

Classificazione dei caratteri

- ✗ **Carattere Qualitativo** (*Mutabile Statistica*): modalità espresse da attributi
 - Genere, Stato civile, Sett. di attività econ., Titolo di studio, Grado militare
- ✗ **Carattere Quantitativo** (*Variabile Statistica*): modalità espresse da numeri
 - *Discreto*: numero finito, o numerabile, di modalità
 - Num. di figli, Num. di pezzi prodotti
 - *Continuo*: numero infinito di modalità
 - Peso, Altezza, Reddito

Pagina 1

Misurazione dei caratteri

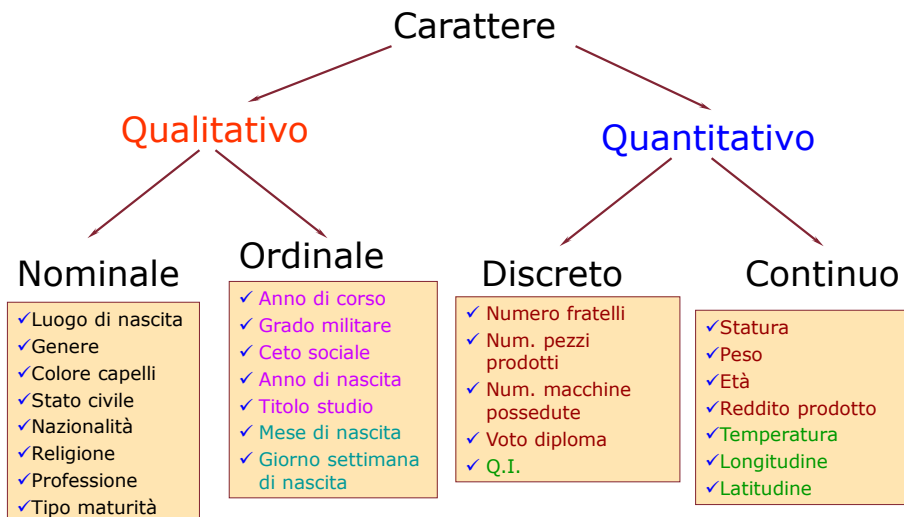
- ✗ Carattere qualitativo
 - *Scala nominale o Mutabile Sconnessa*
 - Se date due sue modalità è possibile affermare solo se queste sono uguali o diverse
 - Nessun ordine tra le modalità
 - Sesso, Stato civile, Sett. attività econ.
 - *Scala Ordinale o Mutabile Ordinata*
 - Esiste un ordine naturale tra le modalità
 - se date due sue modalità è possibile affermare se queste sono uguali e stabilire un ordinamento **oggettivo**, specificando quale precede l'altra
 - *Mutabili rettilinee*: Tit. studio, Grado militare
 - *Mutabili cicliche*: Mese di nascita, Giorno della settimana di nascita

Pagina 2

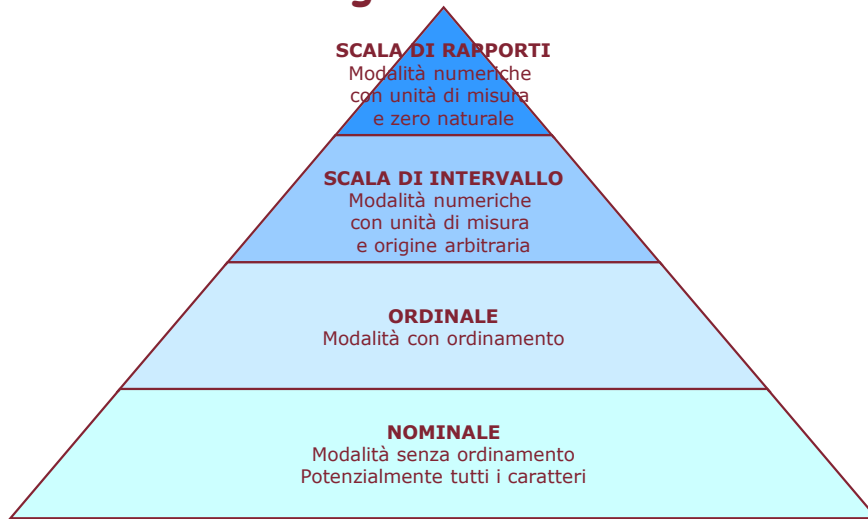
...continua

- ✗ **Caratteri quantitativi:** esiste un'unità di misura delle modalità numeriche
 - *Variabile a scala di intervalli*
 - non esiste uno zero assoluto, naturale e non arbitrario
 - Operazioni = $0 \neq$ e $> 0 <$ e distanza
 - Temperatura, Q.I.
 - *Variabile a scala di rapporti*
 - esiste uno zero assoluto, naturale e non arbitrario
 - Operazioni = $0 \neq$ e $> 0 <$, distanza e rapporto
 - Nati, Morti, Pezzi prodotti, Reddito, Peso

In sintesi

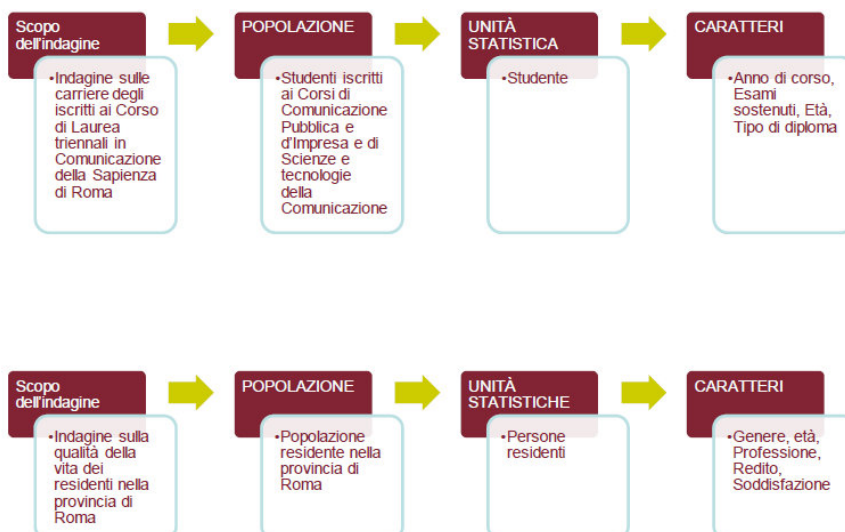


Classificazione gerarchica



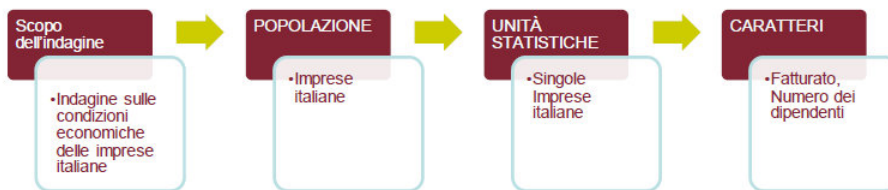
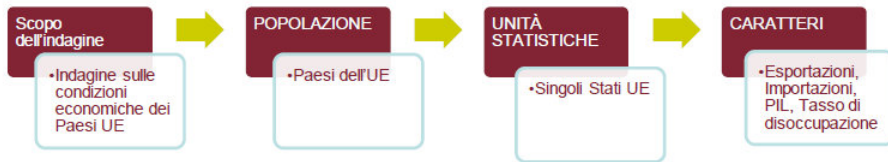
Pagina 5

Popolazione, Unità e Caratteri



Pagina 6

Popolazione, Unità e Caratteri



Pagina 7

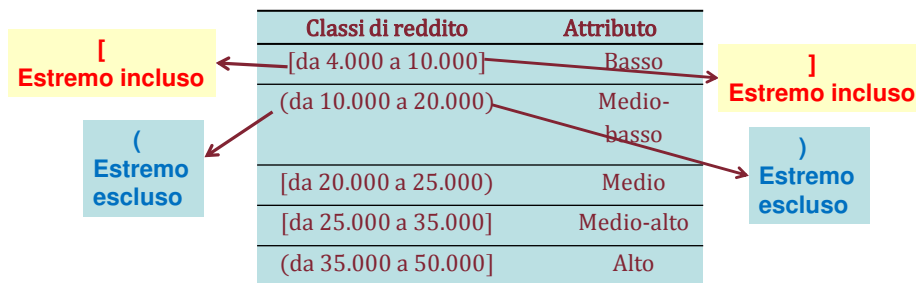
Trasformazioni della scala di misura

DA	A	Nominale	Ordinale	Scala di intervalli	Scala di rapporti
Nominale			Introdurre parametri	Introdurre parametri	Introdurre parametri
Ordinale		Non tenere conto dell'ordine		Introdurre parametri	Introdurre parametri
Scala di intervalli		Suddividere in classi e non tenere conto dell'ordine	Suddividere in classi		
Scala di rapporti		Suddividere in classi e non tenere conto dell'ordine	Suddividere in classi		

Pagina 8

Trasformazione di un carattere: da quantitativo a qualitativo ordinale

- ✓ Supponiamo di aver rilevato il reddito annuo lordo su un collettivo di occupati : il reddito minimo è 4.000 euro, il massimo 50.000 euro
- ✓ Si tratta di un carattere quantitativo a scala di rapporti
- ✓ Può essere trasformato in una mutabile ordinale raggruppando in classi le modalità numeriche ed associando a ciascuna di esse un attributo
- ✓ Le classi devono essere esaustive e non sovrapposte



Pagina 9

Trasformazione di un carattere: da qualitativo a quantitativo

- ✓ Supponiamo di aver rilevato la mutabile ordinale "titolo di studio" in un collettivo di occupati, distinguendo nessun titolo, licenza elementare, media inferiore, diploma, laurea triennale, laurea magistrale
- ✓ La mutabile ordinale può essere trasformata in carattere quantitativo facendo corrispondere a ciascuna modalità il numero di anni scolastici necessari per conseguire il titolo corrispondente

Modalità attributo	Modalità numerica
Nessun titolo	0
Licenza elementare	5
Licenza media inferiore	8
Diploma	13
Laurea triennale	16
Laurea magistrale	18

Pagina 10

Riepilogo sui caratteri

	= o ≠	> o <	+ o -	x o /
Sconnessa	si	no	no	no
Ordinale	si	si	no	no
Scala intervalli	si	si	si	no
Scala rapporti	si	si	si	si


Pagina 11

Tabelle di dati

- ✗ Tabella di dati: insieme di informazioni organizzate secondo diverse righe e diverse colonne
- ✗ Tabella "unità x caratteri" (o tabella "unità x variabili" o *matrice dei dati*)
 - ciascuna *riga* è associata ad una u.s.
 - ciascuna *colonna* è associata ad un carattere
 - una *cella* è la realizzazione del carattere (in colonna) su una u.s. (in riga)
- ✗ Matrice dei dati: può essere una tabella di *microdati* o di *macrodati*

Pagina 12

Trasformazione delle unità statistiche

- ✘ La manipolazione dei dati può riguardare non soltanto i caratteri osservati, ma anche le unità statistiche che possono essere aggregate secondo criteri utili alle finalità dell'indagine
- ✘ Si è rilevato su alcuni individui laureati alcuni caratteri, si può passare
 - ✓ dalle unità statistiche *laureati*
 - 
 - ✓ alle unità statistiche *corsi di laurea*

Pagina 13

La tabella dei microdati

Unità statistica di rilevazione	Corso di laurea	Conclusione del percorso di studio	Voto di laurea
Intervistato 1	Economia	In corso	100
Intervistato 2	Economia	Fuori corso	106
Intervistato 3	Comunicazione	In corso	110
Intervistato 4	Comunicazione	Fuori corso	104
Intervistato 5	Comunicazione	Fuori corso	101
Intervistato 6	Ingegneria	In corso	90
Intervistato 7	Ingegneria	In corso	110

Pagina 14

Trasformazione delle unità statistiche

- ✘ Si vuole focalizzare l'analisi non tanto sui singoli laureati quanto sui corsi di laurea, per far emergere eventuali analogie e differenze in merito ai voti e alla possibilità di licenziare laureati in corso
- ✘ I dati statistici che si andranno ad analizzare non saranno quelli riguardanti i singoli intervistati, ma quelli ottenuti attraverso opportune elaborazioni dei microdati

La tabella dei macrodati

Unità statistica di analisi	Numero laureati in corso	Numero laureati fuori corso	Voto medio di laurea
Economia	1	1	103
Comunicazione	1	2	105
Ingegneria	2	0	100

Matrice unità di analisi x caratteri

- ✘ I laureati sono aggregati in gruppi in base alla medesima modalità del carattere "corso di laurea"
 - tanti gruppi quante sono le modalità del carattere "corso di laurea" → **unità di analisi**
- ✘ Per gli altri caratteri si sceglie una appropriata funzione di aggregazione delle modalità che tenga conto della tipologia del carattere
 - per la mutabile dicotomica "conclusione del percorso di studio" si contano, per ogni unità di analisi, i laureati in corso e quelli fuori corso
 - per la variabile a scala di rapporti "voto di laurea" si calcola il voto medio alla laurea per ciascuna unità di analisi

Pagina 17

Microdati e macrodati

- ✘ **Unità statistica di rilevazione**: unità elementare su cui vengono rilevati i caratteri
 - danno luogo ai **microdati**: insieme di dati riguardanti le unità elementari
- ✘ **Unità statistica di analisi**: unità primaria di interesse per l'analisi, a cui vengono riferiti i caratteri oggetto di studio
 - può coincidere con quella di rilevazione
 - può riferirsi ad aggregazioni di tali unità (la famiglia, la scuola, l'azienda, il corso)
 - da luogo ai **macrodati**, insieme di dati riguardanti le unità aggregate

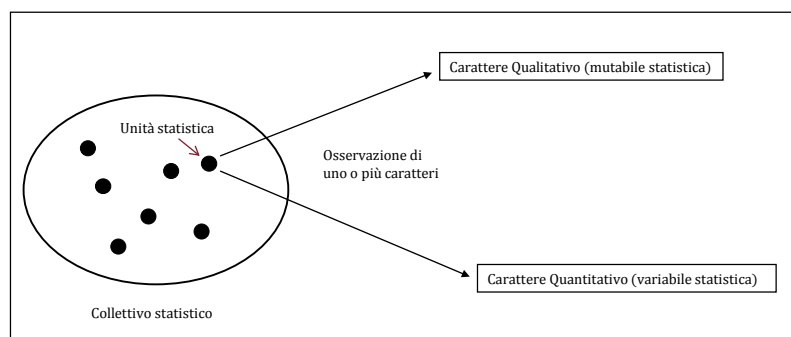
Pagina 18

Tabelle di dati

- ✗ Tabella di dati: insieme di informazioni organizzate secondo diverse righe e diverse colonne
- ✗ Tabella "unità x caratteri" (o tabella "unità x variabili" o *matrice dei dati*)
 - ciascuna *riga* è associata ad una u.s.
 - ciascuna *colonna* è associata ad un carattere
 - una *cella* è la realizzazione del carattere (in colonna) su una u.s. (in riga)
- ✗ Matrice dei dati: può essere una tabella di *microdati* o di *macrodati*

Pagina 19

La struttura dei dati



Pagina 20

Rilevazione dei dati statistici

Da un'indagine condotta ad un anno dalla laurea sui laureati nell'a.a. 2003/2004 ad un corso di laurea triennale della facoltà di Scienze della Comunicazione, si sono ottenuti i seguenti dati:

A	B	C	D	E	F	G	H	I	J	K	M	N	P	Q	
UNTA'	genere	LaureaNO	Test	VotoL	Durata	LavOggi	Settore	Qualifica	Contratto	OreLav	LavAdeguate	UnivOggi	Solo	Reddito	
u1	F	STC	4	101	3	NO, non cerco		0	0	0	0	3	0	0	
u2	M	STC	4	111	3	SI		11	22	3	10	1	1	NO	0
u3	F	SCPO	7	110	3	SI		40	12	3	40	1	3	SI	5
u4	F	STC	3	108	3	NO, in cerca		0	0	0	0	0	3	0	0
u5	M	SCPO	12	99	3	NO, non cerco		0	0	0	0	0	1	0	0
u6	F	COOP	6	110	3	SI		17	22	5	18	1	3	0	1
u7	M	COOP	13	96	3	SI		30	12	4	40	1	3	SI	2
u8	M	STC	4	107	3	SI		15	7	1	12	1	1	SI	1
u9	F	SCPO	7	104	3	NO, non cerco		0	0	0	0	0	2	0	0
u10	F	STC	3	110	3	NO, non cerco		0	0	0	0	0	3	0	0
u11	M	SCPO	16	105	3	SI		14	2	7	16	1	1	SI	3
u12	F	SCPO	4	108	3	NO, non cerco		0	0	0	0	0	1	0	0
u13	F	SCPO	11	108	3	NO, non cerco		0	0	0	0	0	2	0	0
u14	F	SCPO	4	102	4	NO, non cerco		0	0	0	0	0	1	0	0
u15	F	STC	4	103	3	NO, non cerco		0	0	0	0	0	1	0	0
u16	F	SCPO	9	108	4	SI		45	7	6	20	1	1	SI	1
u17	F	SCPO	7	103	3	NO, in cerca		0	0	0	0	0	3	0	0
u18	F	SCPO	11	111	3	NO, in cerca		0	0	0	0	0	1	0	0
u19	M	SCPO	11	105	3	NO, in cerca		0	0	0	0	0	3	0	0
u20	F	STC	2	103	3	NO, non cerco		0	0	0	0	0	1	0	0
u21	F	SCPO	9	109	4	NO, non cerco		0	0	0	0	0	1	0	0
u22	M	STC	1	111	5	SI		18	23	8	36	1	1	SI	5
u23	M	STC	3	104	3	NO, non cerco		0	0	0	0	0	1	0	0
u24	F	SCPO	2	111	4	NO, non cerco		0	0	0	0	0	3	0	0
u25	M	SCPO	5	109	3	NO, non cerco		0	0	0	0	0	1	0	0
u26	F	STC	3	107	5	SI		15	1	6	x	2	1	NO	2

Pagina 21

Matrice dei dati : microdati

A	B	C	D	E	F	G	H	I	J	K	M	N	P	Q	
UNTA'	genere	LaureaNO	Test	VotoL	Durata	LavOggi	Settore	Qualifica	Contratto	OreLav	LavAdeguate	UnivOggi	Solo	Reddito	
u1	F	STC	4	101	3	NO, non cerco		0	0	0	0	3	0	0	
u2	M	STC	4	111	3	SI		11	22	3	10	1	1	NO	0
u3	F	SCPO	7	110	3	SI		40	12	3	40	1	3	SI	5
u4	F	STC	3	108	3	NO, in cerca		0	0	0	0	0	3	0	0
u5	M	SCPO	12	99	3	NO, non cerco		0	0	0	0	0	1	0	0
u6	F	COOP	6	110	3	SI		17	22	5	18	1	3	0	1
u7	M	COOP	13	96	3	SI		30	12	4	40	1	3	SI	2
u8	M	STC	4	107	3	SI		15	7	1	12	1	1	SI	1
u9	F	SCPO	7	104	3	NO, non cerco		0	0	0	0	0	2	0	0
u10	F	STC	3	110	3	NO, non cerco		0	0	0	0	0	3	0	0
u11	M	SCPO	16	105	3	SI		14	2	7	16	1	1	SI	3
u12	F	SCPO	4	108	3	NO, non cerco		0	0	0	0	0	1	0	0
u13	F	SCPO	11	108	3	NO, non cerco		0	0	0	0	0	2	0	0
u14	F	SCPO	4	102	4	NO, non cerco		0	0	0	0	0	1	0	0
u15	F	STC	4	103	3	NO, non cerco		0	0	0	0	0	1	0	0
u16	F	SCPO	9	108	4	SI		45	7	6	20	1	1	SI	1
u17	F	SCPO	7	103	3	NO, in cerca		0	0	0	0	0	3	0	0
u18	F	SCPO	11	111	3	NO, in cerca		0	0	0	0	0	1	0	0
u19	M	SCPO	11	105	3	NO, in cerca		0	0	0	0	0	3	0	0
u20	F	STC	2	103	3	NO, non cerco		0	0	0	0	0	1	0	0
u21	F	SCPO	9	109	4	NO, non cerco		0	0	0	0	0	1	0	0
u22	M	STC	1	111	5	SI		18	23	8	36	1	1	SI	5
u23	M	STC	3	104	3	NO, non cerco		0	0	0	0	0	1	0	0
u24	F	SCPO	2	111	4	NO, non cerco		0	0	0	0	0	3	0	0
u25	M	SCPO	5	109	3	NO, non cerco		0	0	0	0	0	1	0	0
u26	F	STC	3	107	5	SI		15	1	6	x	2	1	NO	2

Pagina 22

Distribuzione unitaria

- ✗ Tabella "unità x variabili" è una distribuzione unitaria multipla
 - *unitaria* perché è rappresentata ciascuna u.s.
 - *multipla* perché sono osservati più caratteri su una medesima u.s.

- ✗ Una sola colonna della tabella → *distribuzione unitaria semplice*

Esempio: Distribuzione ...

Unitaria **Multipla**

A	B	C	D	E	F
UNITA'	sezzo	LaureaNO	Tesi	VotoL	Durata
u1	F	STC	4	101	3
u2	M	STC	4	111	3
u3	F	SCPO	7	110	3
u4	F	STC	3	108	3
u5	M	SCPO	12	99	3
u6	F	COOP	6	110	3
u7	M	COOP	13	96	3
u8	M	STC	4	107	3
u9	F	SCPO	7	104	3
u10	F	STC	3	110	3
u11	M	SCPO	16	105	3
u12	F	SCPO	4	108	3
u13	F	SCPO	11	108	3
u14	F	SCPO	4	102	4
u15	F	STC	4	103	3
u16	F	SCPO	9	108	4
u17	F	SCPO	7	103	3
u18	F	SCPO	11	111	3
u19	M	SCPO	11	105	3
u20	F	STC	2	103	3
u21	F	SCPO	9	109	4
u22	M	STC	1	111	5
u23	M	STC	3	104	3
u24	F	SCPO	2	111	4
u25	M	SCPO	5	109	3
u26	F	STC	3	107	5

Unitaria **Semplice**

A	B
UNITA'	sezzo
u1	F
u2	M
u3	F
u4	F
u5	M
u6	F
u7	M
u8	M
u9	F
u10	F
u11	M
u12	F
u13	F
u14	F
u15	F
u16	F
u17	F
u18	F
u19	M
u20	F
u21	F
u22	M
u23	M
u24	F
u25	F
u26	F

La Distribuzione Statistica di ... Frequenze Assolute

- ✗ Necessità di sintesi
- ✗ Lista delle modalità che il carattere può assumere
- ✗ Conteggio del numero di u.s. del collettivo considerato che presentano la singola modalità (freq. Assoluta)

CARATTERE	FREQUENZE ASSOLUTE
Modalità uno	Numero di unità statistiche del collettivo considerato che presentano la modalità uno del carattere
Modalità due	Numero di unità statistiche del collettivo considerato che presentano la modalità due del carattere
...	...
Modalità kappa	Numero di unità statistiche del collettivo considerato che presentano la modalità kappa del carattere

Pagina 25

La notazione

X : Carattere osservato

$x_1, x_2, \dots, x_i, \dots, x_K$: K modalità assunte del carattere X

➤ x_i modalità generica

➤ x_K ultima modalità

➤ Se X quantitativo

$$x_1 < x_2 < \dots < x_i < \dots < x_K$$

n : numero di u.s. che compongono il collettivo

$n_1, n_2, \dots, n_i, \dots, n_K$: freq. Assol. associate alle K modalità

$$n_1 + n_2 + \dots + n_i + \dots + n_K = n$$

Car. X	FREQUENZE ASSOLUTE
x_1	n_1
x_2	n_2
...	...
x_i	n_i
...	...
x_K	n_K
Totale	n

Pagina 26

...la formalizzazione

CARATTERE	FREQUENZE ASSOLUTE
Modalità uno	Numero di u.s. del collettivo considerato che presentano la modalità uno del carattere
Modalità due	Numero di u.s. del collettivo considerato che presentano la modalità due del carattere
...	...
Modalità i	Numero di u.s. del collettivo considerato che presentano la modalità i del carattere
...	...
Modalità kappa	Numero di u.s. del collettivo considerato che presentano la modalità kappa del carattere

Car. X	FREQUENZE ASSOLUTE
x_1	n_1
x_2	n_2
...	...
x_i	n_i
...	...
x_K	n_K
Totale	n

Pagina 27

Distribuzione dei laureati per genere

- ✗ Distribuzione per genere dei laureati di Scienze della Comunicazione nell'a.a. 2003/2004

Genere	Frequenze assolute
F	17
M	9
Totale	26

Pagina 28

Condizione occupazionale dei laureati

- ✘ Distribuzione per condizione occupazionale dei laureati di Scienze della Comunicazione nell'a.a. 2003/2004

Condizione occupazionale	Frequenze assolute
SI	9
NO, non cerco	13
No, ma cerco	4
Totale	26

Pagina 29

La Distribuzione Statistica di ... Frequenze Relative

- ✘ Consente di confrontare l'informazione relativa ad un medesimo carattere osservato su collettivi di numerosità diversa
- ✘ Frequenze Relative
 - si ottengono dividendo, per ogni modalità, la Frequenza Assoluta corrispondente per la Numerosità Totale del collettivo
 - In sintesi:

$$f_i = \frac{n_i}{n}$$

- $f_1 + f_2 + \dots + f_i + \dots + f_K = 1$

Pagina 30

Distribuzione dei laureati occupati per Durata del Corso di Studi

- ✗ Distribuzione dei laureati di Scienze della Comunicazione nell'a.a. 2003/2004 per Durata del Corso di Studi

Durata Corso di Studi	Frequenze assolute	Frequenze relative
3	20	0.77
4	4	0.15
5	2	0.08
Totale	26	1.00

Pagina 31

La Distribuzione Statistica di ... Frequenze Percentuali

- ✗ Consente di confrontare collettivi di numerosità diversa ed evidenziarne la composizione
- ✗ Frequenze Percentuali
 - si ottengono dividendo, per ogni modalità, la Freq. Ass. per la Num. Totale e moltiplicando per 100

- In sintesi:

$$p_i = \frac{n_i}{n} \times 100$$

- oppure: $p_i = f_i \times 100$

- $p_1 + p_2 + \dots + p_i + \dots + p_K = 100$

Pagina 32

Distribuzione dei laureati per genere

- ✗ Distribuzione per genere dei laureati di Scienze della Comunicazione nell'a.a. 2003/2004

Genere	Frequenze assolute	Frequenze relative	Frequenze percentuali
F	17	0.65	65
M	9	0.35	35
Totale	26	1.00	100

Pagina 33

In sintesi

Carattere X	Frequenza Assoluta n_i	Frequenza Relativa $f_i = n_i/n$	Frequenza Percentuale $p_i = f_i \times 100$
x_1	n_1	f_1	p_1
x_2	n_2	f_2	p_2
\vdots	\vdots	\vdots	\vdots
x_i	n_i	f_i	p_i
\vdots	\vdots	\vdots	\vdots
x_K	n_K	f_K	p_K
Totale	n	1	100

Pagina 34