

# CORSO DI STATISTICA PARTE 1

## Esercitazione N°1

(21 gennaio 2011)

Dr. Mirko Bevilacqua

### Dataset "test d'ingresso"

	SEDE DISTACCATA	GENERE	VOTO DIPLOMA	TIPO SCUOLA	PROVINCIA SEDE SCUOLA	CITTA SCUOLA	PUNTEGGIO TEST
1	SEDE DI CASSINO	F	60	TC	FR	CASSINO	5
2	SEDE DI TERRACINA	M	60	LC	LT	FONDI	14,75
3	SEDE DI CASSINO	F	69	LS	FR	CECCANO	15
4	SEDE DI TERRACINA	M	100	GE	LT	FORMIA	15
5	SEDE DI CASSINO	M	73	TC	FR	CASSINO	17,25
6	SEDE DI CASSINO	M	70	LS	FR	VEROLI	17,5
7	SEDE DI CASSINO	F	72	MP	FR	CASSINO	8,75
8	SEDE DI CASSINO	M	64	LS	FR	PONTECORVO	17,25
9	SEDE DI TERRACINA	M	72	LS	FR	CECCANO	18,5
10	SEDE DI TERRACINA	M	70	TC	FR	CECCANO	9,25
11	SEDE DI CASSINO	M	63	LS	FR	SORA	15,75
12	SEDE DI CASSINO	F	76	MP	FR	CASSINO	7,25
13	SEDE DI CASSINO	F	80	TC	FR	CASSINO	8,5
14	SEDE DI TERRACINA	M	100	TC	LT	FORMIA	26,15
15	SEDE DI CASSINO	M	73	TC	LT	FORMIA	22,3
16	SEDE DI CASSINO	F	73	LS	LT	FORMIA	28,8

### ESERCIZIO N°1

Per ognuna delle variabili che compongono il dataset, si definisca il tipo di variabile (quantitativo/qualitativo) elencando l'insieme delle possibili modalità/intensità.

CARATTERE	TIPO	MODALITA'
<i>Sede distaccata</i>	Qualitativa sconnessa	{"sede Cassino", "sede Terracina"}
<i>Genere</i>	Qualitativa dicotomica	{"M", "F"}
<i>Voto diploma</i>	Quantitativa discreta	{60, 63, 64, ....., 100}
<i>Tipo scuola</i>	Qualitativa sconnessa	{"TC", "LC", "LS", "GE", "MP"}
<i>Provincia sede scuola</i>	Qualitativa sconnessa	{"FR", "LT"}
<i>Città Scuola</i>	Qualitativa sconnessa	{"Cassino", "Formia", "Fondi", "Sora", "Ceccano", "Pontecorvo", "Veroli"}
<i>Punteggio test</i>	Quantitativa continua	{60, 63, 64, ....., 100}

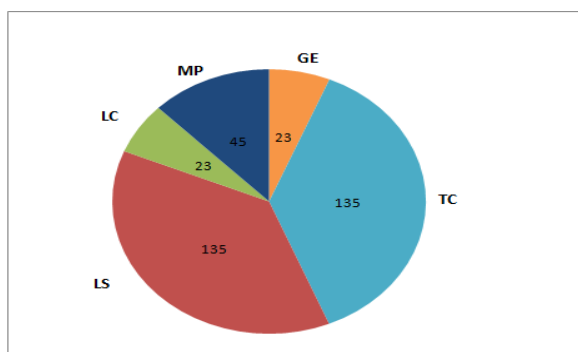
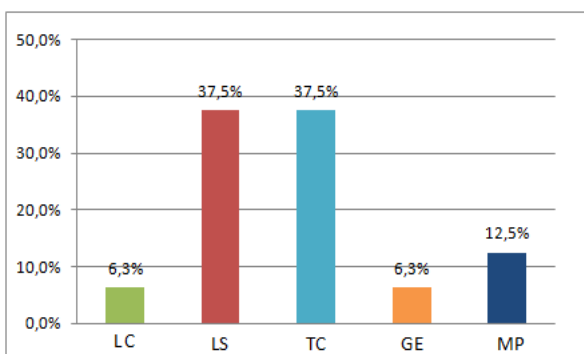
## ESERCIZIO N°2

Si costruiscano le distribuzioni di frequenza semplici per le variabili "Tipo scuola" e "Genere", riportando le frequenze assolute e le frequenze relative. Rappresentare graficamente (grafico a barre e grafico a torta) le distribuzioni ottenute. (PS: Per riprodurre il grafico a torta è necessario calcolare l'ampiezza dell'angolo al centro associato ad ogni modalità).

### a) "Tipo scuola"

<u>Tipo Scuola</u>	$n_i$	$f_i$	$P_i$	$F_i$
LC	1	0,063	6,3%	0,063
LS	6	0,375	37,5%	0,438
TC	6	0,375	37,5%	0,813
GE	1	0,063	6,3%	0,875
MP	2	0,125	12,5%	1
	16	1	100%	

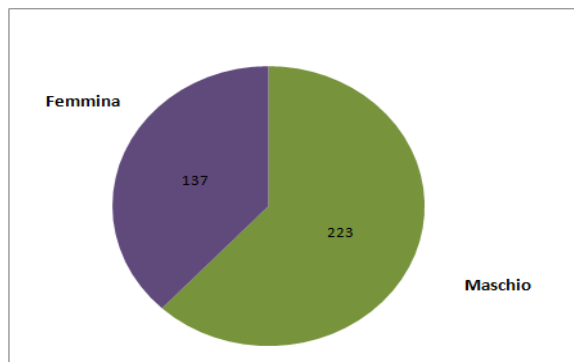
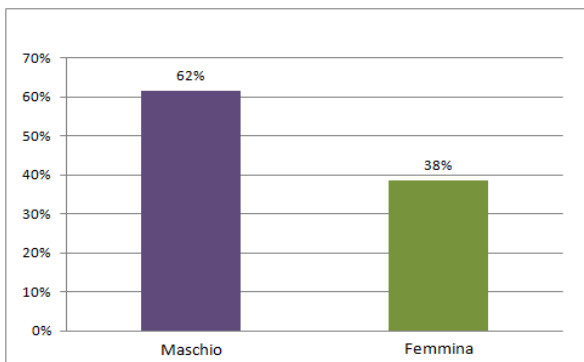
Istogramma di frequenze e grafico a torta:



### b) "Sesso"

<u>Sesso</u>	$n_i$	$f_i$	$P_i$	$F_i$
M	10	0,625	62,5%	0,625
F	6	0,375	37,5%	1
	16	1	100%	

Istogramma di frequenze e grafico a torta:



### ESERCIZIO N° 3

- Costruire la distribuzione di intensità per la "Punteggio test".
- Suddividere la distribuzione in 4 classi equiampie riportando le frequenze (assolute e relative) e la funzione di ripartizione empirica.
- Suddividere la distribuzione in 4 classi equifrequenti riportando le frequenze (assolute e relative) e la funzione di ripartizione empirica.
- Calcolare la percentuale di "Punteggio test" inferiore a 17,5.

**a) Distribuzione ordinata del carattere "punteggio test"**

{5; 7,25; 8,5; 8,75; 9,25; 14,75; 15; 15; 15,75; 17,25; 17,25; 17,5; 18,5; 22,3; 26,15; 28,8}

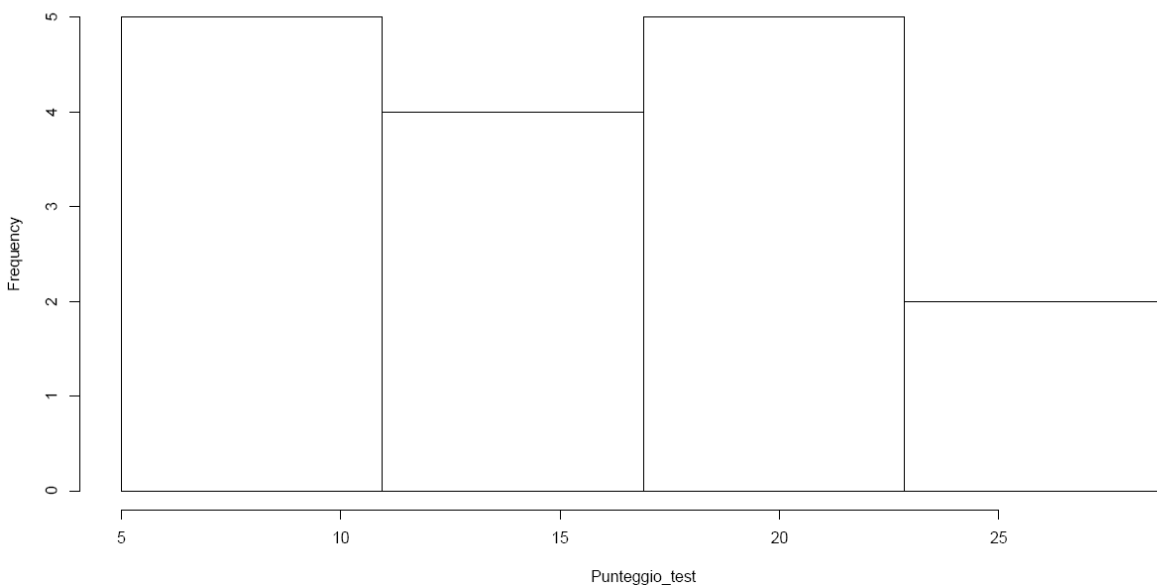
**b) "punteggio test" in 4 classi equiampie:**

Per trovare l'ampiezza delle classi è necessario dividere il campo di variazione della variabile (numeratore)

per 4:  $\frac{28,8 - 5}{4} = 5,95$

<u>Punteggio test</u>	$n_i$	$f_i$	$P_i$	$F_i$	$a_i$	$d_i$
[5,00 – 10,95]	5	0,313	31,3%	0,313	5,95	0,05
]10,95 - 16,90]	4	0,250	25,0%	0,563	5,95	0,04
]16,90 - 22,85]	5	0,313	31,3%	0,875	5,95	0,05
]22,85 - 28,80]	2	0,125	12,5%	1	5,95	0,02
	16	1	100%			

Histogram of Punteggio\_test



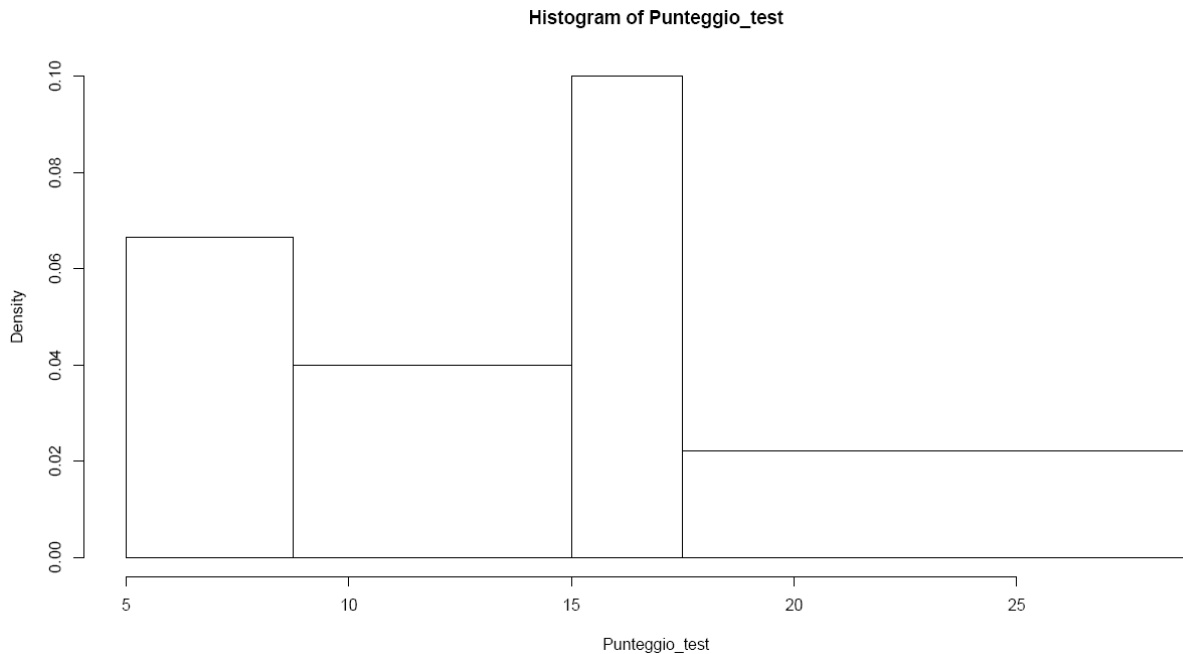
c) "punteggio test" in 4 classi equifrequenti

Per determinare la frequenza assoluta (costante) di ogni classe si divide per 4 l'intero collettivo statistico ( $16:4=4$ )

{5 - 7,25 - 8,5 - **8,75** - 9,25 - 14,75 - 15 - **15** - 15,75 - 17,25 - 17,25 - **17,5** - 18,5 - 22,3 - 26,15 - **28,8**}

(si vengono a creare classi con ampiezza diversa).

<u>Punteggio test</u>	$n_i$	$f_i$	$P_i$	$F_i$	$a_i$	$d_i$
[5,00 - 8,75]	4	0,25	25%	0,25	3,75	0,07
]8,75 - 15,00]	4	0,25	25%	0,50	6,25	0,04
]15,00 - 17,50]	4	0,25	25%	0,75	1,75	0,14
]17,50 - 28,80]	4	0,25	25%	1	11,30	0,02
	16	1	100%			



d) La percentuale di "punteggio test" inferiore a 17,5 è del 75%.