38]	4	0,17	17%
Totale	24	1,00	100%

**FREQUENZE CUMULATE**

Per variabili quantitative e qualitative ordinali.

**Esercizio 5**

La seguente tabella riporta i dati relativi alla variabili *Voto esame* per 350 studenti.

Calcolare le frequenze relative, le frequenze assolute cumulate e le frequenze relative cumulate.

<b>Voto esame</b>	<b><math>n_i</math></b>
< 18	80
18	70
19-22	110
22-25	55
>25	35
Totale	350

Soluzione

<b>Voto esame</b>	<b><math>n_i</math></b>	<b><math>f_i</math></b>	<b><math>N_i</math></b>	<b><math>F_i</math></b>
< 18	80	0,23	80	0,23
18	70	0,20	150	0,43
19-22	110	0,31	260	0,74
22-25	55	0,16	315	0,9
>25	35	0,1	350	1,00
Totale	350	1,00		