Indice

1	Stati	stiche Univariate
	1.1	Importazione di un file .data
	1.2	Medie e variabilità
	1.3	Distribuzioni di frequenze
	1.4	Rappresentazioni Grafiche
	1.5	Analisi dati
	1.6	Esercizi
2	Stati	stiche Bivariate 18
	2.1	Introduzione
	2.2	Tabella pivot 18
	2.3	Medie e variabilità
	2.4	Grafici e tabelle a doppia entrata

Capitolo 2

Statistiche Bivariate

2.1 Introduzione

Nella formazione dei dati statistici si osserva che per ogni unità della popolazione si possono rilevare contemporaneamente uno o più caratteri ottenendo una distribuzione doppia o bivariata. Consideriamo il caso in cui per ogni unità vengano rilevati due caratteri. Si indichino con X e Y i due caratteri in esame; consideriamo dapprima il caso in cui siano entrambi discreti, con un numero di modalità finito, rispettivamente uguale ad h e k. Se le N unità della popolazione non sono molto numerose la distribuzione bivariata si può rappresentare indicando per ciascuna unità le coppie di modalità relative ai due caratteri. Per N elevato, solitamente, i dati vengono rappresentati in tabelle a doppia entrata.

La tabella a doppia entrata fornisce la distribuzione delle frequenze delle unità che presentano congiuntamente la modalità *i-esima* (i = 1, ..., h) del primo carattere e la modalità *j-esima* (j = 1, ..., k) del secondo carattere.

In una tabella a doppia entrata è possibile considerare diverse distribuzioni di frequenza. Le distribuzioni *marginali* che sono due, e si trovano ai MARGINI della tavola rappresentate dalla riga e dalla colonna dei TOTALI; esse rappresentano le 2 distribuzioni semplici dei caratteri X e Y presi singolarmente. Le distribuzioni *condizionate* sono ottenute a partire dalle RIGHE e dalle COLONNE interne alla tabella; ad esempio la distribuzione di Y condizionata alla *i*-esima modalità x_i di X fornisce la distribuzione delle frequenze relative di Y una volta fissata la modalità x_i , (i = 1, ..., h). Analogamente per le distribuzioni condizionate di X dato y_j , (j = 1, ..., k).

2.2 Tabella pivot

In excel lo strumento utilizzato per la costruzione di tabelle a doppia entrata è la tabella pivot.

Un rapporto di tabella pivot è una tabella interattiva che consente di riassumere le informazioni della distribuzione doppia. Una delle caratteristiche più interessanti di quest'analisi interattiva dei dati si ha con la possibilità di visualizzare diversi riepiloghi dei dati filtrandoli e facendo ruotare righe e colonne, visualizzando così solo i dati d'interesse. Ad esempio in questo caso si punta a visualizzare le frequenze assolute di due variabili in una tabella a doppia entrata.

Un rapporto di tabella pivot contiene dei campi, ognuno dei quali corrisponde a una colonna nei dati di origine e riepiloga più righe di informazioni dei dati di origine. I campi di un rapporto di tabella pivot elencano gli elementi dei dati tra le righe o sotto le colonne. Le celle in cui le righe e le

colonne si intersecano mostrano le frequenze degli elementi nella parte superiore della colonna e sul lato sinistro della riga.

Si costruisca una prima tabella di frequenza utilizzando il file di dati *investimenti.txt*, reperibile sul sito internet <u>www.economia.unical.it/statistica/Laboratori/dati/investimenti.txt</u>. Questi dati sono stati raccolti per conoscere le preferenze negli investimenti di individui che svolgono diverse professioni. Il data set in questione consta di N = 330 osservazioni relative a p = 2 variabili discrete (nominali) concernenti rispettivamente la professione e l'area d'investimento preferita per ciascuna unità del colletivo. Le modalità della variabile professione sono:

- medico;
- avvocato;
- commercialista;
- altro;

mentre le modalità della variabile area d'investimento sono:

- fondi azionari;
- fondi obbligazionari;
- azioni;
- titoli di stato;

La procedura da seguire per la costruzione della tabella è la seguente:

• dal menù Dati scegliere **Rapporto tabella pivot e grafico pivot**. Si accede così alla finestra di dialogo "Creazione Guidata tabella pivot e grafico pivot ", indicare dove si trovano i dati da analizzare e il tipo di rapporto da creare. Un rapporto di tabella pivot può essere creato da un elenco Excel, da un database esterno, da più fogli di lavoro Excel o da un altro rapporto di tabella pivot. Le voci che si selezionano in questo caso sono rispettivamente "*elenco o database microsoft excel*" e "*tabella pivot*" (vedi Figura 2.1). Continuare cliccando sul pulsante avanti;

Creazione guidata Tabell	a pivot e grafico pivot - Passaggio 1 di 3	? X
	Indicare dove si trovano i dati da analizzare.	
	Indicare il tipo di rapporto da creare. <u>T</u> abella pivot G <u>r</u> afico pivot (con tabella pivot)	
	Annulla < Indietro Avanti >	Eine

Figura 2.1: Creazione guidata Tabella Pivot.

• nella finestra successiva si deve indicare l'intervallo dove si trovano i dati da analizzare, in questo caso saranno le colonne relative alle distribuzioni delle due variabili (vedi Figura 2.2);

Creazione d	uidata Tab	ella nivot e o	rafico nivot -	Passanni	n 2 di 3	? X
Indicare dov	ve si trovano i	dati da utilizza	re.			
I <u>n</u> tervallo:	1			3	<u>S</u> foglia	
		Annulla	< <u>I</u> ndietro	Avanti	>	Eine
				<u></u>		

Figura 2.2: Creazione guidata Tabella Pivot.

• nell'ultima finestra si sceglie dove collocare la tebella pivot, in un nuovo foglio di lavoro o nello stesso dei dati di origine (vedi figura 2.3);



Figura 2.3: Creazione guidata Tabella Pivot.

infine, prima di ultimare la costruzione della tabella è conveniente impostare da quest'ultima finestra di dialogo le opzioni e il *layout* della tabella pivot. Per quanto riguarda le opzioni, in questo caso, modificare solo le voci relative ai valori da mostrare in caso di errore e in caso di celle vuote, nel primo caso prevedere il valore N.C. e nel secondo caso il valore 0 (vedi Figura 2.4).

9 200		1.15.16	a		1.000
D <mark>pzioni</mark> ta	abella pivot				?
Nome:	Tabella_pivot1	_			
Opzioni di	formattazione			21	
Total	i complessivi per colonna	ę	Layout pagina:	Da sotto a s	opra 🔄
✓ Total ✓ Tabe	i complessivi per rig <u>h</u> e Ila <u>f</u> ormattata automatic	amente	<u>⊂</u> ampi per colonr	na:	0 =
C Subto	otale elementi <u>p</u> agina na ;i etichette	iscosti	🔽 Per i valori di	<u>e</u> rrore, mostra:	N.C.
Mant	ieni formattazione		Per le celle <u>v</u> u	uote, mostra:	0
Etiche	ette elementi su <u>og</u> ni pa	gina stampata	📕 Titolį di stamp)a	
Conti Onzioni da	rassegna totali con *				
Opzioni o	origine dati:		Opzioni dati este	rni:	
V Salva	i dati con il layout della	tabella	🔽 Salva passwo	bit	
Attiv.	a mostra <u>d</u> ettaglio		C Query in seco	ondo piano	
	orna all'apertura		Cttimizza mer	moria	
C Aggie	oma ogni 🚺 🚖	minuti			
				ОК	Annulla

Figura 2.4: Opzioni tabella pivot

Per quanto riguarda le impostazioni del *layout* seguire la seguente procedura (vedi Figura 2.5):

Creazione guidata Tabel	la pivot e grafico	pivot - Layout	? ×
internal de la constantina de		Impostare la tabella pivot tr. pulsanti dei campi sulla destr varie aree del diagramma su	ascinando i 'a nelle Illa sinistra,
PAGINA	Area di i Profession Con <u>R</u> IGA	n <u>C</u> OLONNA nteggio di Area di DATI	Profession Area di ini
	· · ·	2	Annulla

Figura 2.5: Layout tabella pivot

- dal gruppo di pulsanti campo sulla destra, trascinare i campi desiderati sulle aree RIGA e COLONNA del diagramma; prevedere professione su riga e area di investimento su colonna;
- 2. trascinare i campi contenenti i dati che si desidera riepilogare nell'area DATI, in questo caso si può scegliere indistintamente uno dei due campi, per impostare il tipo di riepilogo da calcolare posizionare il cursore sul campo inserito nell'area dati e cliccare due volte si aprirà la finestra di dialogo"" quindi selezionare la voce conta (vedi figura 2.6);



Figura 2.6: Campo Tabella Pivot

Dopo aver cliccato su OK nella finestra di dialogo relativa al layout confermare tutte le operazioni con il pulsante *fine*.

Completate le operazioni la tabella si presenterà come in figura (2.7).

CAPITOLO 2. STATISTICHE BIVARIATE

Conteggio di Area di investimento	Area di investimento 👻	e e e e e e e e e e e e e e e e e e e			
Professione 👻	Fondi azionari	Fondi Obbligazionari	Azioni	Titoli di stato	Totale complessivo
Medico	30	25	15	0	70
Awocato	29	34	12	6	81
Commercialista	50	35	29	15	129
Altro	21	14	10	5	50
Totale complessivo	130	108	66	26	330

Figura 2.7: Distribuzione di frequenza

Nella tabella ottenuta si osservano i valori delle distribuzioni di frequenze marginali e condizionate. Con le tabelle pivot è possibile visualizzare le frequenze relative solo ad alcune delle modalità delle variabili in questione, così, diventa immediata la costruzione di tabelle ralative alle singole distribuzioni marginali e condizionate; inoltre in questo modo è anche possibile visualizzare le distribuzioni di frequenza relative ad un sottoinsieme dell'intero collettivo con determinate caratteristiche. La procedura per la selezione delle modalità è mostrata in figura 2.8.

Conteggio di Area di investimento	Area di investimento 👻				
Professione 🗸 🗸	Fondi azionari				
V Medico	30				
🖌 🖌 Avvocato	29				
Commercialista	50				
Altro	2				
	130				
	1				
OK Annulla					

Figura 2.8: Selezione modalità.

In figura 2.9 è visualizzata la tabella di frequenza relativa solo ad alcune delle variabili in questione.

Conteggio di Area di investimento	Area di investimento 👻	
Professione 🗸	Fondi azionari	Totale complessivo
Medico	30	30
Awocato	29	29
Totale complessivo	59	59

Figura 2.9: Distribuzione

Utilizzando il data set *esami.txt* reperibile sul sito internet <u>www.economia.unical.it/STATistica/Laboratori/dati/wine6.dat</u> si costruisca una seconda tabella pivot. Questo set di dati è relativo alla distribuzione di studenti secondo il voto riportato in Statistica e in Analisi Matematica e consta di 40 osservazioni; ovviamente le variabili in questione sono discrete (quantitative) e le modalità sono individuate da tutti i possibili voti.

Conteggio	Analisi 👻													
Statistica 👻	18	19	20	21	22	23	24	25	26	27	28	29	30	Totale
18	0	0	1	0	0	0	0	0	0	0	0	0	0	1
19	1	0	0	0	0	0	0	0	0	0	0	0	0	1
20	1	0	0	1	0	1	0	0	0	0	0	0	0	3
21	0	0	1	0	0	2	0	0	0	0	0	0	0	3
22	0	1	0	0	0	1	1	1	0	0	0	0	0	4
23	0	1	0	2	0	0	0	0	0	0	0	0	0	3
24	0	0	0	0	0	0	1	3	0	0	1	0	0	5
25	0	0	Ö	0	2	0	1	0	2	0	0	0	0	5
26	0	0	0	0	1	0	0	0	0	2	0	0	0	3
27	0	0	0	0	0	0	1	1	1	0	0	1	0	4
28	0	Û	0	0	0	1	Û	1	0	1	0	0	0	3
29	0	0	0	0	0	0	0	0	1	0	1	0	0	2
30	0	0	0	0	0	0	0	0	0	1	1	0	1	3
Totale	2	2	2	3	3	5	4	6	4	4	3	1	1	40

Ultimati i passaggi per la costruzione di una tabella a doppia entrata la tabella si presenterà come in figura 2.10.

Figura 2.10: Distribuzione

Consideriamo ora un altro insieme di dati, contenente variabili continue. A questo punto si costruisca una tabella a doppia entrata che coinvolga le variabili colore ed intensità colore relative al data set *wine recognition*; si coinvolgano tutte le unità che costituiscono il collettivo e si ricordi che, in questo caso, le variabili sono continue sarà quindi necessario costruire delle classi di frequenza per poter calcolare le distribuzioni di frequenza.

Rieseguendo la procedura sopra descritta e prevedendo nel layout la variabile colore su riga e la variabile intensità colore su colonna la tabella ottenuta si presenterà come in figura 2.11.

CAPITOLO 2. STATISTICHE BIVARIATE

Conteggio di Colore	Colore -							
Intensità colore 👻	0.90	0.91	0.91	0.92	0.93	0.94	0.95	0.96
1.28	0	0	0	0	1	0	0	0
1.74	0	0	0	0	0	0	0	0
1.90	0	0	0	0	0	0	0	0
1.95	i 0	0	0	0	0	0	1	0
2.00	1	0	0	0	0	0	0	0
2.06	i 0	0	0	0	0	1	0	0
2.08	0	0	0	1	0	0	0	0
2.15	i O	0	0	0	0	0	0	0
2.20	0	0	0	0	0	0	0	0
2.30	0	0	0	0	0	0	0	0
2.40	0	0	0	0	0	0	0	0
2.45	i 0	0	0	0	0	0	0	0
2.50	0	0	0	0	0	0	1	0
2.57	0	0	0	0	0	0	0	0
2.60	0	0	0	0	0	0	0	0
2.62	2 0	0	0	0	0	0	0	0
2.65	i 0	0	0	0	0	0	0	2
2.80	0	0	0	0	0	0	0	0
2.85	i 0	0	0	0	0	0	0	0
2.90	0	0	0	0	1	.0	0	0

Figura 2.11: Distribuzione di frequenza.

Evidentemente la tabella in figura 2.11 non è quella definitiva, è infatti necessario raggruppare i dati per classi, la procedura da seguire è la seguente:

• posizionare il cursore sulla cella intensità colore o colore cliccare sul pulsanre destro del mouse e dal menù selezionare raggruppa e struttura quindi raggruppa (vedi figura 2.12);



Figura 2.12: Creazione classi di frequenza

• a questo punto si aprirà una finestra di dialogo in cui si deve specificare l'ampiezza della classe

che si pone pari a 2.35 per la variabile intensità colore e 0.25 per la variabile colore (vedi figura 2.13).

Raggruppa	? ×
Raggruppamento auto	matico
Fino a:	7.8
Raggruppa <u>p</u> er:	1
	K Annulla

Figura 2.13: Creazione classi di frequenza

Completate le operazioni la tabella pivot si presenterà come in figura 2.14.

Conteggio di Colore	Colore 👻				la a	
Intensità colore 👻	0.48-0.73	0.73-0.98	0.98-1.23	1.23-1.48	1.48-1.73	Totale complessivo
1.28-3.62	1	24	19	13	ୀ	58
3.62-5.96	7	23	31	9	0	70
5.96-8.3	9	8	12	1	0	30
8.3-10.64	13	1	1	1	0	16
10.64-12.98	2	1	0	0	0	3
12.98-15.32	1	0	0	0	0	1
Totale complessivo	33	57	63	24	া	178

Figura 2.14: Tabella a doppia entrata

Partendo da una tabella pivot e sfruttando gli strumenti presenti in *excel* è possibile creare tabelle che riportino altri tipi di valori, come:

- frequenze relative (si ottengono dividendo le frequenze assolute per *N*);
- frequenze percentuali condizionate (si ottengono dividendo le distribuzioni parziali per la corrispettiva frequenza marginale).

Un esempio di questo tipo di tabelle relative alle variabili colore ed intensità colore è riportato in figura 2.15.

CAPITOLO 2. STATISTICHE BIVARIATE

Frequenze assolute					Fragmanza residiva								
1 4	0.48- 0.73	0.73-0.98	0.98- 1.23	1.23- 1.48	1.48- 1.73	Totale	i	0.48-0.73	0.73- 0.98	0.98- 1.23	1.23- 1.48	1.48- 1.73	Totale
1.28-3.63	1	24	19	13	1	58	1.28-3.63	0.01	0.13	0.11	0.07	0.01	0.33
3.63-5.98	7	23	31	9	0	70	3.63-5.98	0.04	0.13	0.17	0.05	0.00	0.39
5.98-8.33	9	8	12	1	0	30	5.98-8.33	0.05	0.04	0.07	0.01	0.00	0.17
8.33-10.68	13	1	1	1	0	16	8.33-10.68	0.07	0.01	0.01	0.01	0.00	0.09
10.68.13.03	3	1	0	0	0	4	10.68-13.03	0.02	0.01	0.00	0.00	0.00	0.02
10100-10100									0.00	0.00	0.40	0.04	4.00
Totale	33	57	63	24	1	178	Totale	0.19	0.32	0.35	0.13	0.01	1.00
Frequence	33	57 M condum	63 mate dal c tafore	24 color a risp	alto affa	178	Freques	0.19 Te percentina	0.32 Ø conditio	0.35 mate deá tiore	U.13	celare ris	1.00
Frequence,	33 0.48- 0.73	57 0.73- 0.98	63 wate def c afore 0.98- 1.23	24 olove risp 1.23- 1.48	1 alle afr 1.48- 1.73	178 Intensitio Totale	Frequent	0.19 •• #••Certhe 0.48-0.73	0.32 Conduit C	0.35 wale dea tore 0.98- 1.23	0.13 Withersoftle 1.23- 1.48	0.01 colore ris 1.48- 1.73	n.oo getto Tatale
Frequence	33 0.48- 0.73 1.72	57 0.73- 0.98 41.38	63 while def (24 color a <i>rise</i> 1.23- 1.48 22.41	1 alle alle 1.48- 1.73 1.72	178	Frequence i C	0.19 0.48-0.73 3.03	0.32 0.73- 0.98 42.11	0.35 male dea tive 0.98- 1.23 30.16	0.13	1.48- 1.73	1.00
Frequenze 1.28-3.63 3.63-5.98	33 0.48- 0.73 1.72 10.00	57 0.73- 0.98 41.38 32.86	63 0.98- 1.23 32.76 44.29	24 1.23- 1.48 22.41 12.86	1.48- 1.73 1.72 0.00	178	Frequent i C 1.28-3.63 3.63-5.98	0.19 0.48-0.73 3.03 21.21	0.32 0.73- 0.98 42.11 40.35	0.35 0.98- 1.23 30.16 49.21	1.23- 1.48 54.17 37.50	0.01 Color o ris 1.48- 1.73 100.00 0.00	1.00 perfo Tetale 32.58 39.33
Frequenze 1.28-3.63 3.63-5.98 5.98-8.33	33 0.48- 0.73 1.72 10.00 30.00	57 0.73- 0.98 41.38 32.86 26.67	63 0.98- 1.23 32.76 44.29 40.00	24 1.23 1.48 22.41 12.86 3.33	1.48- 1.73 1.72 0.00 0.00	178 Tetain 100.00 100.00 100.00	Frequence i C 1.28-3.63 3.63-5.98 5.98-8.33	0.19 0.48-0.73 3.03 21.21 27.27	0.32 0.73- 0.98 42.11 40.35 14.04	0.35 0.98- 1.23 30.16 49.21 19.05	0.13 1.23- 1.48 54.17 37.50 4.17	0.01 color o 75 1.48- 1.73 100.00 0.00 0.00	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
Frequence 1.28-3.63 3.63-5.98 5.98-8.33 8.33-10.68	33 0.48- 0.73 1.72 10.00 30.00 81.25	57 0.73- 0.98 41.38 32.86 26.67 6.25	63 0.98- 1.23 32.76 44.29 40.00 6.25	24 1.23- 1.48 22.41 12.86 3.33 6.25	1 1.48- 1.73 1.72 0.00 0.00 0.00 0.00	178 100.00 100.00 100.00 100.00 100.00	Frequence i E 1.28-3.63 3.63-5.98 5.98-8.33 8.33-10.68	0.19 0.48-0.73 3.03 21.21 27.27 39.39	0.32 0.73- 0.98 42.11 40.35 14.04 1.75	0.35 0.98- 1.23 30.16 49.21 19.05 1.59	1.23- 1.48 54.17 37.50 4.17 4.17	0.01 1.48- 1.73 100.00 0.00 0.00 0.00	1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00
Frequences i C 1.28-3.63 3.63-5.98 5.98-8.33 8.33-10.68 10.68-13.03	33 0.48- 0.73 1.72 10.00 30.00 81.25 75.00	57 0.73- 0.98 41.38 32.86 26.67 6.25 25.00	63 	24 1.23- 1.48 22.41 12.86 3.33 6.25 0.00	1 1.48 1.73 1.72 0.00 0.00 0.00 0.00 0.00	178 100.00 100.00 100.00 100.00 100.00 100.00	i C 1.28-3.63 3.63-5.98 5.98-8.33 8.33-10.68 10.68-13.03	0.19 0.48-0.73 3.03 21.21 27.27 39.39 9.09	0.32 0.73- 0.98 42.11 40.35 14.04 1.75 1.75	0.35 0.98- 1.23 30.16 49.21 19.05 1.59 0.00	1.23 1.48 54.17 37.50 4.17 4.17 0.00	0.01 1.48- 1.73 100.00 0.00 0.00 0.00 0.00 0.00	1.00 Tetair 32.58 39.33 16.85 8.99 2.25

Figura 2.15: Distribuzioni

2.3 Medie e variabilità

La procedura per il calcolo delle funzioni statistiche di distribuzioni condizionate e marginali nelle tabelle a doppia entrata, grazie alle proprità delle tabelle pivot, è abbastanza semplice. Infatti, posizionando il cursore sul campo incluso nell'area dati e cliccando sul pulsante destro del mouse e possibile selezionare dal menù la voce "Impostazioni campo", sarà visualizzata la finestra "Campo tabella pivot" e da essa si selezionerà la voce d'interesse (vedi figura 2.6).

Considerando la distribuzione doppia relativa alle variabili colore ed intensità colore, generare due copie della tabella pivot e prevedere in una come campo dati la variabile intensità colore e nell'altra la variabile colore; calcolare quindi per ognuna di esse le seguenti statistiche:

- media;
- dev.st;
- dev.st.pop;
- var.;
- var.pop.;
- min;
- max;

Il risultato di tale operazioni deve produrre due tabelle come quelle riportate in figura 2.16.

	1	1000	and a factor of the		Contractor and	0.0000000000000000000000000000000000000	
	Stat	istiche ma	erginali e	condizion	ate per in	tensità colore	
Media di Colore	Colore 👻						
Intensità colore 🛛 👻	0.48-0.73	0.73-0.98	0.98-1.23	1.23-1.48	1.48-1.73	Totale complessivo	
1.28-3.63	0.70	0.88	1.08	1.32	1.71	1.05	
3.63-5.98	0.65	0.85	1.09	1.24	0.00	0.99	
5.98-8.33	0.62	0.87	1.08	1.28	0.00	0.89	
8.33-10.68	0.61	0.74	1.12	1.23	0.00	0.69	
10.68-13.03	0.54	0.85	0.00	0.00	0.00	0.62	
Totale complessivo	0.62	0.86	1.08	1.29	1.71	0.96	
	Stat	istiche ma	siginali e	condizion	ate per ini	tensità colore	
Media di Intensità colore	Colore 👻						
Intensità colore 🛛 👻	0.48-0.73	0.73-0.98	0.98-1.23	1.23-1.48	1.48-1.73	Totale complessivo	
1.28-3.63	3.40	2.61	2.84	2.82	1.90	2.73	
3.63-5.98	4.99	4.97	4.73	4.65	0.00	4.82	
5.98-8.33	7.62	6.62	6.68	7.30	0.00	6.97	
8.33-10.68	9.43	8.66	8.90	8.70	0.00	9.30	
10.68-13.03	11.85	10.68	0.00	0.00	0.00	11.56	
Totale complessivo	8.03	4.37	4.60	3.94	1.90	5.06	

Figura 2.16:	Statistiche
--------------	-------------

2.4 Grafici e tabelle a doppia entrata

Come si è visto nel precedente paragrafo le distribuzioni statistiche doppie possono coinvolgere coppie di caratteri quantitativi o coppie di caratteri qualitativi (ovviamente si può presentare anche il caso di un carattere qualitativo ed uno quantitativo); per queste distribuzioni esistono diversi tipi di rappresentazioni grafiche. In particolare nel caso di due caratteri quantitativi, continui e discreti, uno dei grafici più utilizzati è *nuvola dei punti*. La nuvola dei punti è un diagramma cartesiano in cui sull'asse delle ascisse e su quello delle ordinate sono riportati, avendo scelto unità di misura per ciascun asse, i valori assunti dalle due variabili in questione. La nuvola dei punti consente di visualizzare la dispersione tra le unità statistiche ossia la loro vicinanza o distanza indicanti rispettivamente la loro somiglianza o dissomiglianza rispetto ai due caratteri contemporaneamente considerati. Importante è inoltre la forma assunta dalla nuvola dei punti in quanto fornisce indicazioni sul tipo di relazione esistente tra le due variabili: ad esempio, se i punti assumono una forma sferica le due variabili sono scarsamente correlate o indipendenti ossia il variare di una non influisce sul variare dell'altra (la costruzione di questo grafico sarà trattata nel paragrafo successivo relativamente alle variabili colore ed intensità colore).

Nel caso di due caratteri qualitativi la rapresentazione grafica più adatta alla distribuzione doppia di contingenza è il *grafico a barre verticali a tre dimensioni o ortogramma a colonne* in esso sono rappresentati i due caratteri e le frequenze di associazione delle modalità dei due caratteri.

Inoltre in quest'ultimo caso un'utile rappresentazione per le distribuzioni parziali è data dal *diagramma circolare o aereogramma circolare* detto anche "*grafico a torta*"; fissando la modalità di una varianile si possono osservare le variazioni percentuali nelle modalità dell'altra variabile. Ad esempio, nel caso della distribuzione doppia relativa agli investimenti, si può osservare la ripartizione degli investimenti in una categoria professionale.

Questi ultimi due grafici possono essere costruiti con una rapida procedura partendo dalla tabella

pivot. Come si è visto nel paragrafo precedente, durante la costruzione guidata di una tabella pivot, nella prima finestra di dialogo si deve indicare il tipo di rapporto da creare (vedi Figura 2.1), se si sceglie la voce "*Grafico pivot (con tabella)*" in modo automatico insieme alla tabella pivot viene generato, in un nuovo foglio elettronico, il grafico ad essa associato. Invece nel caso in cui è stata selezionata la voce tabella pivot, una volta ultimata la procedura di costruzione della tabella, per la costruzione del grafico basta posizionare il cursore sulla tabella cliccare con il pulsante destro del mouse e scegliere la voce "*Grafico pivot*" (vedi Figura 2.17).



Figura 2.17: Grafico pivot

Per le impostazioni relative al tipo di grafico si procede come illustrato nel capitolo precedente, l'unica differenza è che in un rapporto di grafico pivot, la formattazione e altre modifiche apportate ad una serie di dati non vengono mantenute quando si modifica la visualizzazione del grafico o del rapporto della tabella pivot associato.

Nelle figure 2.18 e 2.19 sono riportate due rapprasentazioni grafiche relative alla distribuzione di frequenza ottenuta dal set di dati investimenti.txt.



Figura 2.18: *Diagramma a barre tridimensionale*.



Figura 2.19: Grafico a torta.