

# Statistica per esempi

## La statistica spiegata attraverso esempi concreti

I primi 6 capitoli del volume scritto nel 1988 da Enzo Lombardo e Alberto Zuliani, docenti di statistica presso La Sapienza di Roma, sono stati rivisitati dall'Istat e dai due autori e rilasciati sul sito dell'Istat (il libro è fuori commercio).

Le tavole sono tratte dall'Annuario statistico italiano 2002. Per agevolarne la lettura, alcune tavole non sono riportate integralmente, data la loro complessità, ma soltanto per le parti trattate nel testo. In questi casi, chiaramente indicati, un link permette di visualizzare la tavola completa.

### Indice

1. Un gioco: ritrovarsi nelle statistiche ufficiali
2. Una lettura statistica dei giornali
3. Uomini, donne, sport
4. Quanti eravamo, quanti siamo, quanti saremo
5. Il campionato di calcio italiano di serie A
6. La pesca sportiva in acque dolci

# 1. Un gioco: ritrovarsi nelle statistiche ufficiali

## Che cosa si impara nel capitolo 1

Il primo capitolo ha per protagonista Paolo, un tredicenne che gioca a ritrovarsi nelle statistiche ufficiali. Rispettando il diritto alla riservatezza, la statistica ufficiale conta le persone e gli eventi che le riguardano; Paolo si ritrova insieme alla sua famiglia nelle tabelle statistiche che riportano i matrimoni (quello della sorella), i degenti in ospedale (l'operazione per l'appendicite), le vacanze (il soggiorno in montagna con la sua famiglia). Si impara, in generale, a cercare informazioni di proprio interesse in un volume di statistiche e nelle specifiche tavole.

Paolo ha 13 anni. È nato a Roma il 12 ottobre 1989. Vive in questa città con i suoi genitori e con un fratello più grande, Andrea. Nel 2002 ha iniziato a frequentare la terza media nella stessa scuola pubblica dell'anno scorso. Paolo potrebbe essere uno di voi. Gli abbiamo proposto un gioco: ritrovarsi nelle statistiche ufficiali, quelle che l'Istat (Istituto nazionale di statistica) produce e pubblica ogni anno.

Sfogliando insieme a lui l'*Annuario statistico italiano 2002*, abbiamo trovato la tavola 26.1.

**Tavola 26.1 - Popolazione residente per sesso, densità per Km<sup>2</sup>, popolazione presente, famiglie e componenti, numero medio di componenti per famiglia e componenti permanenti delle convivenze, per regione - Censimento 2001**

REGIONI	Popolazione residente			Densità per Km <sup>2</sup>	Popolazione presente	Famiglie		Numero medio di componenti per famiglia	Componenti permanenti delle convivenze
	M	F	MF			Numero	Componenti		
Piemonte	2.011.046	2.155.396	4.166.442	164,0	4.184.901	1.779.054	4.124.575	2,3	41.867
Valle d'Aosta	58.632	60.724	119.356	36,6	120.173	53.237	118.193	2,2	1.163
Lombardia	4.323.073	4.599.390	8.922.463	373,9	8.967.864	3.595.478	8.856.398	2,5	66.065
Trentino-Alto Adige	458.568	478.539	937.107	68,9	959.424	366.158	926.272	2,5	10.835
Bolzano - Bozen	226.415	234.250	460.665	62,3	477.846	172.880	455.705	2,6	4.960
Trento	232.153	244.289	476.442	76,8	481.578	193.278	470.567	2,4	5.875
Veneto	2.188.023	2.302.563	4.490.586	244,2	4.540.026	1.699.235	4.443.890	2,6	46.696
Friuli-Venezia Giulia	567.535	612.840	1.180.375	150,2	1.179.188	496.327	1.165.816	2,3	14.559
Liguria	736.563	824.185	1.560.748	287,9	1.567.889	706.254	1.548.210	2,2	12.538
Emilia-Romagna	1.916.091	2.044.458	3.960.549	179,0	4.035.131	1.638.914	3.928.624	2,4	31.925
Toscana	1.665.794	1.795.041	3.460.835	150,5	3.447.067	1.377.834	3.437.707	2,5	23.128
Umbria	393.978	421.610	815.588	96,5	834.133	309.609	809.905	2,6	5.683
Marche	711.484	752.384	1.463.868	151,0	1.468.526	545.861	1.453.299	2,7	10.569
Lazio	2.380.453	2.595.731	4.976.184	289,2	4.843.576	1.937.353	4.939.701	2,5	36.483
Abruzzo	603.999	640.227	1.244.226	115,3	1.232.454	455.564	1.238.041	2,7	6.185
Molise	154.071	162.477	316.548	71,3	300.143	118.231	315.140	2,7	1.408
Campania	2.754.779	2.897.713	5.652.492	415,9	5.642.397	1.838.826	5.634.330	3,1	18.162
Puglia	1.933.064	2.050.423	3.983.487	205,7	3.918.430	1.362.198	3.967.947	2,9	15.540
Basilicata	293.001	302.726	595.727	59,6	568.967	212.918	593.317	2,8	2.410
Calabria	976.055	1.017.219	1.993.274	132,2	1.945.130	699.220	1.983.149	2,8	10.125
Sicilia	2.351.651	2.514.551	4.866.202	189,3	4.793.417	1.739.972	4.845.529	2,8	20.673
Sardegna	783.093	816.418	1.599.511	66,4	1.584.203	570.845	1.590.797	2,8	8.714
<b>ITALIA</b>	<b>27.260.953</b>	<b>29.044.615</b>	<b>56.305.568</b>	<b>186,9</b>	<b>56.133.039</b>	<b>21.503.088</b>	<b>55.920.840</b>	<b>2,6</b>	<b>384.728</b>
<b>Nord</b>	<b>12.259.531</b>	<b>13.078.095</b>	<b>25.337.626</b>	<b>211,3</b>	<b>25.554.596</b>	<b>10.334.657</b>	<b>25.111.978</b>	<b>2,4</b>	<b>225.648</b>
<b>Centro</b>	<b>5.151.709</b>	<b>5.564.766</b>	<b>10.716.475</b>	<b>183,7</b>	<b>10.593.302</b>	<b>4.170.657</b>	<b>10.640.612</b>	<b>2,6</b>	<b>75.863</b>
<b>Mezzogiorno</b>	<b>9.849.713</b>	<b>10.401.754</b>	<b>20.251.467</b>	<b>164,6</b>	<b>19.985.141</b>	<b>6.997.774</b>	<b>20.168.250</b>	<b>2,9</b>	<b>83.217</b>

Fonte: Istat, 14° Censimento generale della popolazione e delle abitazioni al 21 ottobre 2001, primi risultati

Ecco Paolo: è uno dei 27.260.953 maschi residenti e censiti al 24 ottobre 2001, in occasione del 14° Censimento generale della popolazione (il 14° dall'Unità d'Italia). Meglio ancora, è uno dei 2.380.453 maschi censiti nel Lazio.

Lo ritroviamo anche nella tavola 2.2 (Popolazione residente per regione e classe di età). Alla data del censimento, Paolo aveva da poco compiuto 12 anni. Perciò, eccolo lì: fra i 130.460 coetanei maschi di 10-14 anni che vivono nel Lazio.

La sua famiglia (Tavola 26.1) è una delle 21.503.088 famiglie italiane censite nel 2001. Poiché all'epoca viveva nel Lazio è una delle 1.937.353 famiglie abitanti in questa regione.

Paolo riflette

«Nonostante mia sorella Carla si sia sposata tre anni fa e sia andata a vivere insieme al marito in un'altra casa, la nostra, con quattro persone, è ancora una famiglia grande. Nel Lazio il numero medio di componenti in ogni famiglia è appena 2,5». Riflette ancora: «La professoressa di lettere ci ha detto che nelle regioni del Mezzogiorno si hanno più figli. È proprio vero; lì le famiglie hanno in media 2,9 componenti, rispetto ai 2,6 delle regioni centrali e ai 2,4 nel Nord».

Tavola 2.2 - Popolazione residente per classe di età, regione e sesso al 1° gennaio 2001 \*

REGIONI	Meno di 1	1-4	5-9	10-14	15-24	25-44	45-64	65 e più	Totale
MASCHI E FEMMINE									
Piemonte	35.773	138.626	170.383	170.503	415.032	1.305.472	1.165.264	888.678	4.289.731
Valle d'Aosta	1.153	4.280	5.010	5.015	11.554	38.911	31.873	22.793	120.589
Lombardia	85.209	326.068	394.189	389.439	934.198	2.949.302	2.423.780	1.619.529	9.121.714
Trentino-Alto Adige	10.540	41.796	50.263	48.108	103.606	304.050	226.698	158.062	943.123
<i>Bolzano-Bozen</i>	<i>5.452</i>	<i>22.000</i>	<i>26.490</i>	<i>25.410</i>	<i>53.971</i>	<i>152.714</i>	<i>107.249</i>	<i>71.978</i>	<i>465.264</i>
<i>Trento</i>	<i>5.088</i>	<i>19.796</i>	<i>23.773</i>	<i>22.698</i>	<i>49.635</i>	<i>151.336</i>	<i>119.449</i>	<i>86.084</i>	<i>477.859</i>
Veneto	43.059	166.298	200.361	197.521	475.496	1.477.235	1.165.214	815.669	4.540.853
Liguria	11.200	44.767	57.064	56.827	132.806	469.023	444.311	405.018	1.621.016
Emilia-Romagna	33.896	126.327	151.773	146.673	362.773	1.246.472	1.053.039	887.710	4.008.663
Toscana	28.283	108.785	136.067	139.386	339.595	1.067.056	945.536	782.896	3.547.604
Umbria	6.774	26.305	34.003	35.706	86.958	245.991	216.973	187.772	840.482
Marche	12.471	49.057	63.135	65.214	156.693	436.606	370.086	315.933	1.469.195
Lazio	48.362	194.716	251.247	253.945	580.743	1.681.729	1.371.327	920.233	5.302.302
Abruzzo	10.938	44.163	60.963	64.864	152.113	381.973	309.647	256.622	1.281.283
Molise	2.621	11.239	15.953	17.332	39.950	95.324	76.428	68.330	327.177
Campania	68.081	274.150	370.591	387.930	837.596	1.769.656	1.271.571	802.669	5.782.244
Puglia	42.505	171.317	232.698	249.346	579.750	1.235.319	945.184	630.489	4.086.608
Basilicata	5.648	22.744	32.222	35.902	81.007	181.754	135.818	109.712	604.807
Calabria	19.680	81.558	116.787	131.643	290.504	607.695	454.299	341.122	2.043.288
Sicilia	52.859	217.636	304.920	323.528	691.379	1.496.915	1.155.206	834.257	5.076.700
Sardegna	13.764	54.688	77.293	88.599	220.605	528.193	408.708	256.194	1.648.044
<b>ITALIA</b>	<b>542.368</b>	<b>2.140.683</b>	<b>2.769.342</b>	<b>2.851.511</b>	<b>6.601.630</b>	<b>17.882.708</b>	<b>14.499.840</b>	<b>10.555.935</b>	<b>57.844.017</b>
<b>NORD</b>	<b>230.382</b>	<b>884.325</b>	<b>1.073.463</b>	<b>1.058.116</b>	<b>2.544.737</b>	<b>8.154.497</b>	<b>6.839.057</b>	<b>5.049.706</b>	<b>25.834.283</b>
<b>CENTRO</b>	<b>95.890</b>	<b>378.863</b>	<b>484.452</b>	<b>494.251</b>	<b>1.163.989</b>	<b>3.431.382</b>	<b>2.903.922</b>	<b>2.206.834</b>	<b>11.159.583</b>
<b>MEZZOGIORNO</b>	<b>216.096</b>	<b>877.495</b>	<b>1.211.427</b>	<b>1.299.144</b>	<b>2.892.904</b>	<b>6.296.829</b>	<b>4.756.861</b>	<b>3.299.395</b>	<b>20.850.151</b>
MASCHI									
Piemonte	18.441	71.393	87.468	87.975	212.680	667.918	575.924	360.239	2.082.038
Valle d'Aosta	588	2.202	2.601	2.565	6.032	20.081	16.136	9.179	59.384
Lombardia	44.028	167.762	202.770	200.761	478.710	1.514.274	1.193.072	631.916	4.433.293
Trentino-Alto Adige	5.354	21.329	25.750	24.675	53.010	156.462	113.266	62.613	462.459
<i>Bolzano-Bozen</i>	<i>2.755</i>	<i>11.253</i>	<i>13.601</i>	<i>13.063</i>	<i>27.630</i>	<i>78.922</i>	<i>53.104</i>	<i>28.905</i>	<i>229.233</i>
<i>Trento</i>	<i>2.599</i>	<i>10.076</i>	<i>12.149</i>	<i>11.612</i>	<i>25.380</i>	<i>77.540</i>	<i>60.162</i>	<i>33.708</i>	<i>233.226</i>
Veneto	22.077	85.370	103.033	101.633	243.188	759.591	578.585	322.072	2.215.549
Friuli-Venezia Giulia	4.887	18.673	22.781	22.416	56.380	187.252	162.874	96.591	571.854
Liguria	5.812	23.027	29.374	29.397	67.594	238.197	214.587	160.506	768.494
Emilia-Romagna	17.510	64.935	78.233	75.458	185.618	639.142	516.726	365.883	1.943.505
Toscana	14.530	56.017	69.912	71.558	174.066	541.192	461.200	322.780	1.711.255
Umbria	3.417	13.792	17.569	18.314	44.734	124.322	106.514	78.987	407.649
Marche	6.518	25.352	32.588	33.515	80.061	222.268	181.868	132.784	714.954
Lazio	24.994	100.195	129.177	130.460	296.474	833.250	661.181	379.184	2.554.915
Abruzzo	5.738	22.591	31.383	33.387	77.967	191.727	153.144	108.137	624.074
Molise	1.340	5.845	8.191	8.921	20.429	48.076	37.934	28.913	159.649
Campania	34.974	140.916	189.985	198.606	426.172	876.421	623.141	329.045	2.819.260
Puglia	21.834	88.219	120.211	128.531	296.125	610.708	457.940	266.325	1.989.893
Basilicata	2.979	11.837	16.560	18.524	41.637	91.291	67.033	48.204	298.065
Calabria	10.123	41.817	60.072	67.682	148.255	303.196	226.189	145.931	1.003.265
Sicilia	27.228	111.460	156.421	166.042	352.542	739.724	558.099	353.038	2.464.554
Sardegna	7.099	28.442	39.881	45.561	113.614	265.968	201.453	108.730	810.748
<b>ITALIA</b>	<b>279.471</b>	<b>1.101.174</b>	<b>1.423.960</b>	<b>1.465.981</b>	<b>3.375.288</b>	<b>9.031.060</b>	<b>7.106.866</b>	<b>4.311.057</b>	<b>28.094.857</b>
<b>NORD</b>	<b>118.697</b>	<b>454.691</b>	<b>552.010</b>	<b>544.880</b>	<b>1.303.212</b>	<b>4.182.917</b>	<b>3.371.170</b>	<b>2.008.999</b>	<b>12.536.576</b>
<b>CENTRO</b>	<b>49.459</b>	<b>195.356</b>	<b>249.246</b>	<b>253.847</b>	<b>595.335</b>	<b>1.721.032</b>	<b>1.410.763</b>	<b>913.735</b>	<b>5.388.773</b>
<b>MEZZOGIORNO</b>	<b>111.315</b>	<b>451.127</b>	<b>622.704</b>	<b>667.254</b>	<b>1.476.741</b>	<b>3.127.111</b>	<b>2.324.933</b>	<b>1.388.323</b>	<b>10.169.508</b>

Fonte: Istat, Popolazione residente comunale per sesso, anno di nascita e stato civile

\* Per agevolare la lettura la tavola non è stata riportata integralmente. La tavola completa si può consultare cliccando: Tavola 2.2

La sorella di Paolo, Carla, quando si è sposata, nel 1998, aveva appena 21 anni e il marito era più grande di lei di due. Eccoli, tutti e due, nella tavola 2.4 (Matrimoni per classe di età della sposa e classe di età dello sposo): sono una delle 17.042 coppie, in cui marito e moglie avevano fra 20 e 24 anni, che si sono sposate nel 1998.

Paolo riflette

«Certo, mia sorella e mio cognato si sono sposati molto giovani! Sono soltanto 17.042 le coppie con la stessa combinazione di età dei coniugi rispetto alle 280.034 coppie che si sono sposate nel 1998, cioè meno di una ogni 16».

**Tavola 2.4 - Matrimoni per combinazione di stato civile degli sposi, classe di età della sposa, rito e classe di età dello sposo - Anno 1998 \***

CLASSI DI ETA' DELLA SPOSA	CLASSI DI ETA' DELLA SPOSA												Totale	
	16	17	18-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60 e più		
<b>TOTALE DEGLI SPOSI</b>														
<b>TOTALE</b>														
16	-	3	-	-	-	-	-	-	-	-	-	-	-	3
17	10	4	5	2	4	2	-	-	-	-	-	-	-	27
18-19	57	81	524	349	66	17	5	-	-	-	-	-	-	1.099
20-24	126	365	5.230	17.042	4.622	560	94	21	12	8	3	2	28.085	
25-29	24	71	2.884	41.985	57.446	8.348	893	147	43	28	12	12	111.893	
30-34	5	10	468	11.641	43.544	26.099	3.595	396	93	30	16	15	85.912	
35-39	-	4	106	1.783	8.408	11.556	5.119	974	190	45	10	14	28.209	
40-44	-	2	31	452	1.657	3.328	3.043	1.467	409	108	32	6	10.535	
45-49	-	-	11	161	491	1.089	1.371	1.191	670	229	61	15	5.289	
50-54	-	-	6	62	209	404	647	773	626	450	124	49	3.350	
55-59	-	-	4	27	85	129	249	429	481	418	235	104	2.161	
60 e più	-	-	3	46	65	107	163	346	497	632	617	995	3.471	
<b>Totale</b>	<b>222</b>	<b>540</b>	<b>9.272</b>	<b>73.550</b>	<b>116.597</b>	<b>51.639</b>	<b>15.179</b>	<b>5.744</b>	<b>3.021</b>	<b>1.948</b>	<b>1.110</b>	<b>1.212</b>	<b>280.034</b>	
<b>RITO RELIGIOSO</b>														
16	-	1	-	-	-	-	-	-	-	-	-	-	-	1
17	6	1	4	1	4	1	-	-	-	-	-	-	-	17
18-19	29	55	270	186	39	10	2	-	-	-	-	-	-	591
20-24	81	245	3.773	14.057	3.689	367	31	4	2	5	1	-	22.255	
25-29	17	49	2.166	37.437	51.800	6.706	523	45	4	15	4	7	98.773	
30-34	4	7	319	9.842	38.353	21.126	2.232	129	25	7	4	10	72.058	
35-39	-	2	57	1.146	6.216	7.927	2.777	347	36	9	2	1	18.520	
40-44	-	1	7	177	784	1.642	1.293	459	72	20	4	1	4.460	
45-49	-	-	-	38	136	309	407	322	124	35	11	-	1.382	
50-54	-	-	2	4	37	72	112	136	103	56	17	6	545	
55-59	-	-	1	4	13	18	34	77	102	65	35	17	366	
60 e più	-	-	2	9	16	13	11	31	85	108	117	221	613	
<b>Totale</b>	<b>137</b>	<b>361</b>	<b>6.601</b>	<b>62.901</b>	<b>101.087</b>	<b>38.191</b>	<b>7.422</b>	<b>1.550</b>	<b>553</b>	<b>320</b>	<b>195</b>	<b>263</b>	<b>219.581</b>	
<b>RITO CIVILE</b>														
16	-	2	-	-	-	-	-	-	-	-	-	-	-	2
17	4	3	1	1	-	1	-	-	-	-	-	-	-	10
18 - 19	28	26	254	163	27	7	3	-	-	-	-	-	-	508
20 - 24	45	120	1.457	2.985	933	193	63	17	10	3	2	2	5.830	
25 - 29	7	22	718	4.548	5.646	1.642	370	102	39	13	8	5	13.120	
30 - 34	1	3	149	1.799	5.191	4.973	1.363	267	68	23	12	5	13.854	
35 - 39	-	2	49	637	2.192	3.629	2.342	627	154	36	8	13	9.689	
40 - 44	-	1	24	275	873	1.686	1.750	1.008	337	88	28	5	6.075	
45 - 49	-	-	11	123	355	780	964	869	546	194	50	15	3.907	
50 - 54	-	-	4	58	172	332	535	637	523	394	107	43	2.805	
55 - 59	-	-	3	23	72	111	215	352	379	353	200	87	1.795	
60 e più	-	-	1	37	49	94	152	315	412	524	500	774	2.858	
<b>Totale</b>	<b>85</b>	<b>179</b>	<b>2.671</b>	<b>10.649</b>	<b>15.510</b>	<b>13.448</b>	<b>7.757</b>	<b>4.194</b>	<b>2.468</b>	<b>1.628</b>	<b>915</b>	<b>949</b>	<b>60.453</b>	

Fonte: Istat, Indagine sui matrimoni

\* Per agevolare la lettura la tavola non è stata riportata integralmente. La tavola completa si può consultare cliccando: Tavola 2.4

Guardiamo adesso la tavola 3.2. Nel 1999 Paolo ha tolto l'appendice. In casa hanno preferito un ospedale pubblico («È più sicuro», ha detto il papà). Quindi c'è anche Paolo: è stato in uno dei 218 istituti di cura pubblici del Lazio; in uno dei 32.950 letti. I tre giorni di degenza di Paolo sono compresi nei 9.182.458 complessivi.

Tavola 3.2 - Istituti di cura, posti-letto, degenze, giornate di degenza e personale per regione - Anno 1999\*

REGIONI	Istituti	Posti letto		Degenze		Giornate di degenza		Degenza media (c)
		Numero	Per 1.000 abitanti	Numero	Tasso di ospedalizzazione (a)	Numero	Tasso di utilizzo dei posti letto (b)	
Piemonte	105	22.481	5,2	640.597	149,4	6.158.555	71,1	9,6
Valle d'Aosta	1	457	3,8	16.817	139,9	151.410	90,3	9,0
Lombardia	139	46.099	5,1	1.615.000	178,5	12.609.111	75,9	7,8
Trentino-Alto Adige	35	4.941	5,3	178.691	191,5	1.394.899	77,9	7,8
<i>Bolzano-Bozen</i>	16	2.827	6,1	99.904	216,7	772.547	75,8	7,7
<i>Trento</i>	19	2.114	4,5	78.787	167,0	622.352	80,7	7,9
Veneto	99	22.785	5,1	737.969	164,0	6.467.438	79,5	8,8
Friuli-Venezia Giulia	25	6.605	5,6	187.829	158,6	1.552.497	66,1	8,3
Liguria	40	8.513	5,2	297.341	182,5	2.473.363	81,0	8,3
Emilia-Romagna	91	20.173	5,1	696.402	175,4	5.523.851	79,0	7,9
Toscana	73	16.289	4,6	562.898	159,3	4.452.206	76,2	7,9
Umbria	16	3.494	4,2	147.799	177,2	955.531	76,2	6,5
Marche	51	7.267	5,0	256.738	176,1	1.926.281	75,0	7,5
Lazio	218	32.950	6,3	915.666	174,1	9.182.458	80,4	10,0
Abruzzo	35	6.596	5,2	248.088	194,1	1.929.932	80,5	7,8
Molise	9	1.608	4,9	61.575	187,5	460.461	84,4	7,5
Campania	145	21.488	3,7	931.896	161,0	5.958.227	78,0	6,4
Puglia	105	19.050	4,7	807.101	197,5	4.966.382	72,3	6,2
Basilicata	14	2.456	4,0	96.565	159,1	609.916	68,3	6,3
Calabria	74	9.052	4,4	352.549	171,3	2.389.847	73,6	6,8
Sicilia	129	20.025	3,9	838.585	164,7	5.136.768	71,6	6,1
Sardegna	46	8.109	4,9	248.962	150,6	1.838.100	64,1	7,4
<b>NORD</b>	<b>535</b>	<b>132.054</b>	<b>5,1</b>	<b>4.370.646</b>	<b>170,3</b>	<b>36.331.124</b>	<b>76,1</b>	<b>8,3</b>
<b>CENTRO</b>	<b>358</b>	<b>60.000</b>	<b>5,4</b>	<b>1.883.101</b>	<b>169,9</b>	<b>16.516.476</b>	<b>78,3</b>	<b>8,8</b>
<b>MEZZOGIORNO</b>	<b>557</b>	<b>88.384</b>	<b>4,2</b>	<b>3.585.321</b>	<b>171,6</b>	<b>23.289.633</b>	<b>73,6</b>	<b>6,5</b>
<b>ITALIA</b>	<b>1.450</b>	<b>280.438</b>	<b>4,9</b>	<b>9.839.068</b>	<b>170,7</b>	<b>76.137.233</b>	<b>75,8</b>	<b>7,7</b>

Fonte: Istat, Struttura ed attività degli istituti di cura

(a) Rapporto tra degenze e popolazione media residente per 1.000.

(b) Rapporto tra giornate di degenza effettive e giornate di degenza potenziali per 100. Giornate di degenza potenziali = posti letto per i 365 o 366 giorni dell'anno.

(c) Rapporto tra giornate di degenza e degenze.

\* Per agevolare la lettura la tavola non è stata riportata integralmente.

#### Paolo riflette

«In un anno, negli ospedali pubblici del Lazio, vengono ricoverate 176 persone ogni 1000 che abitano nella regione, magari un poco meno perché qualcuno andrà in ospedale più di una volta ed è contato come più persone. Pensavo che fossero meno quelli che hanno bisogno di cure così importanti. Io ho avuto tante difficoltà a trovare posto, eppure, quell'anno, sono stati occupati, in media, circa 80 posti letto ogni 100 disponibili negli ospedali pubblici della regione. Però, non è un'utilizzazione bassa! In tutta Italia sono stati occupati circa 76 posti letto ogni 100 disponibili. Io sono stato dimesso dopo tre giorni, ma in media la durata della degenza è risultata di dieci giorni, la più alta fra tutte le regioni. Ora ricordo che molti malati aspettavano un poco prima che iniziassero a fare loro le analisi di laboratorio. Forse le degenze si allungano per questo?».

Nella tavola 7.1 possiamo trovare ancora una volta Paolo: c'è la sua scuola, una delle 613 scuole medie del Lazio; la sua classe, una delle 7.619; e lui, uno dei 160.210 alunni.

**Tavola 7.1 - Scuole, classi, alunni e insegnanti delle scuole materne, elementari e medie per regione - Anno scolastico 2000-2001**

ANNI SCOLASTICI REGIONI	Scuole materne				Scuole elementari				Scuole medie			
	Scuole	Sezioni	Bambini	Insegnanti	Scuole	Classi	Alunni	Insegnanti	Scuole	Classi	Alunni	Insegnanti
1996-97	25.944	67.340	1.577.537	123.423	19.906	160.407	2.810.040	289.504	9.119	92.451	1.852.247	231.396
1997-98	25.825	67.790	1.588.020	123.930	19.406	161.294	2.820.919	282.403	8.840	89.534	1.809.059	220.148
1998-99 (a)	25.666	68.199	1.577.696	123.602	19.073	155.940	2.859.379	281.909	8.695	86.904	1.775.009	208.620
1999-2000 (a)	25.208	68.168	1.582.527	125.745	19.068	154.783	2.821.085	283.152	8.496	85.744	1.774.726	205.921
ANNO SCOLASTICO 2000-2001 (a)												
Piemonte	1.625	4.119	100.794	8.315	1.489	10.026	173.430	19.489	543	5.137	106.883	12.988
Valle d'Aosta	91	178	3.021	344	87	380	5.053	652	17	181	3.198	581
Lombardia	3.076	9.655	237.577	17.800	2.543	21.203	395.609	41.489	1.194	11.320	239.751	28.986
Trentino-Alto Adige	608	1.406	28.467	3.642	591	3.807	50.384	6.480	171	1.539	29.675	3.962
<i>    Bolzano-Bozen</i>	<i>320</i>	<i>694</i>	<i>13.768</i>	<i>1.773</i>	<i>338</i>	<i>2.168</i>	<i>26.679</i>	<i>3.528</i>	<i>87</i>	<i>839</i>	<i>15.764</i>	<i>2.255</i>
<i>    Trento</i>	<i>288</i>	<i>712</i>	<i>14.699</i>	<i>1.869</i>	<i>253</i>	<i>1.639</i>	<i>23.705</i>	<i>2.952</i>	<i>84</i>	<i>700</i>	<i>13.911</i>	<i>1.707</i>
Veneto	1.752	5.201	122.509	8.326	1.578	11.696	202.886	21.345	629	5.896	22.146	14.852
Friuli-Venezia Giulia	456	1.141	25.730	2.254	380	2.614	43.631	5.320	164	1.391	27.424	3.522
Liguria	588	1.455	33.915	2.868	515	3.385	58.544	6.525	180	1.743	35.955	4.545
Emilia-Romagna	1.462	3.783	89.932	7.556	1.063	8.287	153.555	16.412	438	4.322	90.521	10.599
Toscana	1.353	3.415	80.916	6.967	1.079	7.740	138.783	14.808	392	4.082	86.816	9.722
Umbria	426	902	19.820	1.799	322	2.129	34.462	3.697	108	1.084	21.955	2.555
Marche	638	1.638	37.911	3.388	504	3.638	64.454	6.561	224	1.920	39.570	4.537
Lazio	1.893	5.887	136.189	10.729	1.464	13.591	257.557	26.002	613	7.619	160.210	18.230
Abruzzo	688	1.545	35.098	3.096	512	3.577	62.520	6.247	228	1.995	40.204	4.457
Molise	184	400	8.453	784	163	988	16.040	1.632	92	558	10.519	1.259
Campania	3.144	8.850	202.032	17.127	2.050	19.776	375.891	35.646	806	11.550	243.391	27.731
Puglia	1.690	5.855	134.271	10.660	831	11.496	236.121	21.041	448	6.951	154.822	16.083
Basilicata	333	830	17.705	1.730	241	1.869	32.709	3.565	142	1.138	22.081	2.975
Calabria	1.485	3.038	64.700	5.991	1.106	7.087	115.258	12.473	455	4.096	79.374	10.204
Sicilia	2.759	6.842	155.325	11.690	1.742	16.063	311.875	29.395	717	9.819	203.769	24.511
Sardegna	790	1.963	42.091	3.906	594	4.642	81.497	8.565	345	3.000	58.686	7.530
<b>ITALIA</b>	<b>25.041</b>	<b>68.103</b>	<b>1.576.456</b>	<b>128.972</b>	<b>18.854</b>	<b>153.994</b>	<b>2.810.259</b>	<b>287.344</b>	<b>7.906</b>	<b>85.341</b>	<b>1.776.950</b>	<b>209.829</b>
<b>Nord</b>	<b>9.658</b>	<b>26.938</b>	<b>641.945</b>	<b>51.105</b>	<b>8.246</b>	<b>61.398</b>	<b>1.083.092</b>	<b>117.712</b>	<b>3.336</b>	<b>31.529</b>	<b>655.553</b>	<b>80.035</b>
<b>Centro</b>	<b>4.310</b>	<b>11.842</b>	<b>274.836</b>	<b>22.883</b>	<b>3.369</b>	<b>27.098</b>	<b>495.256</b>	<b>51.068</b>	<b>1.337</b>	<b>14.705</b>	<b>308.551</b>	<b>35.044</b>
<b>Mezzogiorno</b>	<b>11.073</b>	<b>29.323</b>	<b>659.675</b>	<b>54.984</b>	<b>7.239</b>	<b>65.498</b>	<b>1.231.911</b>	<b>118.564</b>	<b>3.233</b>	<b>39.107</b>	<b>812.846</b>	<b>94.750</b>

Fonte: Istat, Scuole materne statali e non statali; scuole elementari statali e non statali; scuole medie statali e non statali  
(a) Dati provvisori.

«Guarda qui! (Tavola 8.12) - dice Paolo - non credevo che la statistica si sarebbe ricordata di quella volta che sono stato al cinema nel 2000. Invece, fra i biglietti venduti nelle sale cinematografiche, c'è anche il mio e quello dei miei genitori e di mio fratello. Eravamo andati a vedere *Guerre stellari: la minaccia fantasma*».

**Tavola 8.12 - Giorni di spettacolo e biglietti venduti per il cinematografo, per capoluogo di provincia e regione - Anno 2000 (a)**

ANNI REGIONI	Totale			Di cui capoluogo di provincia		
	Giorni di spettacolo	Biglietti venduti		Giorni di spettacolo	Biglietti venduti	
		Totale (in migliaia)	Per abitante (b)		Totale (in migliaia)	Per abitante (b)
1996	585.901	96.512	1,7	325.852	64.871	3,7
1997	645.555	102.782	1,8	365.414	66.884	3,8
1998	667.092	118.504	2,1	354.889	73.610	4,2
1999	727.895	103.483	1,8	387.158	62.913	3,6
<b>2000 - PER REGIONE</b>						
Piemonte	64.392	8.049	1,9	30.222	4.928	3,7
Valle d' Aosta	2.497	193	1,6	993	97	2,8
Lombardia	125.896	18.212	2,0	56.577	10.430	5,0
Trentino- Alto Adige	11.284	1.070	1,1	4.500	517	5,3
Veneto	66.401	8.621	1,9	25.127	3.524	3,5
Friuli-Venezia Giulia	15.592	1.782	1,5	10.799	1.248	3,1
Liguria	34.673	3.903	2,4	21.737	2.585	3,1
Emilia-Romagna	90.418	10.541	2,6	52.774	7.016	4,7
Toscana	60.703	8.131	2,3	34.101	5.121	4,1
Umbria	10.450	1.153	1,4	4.380	654	2,5
Marche	27.569	2.874	2,0	10.980	1.171	4,2
Lazio	117.640	14.786	2,8	90.667	12.585	4,3
Abruzzo	16.934	2.248	1,8	4.695	652	2,2
Molise	2.880	207	0,6	2.028	148	2,0
Campania	44.853	5.938	1,0	24.062	3.497	2,6
Puglia	39.790	4.399	1,1	12.607	1.774	2,0
Basilicata	3.344	274	0,5	1.328	154	1,2
Calabria	8.889	949	0,5	3.958	551	1,2
Sicilia	45.470	6.093	1,2	23.925	3.899	2,3
Sardegna	10.223	1.490	0,9	4.793	975	2,7
<b>ITALIA</b>	<b>799.898</b>	<b>100.911</b>	<b>1,7</b>	<b>420.253</b>	<b>61.526</b>	<b>3,6</b>
<b>Nord</b>	<b>411.153</b>	<b>52.370</b>	<b>2,0</b>	<b>202.729</b>	<b>30.345</b>	<b>4,1</b>
<b>Centro</b>	<b>216.362</b>	<b>26.943</b>	<b>2,4</b>	<b>140.128</b>	<b>19.532</b>	<b>4,1</b>
<b>Mezzogiorno</b>	<b>172.383</b>	<b>21.598</b>	<b>1,0</b>	<b>77.396</b>	<b>11.650</b>	<b>2,2</b>

Fonte: Istat, Attività ricreative e sportive

(a) A seguito delle nuove norme fiscali (D.Lgs. 60/1999), dal 1° gennaio 2000, sono cambiate le modalità di rilevazione ed elaborazione dei dati della S.I.A.E..

(b) Per il calcolo dei valori relativi sono stati utilizzati i dati sulla popolazione residente al 31 dicembre dei rispettivi anni.

#### Paolo riflette

«Che grandi differenze ci sono nel paese! Noi nel Lazio siamo fortunati. Abbiamo molte sale cinematografiche e le frequentiamo; comperiamo ogni anno 2,8 biglietti ciascuno in media. In Basilicata e Calabria, soltanto 0,5 biglietti per persona; quindi lì ognuno va al cinema una volta ogni due anni».

Nella tavola 9.17 c'è lo sciopero che il padre di Paolo ha fatto nel 2001. Lui e i suoi colleghi chiedevano un aumento di stipendio e migliori condizioni di lavoro. In quell'anno, nel Lazio, sono state perse 299.000 ore di lavoro.

**Tavola 9.17 - Conflitti di lavoro, lavoratori partecipanti e ore di lavoro perdute per natura del conflitto e regione - Anni 1997-2001 (valori assoluti, lavoratori ed ore in migliaia)**

ANNI REGIONI	Conflitti originati dal rapporto di lavoro			Conflitti estranei al rapporto di lavoro			Totale		
	Conflitti	Lavoratori partecipanti	Ore di lavoro perdute	Conflitti	Lavoratori partecipanti	Ore di lavoro perdute	Conflitti	Lavoratori partecipanti	Ore di lavoro perdute
1997	920	718	8.150	3	19	149	923	737	8.299
1998	1.097	386	3.807	6	49	256	1.103	435	4.063
1999	753	935	6.364	-	-	-	753	935	6.364
2000	964	668	6.113	2	19	76	966	687	6.189
2001 - PER REGIONE									
Piemonte	78	84	574	-	-	-	78	84	574
Valle d'Aosta	21	5	39	-	-	-	21	5	39
Lombardia	105	307	2.021	-	-	-	105	307	2.021
Trentino-Alto Adige	44	49	311	-	-	-	44	49	311
<i>Bolzano-Bozen</i>	15	16	131	-	-	-	15	16	131
<i>Trento</i>	31	33	180	-	-	-	31	33	180
Veneto	98	82	506	-	-	-	98	82	506
Friuli-Venezia Giulia	42	21	113	1	3	3	43	21	115
Liguria	56	95	515	1	2	14	57	95	529
Emilia-Romagna	103	263	1.759	1	30	60	104	263	1.819
Toscana	83	26	195	1	14	54	84	26	250
Umbria	14	16	91	1	6	6	15	16	97
Marche	62	29	207	2	6	7	64	29	214
Lazio	95	41	299	-	-	-	95	41	299
Abruzzo	24	7	36	-	-	-	24	7	36
Molise	15	1	18	-	-	-	15	1	18
Campania	38	8	87	-	-	-	38	8	87
Puglia	29	7	43	-	-	-	29	7	43
Basilicata	29	8	58	-	-	-	29	8	58
Calabria	18	4	83	-	-	-	18	4	83
Sicilia	33	2	16	-	-	-	33	2	16
Sardegna	40	8	66	-	-	-	40	8	66
<b>ITALIA (a)</b>	<b>741</b>	<b>1.065</b>	<b>7.038</b>	<b>5</b>	<b>60,2</b>	<b>144</b>	<b>746</b>	<b>1.065</b>	<b>7.182</b>
<b>Nord-Centro(a)</b>	<b>632</b>	<b>1.019</b>	<b>6.631</b>	<b>5</b>	<b>60,2</b>	<b>144</b>	<b>637</b>	<b>1.019</b>	<b>6.775</b>
<b>Mezzogiorno (a)</b>	<b>109</b>	<b>46</b>	<b>407</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>109</b>	<b>46</b>	<b>407</b>

Fonte: Istat, Conflitti di lavoro e conflitti non originati da vertenze di lavoro

(a) Il totale del numero dei conflitti non corrisponde alla somma per regione, perché i conflitti che coinvolgono più regioni vengono conteggiati una sola volta a livello di ripartizione geografica. Per questo motivo, per determinare la ripartizione Nord separata da quella del Centro è corretto sommare i lavoratori partecipanti e le ore di lavoro perdute.

Nella tavola 18.10 è ricordata la settimana bianca di Paolo nel febbraio del 2001. Con i genitori stava in una pensione a Moena. Paolo è uno dei 2.616.299 clienti italiani contati nel febbraio di quell'anno.

**Tav 18.10 - Arrivi, presenze e permanenza media degli italiani e stranieri negli esercizi ricettivi per mese - Anno 2001 (a)**

ANNI MESI	Italiani			Stranieri			Totale		
	Arrivi	Presenze	Permanenza media (b)	Arrivi	Presenze	Permanenza media (b)	Arrivi	Presenze	Permanenza media (b)
1997	40.671.342	173.916.564	4,3	29.963.670	118.359.759	4,0	70.635.012	292.276.323	4,1
1998	41.371.579	178.266.093	4,3	30.941.982	121.242.294	3,9	72.313.561	299.508.387	4,1
1999	42.475.852	181.646.770	4,3	31.845.086	126.667.959	4,0	74.320.938	308.314.729	4,2
2000	44.924.162	198.528.158	4,3	35.107.475	140.356.985	4,0	80.031.637	338.885.143	4,2
ANNO 2001									
Gennaio	2.507.312	9.060.797	3,5	1.182.345	4.772.895	3,7	3.689.657	13.833.692	3,0
Febbraio	2.616.299	8.691.352	3,2	1.603.715	6.206.629	3,5	4.220.014	14.897.981	3,0
Marzo	3.016.936	9.075.979	2,9	2.061.695	6.790.619	3,1	5.078.631	15.866.598	3,0
Aprile	3.905.265	10.948.733	2,7	3.240.157	10.297.461	3,1	7.145.422	21.246.194	2,9
Maggio	3.597.776	10.942.078	3,0	3.862.699	13.394.645	3,3	7.460.475	24.336.723	3,1
Giugno	5.136.659	22.350.024	4,5	4.659.432	20.028.077	4,5	9.796.091	42.378.101	4,5
Luglio	5.593.944	36.047.301	6,2	5.162.747	26.338.159	4,9	10.756.691	62.385.460	5,6
Agosto	7.134.131	53.276.358	7,6	4.395.192	23.025.775	5,2	11.529.323	76.302.133	6,7
Settembre	4.121.338	18.667.913	4,5	4.304.808	18.041.436	4,2	8.426.146	36.709.349	4,3
Ottobre	2.970.603	8.681.136	2,8	2.849.065	9.728.357	3,2	5.819.668	18.409.493	3,1
Novembre	2.476.679	6.836.866	2,7	1.370.330	4.193.990	2,8	3.847.009	11.030.856	2,8
Dicembre	2.820.335	8.564.525	2,8	1.061.637	3.599.164	3,2	3.881.972	12.163.689	2,9
<b>Totale</b>	<b>45.897.277</b>	<b>203.143.062</b>	<b>4,4</b>	<b>35.753.822</b>	<b>146.417.207</b>	<b>4,0</b>	<b>81.651.099</b>	<b>349.560.269</b>	<b>4,2</b>

Fonte: Istat, Movimenti dei clienti negli esercizi ricettivi

(a) Dati provvisori

(b) Il valore della permanenza media è determinato dal rapporto tra le presenze e gli arrivi

Paolo si è stancato. Però ha capito due cose:

1. se sfoglia *l'Annuario statistico italiano*, può trovare informazioni su molti aspetti della vita del paese e potrà consultarlo ancora per soddisfare altre curiosità (la presenza di parchi marini nella sua regione, quanti libri si pubblicano in Italia, quanto si spende per andare al cinema o allo stadio, quante persone si dedicano ad attività di volontariato ecc.);
2. egli, come ogni cittadino, è protagonista delle statistiche, in quanto protagonista del vivere sociale e contribuisce a scrivere la storia del paese: sintetizzata nei numeri, un anno di questa storia è lì, nell'Annuario.

### Come si cercano i dati in una tabella

Paolo quest'anno avrà gli esami finali di scuola media. L'evento gli dà qualche preoccupazione, come a tutti i suoi compagni. È curioso di saperne di più. Chissà se può trovare qualche dato nell'Annuario statistico italiano del 2002? Va all'indice analitico e trova a pagina 730 l'indicazione "scrutini p. 161".

Va allora alla pagina 161 dell'Annuario. Ci sono due prospetti. Il primo (Prospetto 7.1\*) contiene i risultati degli scrutini, o meglio il numero di alunni respinti per 100 scrutinati nei diversi tipi di scuole statali, come quella che frequenta, nell'anno scolastico 2000-01. Forse contiene l'informazione che cerca. La prima colonna indica appunto il tipo di scuola: cerca le scuole medie inferiori. La prima riga indica l'anno di corso: cerca il terzo anno, l'ultimo della scuola media, quello dell'esame finale. All'incrocio, i respinti risultano 0,2 ogni 100 scrutinati, cioè 2 ogni 1000. Paolo pensa che non sarà uno di loro. Attenzione, però, c'è una nota. Bisogna leggerla. Ci potrebbe essere qualche precisazione utile. Infatti, il valore di 0,2 per 100 è riferito soltanto a coloro che sono stati ammessi a sostenere gli esami e non al totale degli scrutinati. Tuttavia, per il secondo anno di corso, quando inoltre gli insegnanti sono più severi, i respinti, questa volta proprio su 100 scrutinati, sono risultati 3,3, cioè 33 ogni 1000. Se mai la stessa severità fosse mantenuta anche nel terzo anno di corso, Paolo si sente comunque rincuorato. Non pensa di essere in una posizione così bassa nella graduatoria dei valori. Scorrendo la riga capisce anche che le maggiori difficoltà nella scuola media si incontrano al primo anno. Effettivamente, anche a lui il passaggio dalle elementari aveva posto qualche problema di adattamento.

Paolo è incuriosito e legge anche il prospetto 7.2\* che considera gli alunni licenziati nelle scuole medie statali nell'anno scolastico 2000-01, secondo il giudizio riportato. Ormai ha imparato a leggere una tabella. Entra nella prima colonna su "Centro" (il Lazio è una delle regioni dell'Italia centrale) e va subito alla colonna del giudizio "ottimo", quello al quale aspira. Chissà se sarà effettivamente fra i circa 16 studenti ogni 100 che prevedibilmente avranno quel giudizio come è avvenuto l'anno scorso? Per una volta vorrebbe vivere nel Sud. Lì circa 21 studenti su 100 hanno avuto il giudizio massimo. Saranno più bravi che altrove? Oppure gli insegnanti saranno stati più generosi nella valutazione? Paolo si spiega ora perché la tabella riporta le composizioni percentuali. Se, per ciascuna ripartizione territoriale, il totale degli alunni licenziati viene fatto pari a 100, i dati relativi ai diversi giudizi possono essere confrontati senza difficoltà: il giudizio "sufficiente" è proporzionalmente più frequente nel Nord rispetto al Sud e viceversa, come Paolo aveva già rilevato, avviene per l'opposto giudizio "ottimo". Se la tabella avesse presentato i dati effettivi relativi ai licenziati (i valori assoluti, come si dice), la comparazione sarebbe risultata più difficoltosa.

Paolo torna a considerare il prospetto 7.1. Anche in esso i valori presentati sono percentuali, ma vengono usati per uno scopo un poco diverso. Non compare il 100 che è comunque il riferimento per ciascuno dei valori scritti nella tabella. Quindi, tutti i valori sono confrontabili? In qualche modo sì, tenendo conto tuttavia delle differenze e quindi dei motivi di incomparabilità sostanziali: che significato avrebbe paragonare gli insuccessi nella scuola elementare con quelli delle superiori? È invece interessante capire che i respinti al primo anno degli istituti professionali e tecnici sono proporzionalmente assai più numerosi che non nei licei classico e scientifico. Per capirne i motivi, l'approfondimento dovrà avvalersi di altri contributi disciplinari: della sociologia, della pedagogia, dell'economia.

Paolo ha capito che ha ancora molto da imparare e anche che se i dati raccolti (nel caso del prospetto 7.1, i risultati degli scrutini per ciascuno studente) vengono messi tutti insieme, sistemati in una tabella, o elaborati con qualche metodo suggerito dalla statistica, fanno capire meglio le situazioni e i fenomeni con i quali si entra in contatto.

\*Nota:  
 Prospetto 7.1 e prospetto 7.2

### Prospetto 7.1

#### Risultati degli scrutini nelle scuole statali - Anno scolastico 2000-2001

TIPI DI SCUOLA	Respinti per 100 scrutinati				
	1° anno	2° anno	3° anno	4° anno	5° anno (a)
Scuole elementari	0,4	0,3	0,2	0,1	0,2
Scuole medie inferiori	4,7	3,3	(a) 0,2		
Scuole superiori	18,2	12,7	11,0	10,2	2,2
Istituti professionali	26,5	19,3	(b) 5,0	15,6	3,8
Istituti tecnici	19,7	13,8	14,7	11,6	2,7
Istituti magistrali	14,6	8,8	9,1	(a) 2,2	
Licei scientifici	6,1	5,4	6,0	3,4	0,8
Licei ginnasi	8,6	5,4	7,5	5,1	0,7
Istituti d'arte	18,3	14,9	(b) 5,1	12,6	1,6
Licei artistici	18,3	10,7	10,5	(a) 1,6	

Fonte: Ufficio Statistico MIUR, Indagine campionaria sui risultati degli scrutini nella scuola statale.

(a) I dati sull'ultimo anno di corso si riferiscono agli esiti degli esami del diploma conclusivo di Stato e non includono i non ammessi a sostenere gli esami.

(b) I dati del terzo anno degli istituti professionali e degli istituti d'arte si riferiscono agli esiti degli esami di licenza e di qualifica professionale e non includono i non ammessi a sostenere gli esami.

### Prospetto 7.2

#### Anni licenziati nella scuola media statale per giudizio riportato agli esami - Anno scolastico 2000-2001 (composizione percentuale)

RIPARTIZIONI TERRITORIALI	Giudizio riportato agli esami				Totale
	Sufficiente	Buono	Distinto	Ottimo	
Nord	42,5	26,3	17,8	13,4	100,0
Centro	40,2	25,7	18,3	15,8	100,0
Sud	37,0	24,1	18,2	20,7	100,0
Isole	44,5	22,8	15,2	17,5	100,0
Totale	40,6	25,0	17,7	16,7	100,0

Fonte: Ufficio Statistico MIUR, Indagine campionaria sui risultati degli esami di licenza nelle scuole medie statali.

## Esercitazioni

- 1) Con un gruppo di amici, scegliete 20 tabelle dell'*Annuario statistico italiano* (meglio se attinenti alla vostra condizione di giovani: scuola, spettacoli, consumi ecc.) e provate a ritrovarvi nei numeri. Vince chi vi si ritrova più spesso. Ciò può avvenire perché è più perspicace, oppure perché nell'anno (o negli anni) al quale i dati si riferiscono ha avuto una vita più movimentata. Ed è per questo che le statistiche gli danno più spazio.
- 2) Scelto un fenomeno che vi interessa, consultate uno degli ultimi annuari e rintracciate le informazioni presenti più vicine alla questione posta.

## 2. Una lettura statistica dei giornali

### Che cosa si impara nel capitolo 2

Si può valutare e confrontare il contenuto, ad esempio di tre diversi settimanali, riducendo gli aspetti qualitativi a valori numerici: quante pagine sono dedicate alla pubblicità, quante all'attualità e quante agli altri argomenti. Il confronto risulta efficace se calcoliamo le percentuali, le **frequenze assolute**, le **frequenze relative** e se rappresentiamo i risultati attraverso il **diagramma a nastri suddivisi**. Il capitolo fornisce la regola per approssimare le cifre.

**Frequenza assoluta.** Numero di elementi che hanno una determinata caratteristica, che presentano una determinata modalità del carattere di interesse

**Frequenza relativa.** Rapporto tra la frequenza assoluta degli elementi che presentano una determinata modalità del carattere di interesse e il totale degli elementi da analizzare

**Diagramma a nastri suddivisi.** Rappresentazione grafica di un carattere, costituita da una successione di nastri (a forma di rettangolo) aventi superficie proporzionale alle frequenze secondo le quali si presentano le diverse modalità del carattere. Le altezze dei rettangoli sono costanti.

Per confrontare un giornale con un altro, possiamo prenderne in considerazione la struttura grafica (articoli di fondo, rubriche, illustrazioni e così via), il modo di realizzare i titoli (a grandi caratteri, con sottotitoli, sull'intera pagina eccetera), o il modo di dare le informazioni (con poche righe di commento o con lunghe considerazioni).

Proviamo a esaminare, a titolo esemplificativo, tre riviste dalle caratteristiche diverse (un numero di ciascuna) per confrontarle e per ricavare alcune informazioni sulle loro particolarità:

Donna moderna (A) settimanale femminile, Mondadori Editore, n. 3, 2003  
Oggi (B) settimanale di attualità, Rizzoli Editore, n. 7, 2003  
Panorama (C) settimanale di politica, attualità e cultura, Mondadori Editore, n. 52, 2003

Per ciascuna di queste riviste abbiamo rilevato:

- il numero di pagine complessivo, escluse la prima e l'ultima di copertina;
- le pagine (o frazioni di pagina) dedicate alla pubblicità diretta, di marchi e prodotti vari;
- il numero delle altre pagine (per differenza).

Le pagine della terza categoria sono state successivamente suddivise in tre classi:

- articoli di moda, di arredamento, di cure estetiche;
- articoli di attualità;
- altri testi (sommari, composizione della redazione, oroscopi, critiche cinematografiche, teatrali, televisive, programmi televisivi, recensioni di libri, vignette e fumetti).

La soluzione adottata ha chiaramente carattere soggettivo e, a seconda delle vostre esigenze, potrete scegliere la classificazione che più riterrete utile al vostro caso e ai vostri fini. Ovviamente la vostra scelta va argomentata e spiegata con chiarezza.

Nella nostra rilevazione abbiamo contato il numero delle pagine per le varie categorie della classificazione e abbiamo misurato gli spazi, suddividendoli per testo scritto o fotografico.

Che cosa possiamo leggere su questa tavola? Molte informazioni: ad esempio, la pubblicità è largamente presente nella rivista C, mentre sono poco presenti gli articoli di moda; nel settimanale A le fotografie che accompagnano gli articoli occupano molto spazio.

Ma le tre riviste hanno un numero di pagine diverso e quindi le comparazioni diventano difficili. Per un confronto più preciso, è meglio considerare delle proporzioni o delle percentuali, in modo da rendere uguale a 1 o a 100 il numero totale di pagine e quindi più facili i raffronti.

**Tavola 1 - Suddivisione delle pagine delle tre riviste (A, Donna moderna - B, Oggi - C, Panorama) secondo la classificazione adottata (a)**

DESCRIZIONE		A	B	C
a) pagine totali		142	130	370
b) pubblicità		50	44	177
c1) articoli di moda, di arredamento, di cure estetiche	t.s.	14	10	5
	t.f.	30	4	3
	tot.	44	14	8
c2) articoli di attualità	t.s.	15	23	99
	t.f.	6	18	17
	tot.	21	41	116
c3) altri testi (sommari, composizioni redazionali, oroscopi, programmi televisivi, critiche cinematografiche, ecc.)	t.s.	21	25	61
	t.f.	6	6	8
	tot.	27	31	69

(a) Vedi il testo per le specificazioni. I simboli seguenti stanno a significare:  
t.s. = testo scritto  
t.f. = testo fotografato.

Ad esempio, che percentuale di pagine Donna moderna dedica alla pubblicità? Bisogna impostare una proporzione:

$$\frac{\text{pagine di pubblicità}}{\text{pagine in totale}} = \frac{x}{100}$$

e moltiplicando per 100 il 1° e il 2° membro avremo:

$$100 \frac{50}{142} = 100 \frac{x}{100}$$

$$x = 100 \frac{50}{142} = 35,2$$

Su questa base procediamo a formulare di nuovo i dati nella tavola 2; questa volta si tratta di valori percentuali.

**Tavola 2 - Valori percentuali per ciascuna delle riviste analizzate nella Tavola 1**

DESCRIZIONE	A	B	C
b) pagine di pubblicità	35,2	33,9	47,8
c1) articoli di moda	31,0	10,8	2,2
c2) articoli di attualità	14,8	31,5	31,4
c3) altri testi	19,0	23,9	18,6
a) pagine totali	100	100,1 (a)	100

(a) per le approssimazioni numeriche utilizzate nel calcolo, il totale non dà esattamente 100. Quando ciò accade è preferibile lasciare in evidenza l'approssimazione (in questo caso a un millesimo) piuttosto che "far tornare" comunque le cifre.

Ora possiamo avere altre e più precise indicazioni:

- Panorama ha la quota maggiore di pubblicità (47,8 per cento);
- nella tavola 1 risulta che Oggi dedica alla pubblicità 44 pagine e Donna moderna 50; in realtà le pagine pubblicitarie rappresentano 35,2 per cento di Donna moderna e il 33,9 per cento di Oggi; gli articoli veri e propri coprono il 14,8 per cento delle pagine di Donna moderna, il 31,5 per cento delle pagine di Oggi e il 31,4 per cento delle pagine di Panorama; lo spazio dedicato agli altri testi (critiche, recensioni di film, dischi, libri eccetera) è leggermente più ampio in Oggi (23,9 per cento del totale), quasi uguale in Panorama (18,6 per cento) e in Donna moderna (19,0 per cento).

Da queste prime analisi appaiono abbastanza chiaramente le diverse caratteristiche dei tre settimanali.

Analizziamo ora separatamente il modo in cui vengono presentate le informazioni e verifichiamo le proporzioni fra testo fotografico e testo scritto in tutte le pagine che non siano a carattere pubblicitario (Tavola 3).

**Tavola 3 - Classificazione delle pagine non dedicate a messaggi pubblicitari delle tre riviste, suddivise a seconda che il testo sia scritto (t.s.) oppure che sia di carattere fotografico (t.f.)**

RIVISTE	Frequenze assolute		
	t.s.	t.f.	Totale
A	50	42	92
B	58	28	86
C	165	28	193

Anche in questo caso conviene calcolare le proporzioni in modo da rendere pari a 100 il totale delle pagine esaminate e rendere così più chiara la differenza fra testo scritto e fotografico in ciascuna delle tre riviste permettendo opportuni confronti (Tavola 4).

**Tavola 4 - Distribuzione percentuale delle pagine secondo i dati della tavola 3**

RIVISTE	Frequenze assolute		
	t.s.	t.f.	Totale
A	54,3	45,7	100
B	67,4	32,6	100
C	85,5	14,5	100

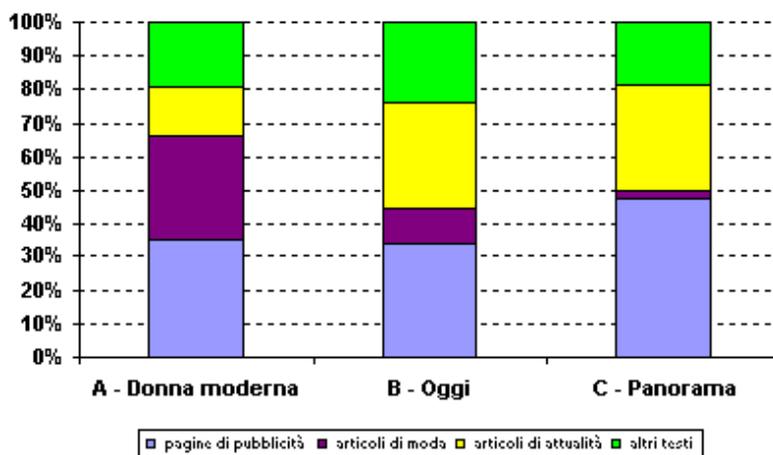
Cominciamo col rilevare che Donna moderna e Oggi contengono poco testo scritto, quindi sono diverse dai libri comuni nei quali il testo scritto occupa gran parte delle pagine. Inoltre:

- la rivista femminile contiene meno "scritto" delle altre due (54,3 per cento);
- la parte occupata da testi scritti aumenta progressivamente, passando dalla rivista al settimanale d'informazione, al settimanale politico-culturale (da Donna moderna a Oggi a Panorama), così qualificandoli chiaramente e permettendo di farsi un'idea più precisa sul lettore di ciascuna delle tre testate;
- la parte occupata da testi fotografici aumenta progressivamente passando da Panorama e Oggi a Donna moderna.

Per le nostre tavole abbiamo considerato come unità statistica, cioè come unità elementare su cui effettuare la rilevazione, la singola pagina. Il *carattere statistico* che abbiamo rilevato riguarda la superficie della pagina a seconda del contenuto (Tavole 1 e 2) e del tipo di testo utilizzato (Tavole 3 e 4).

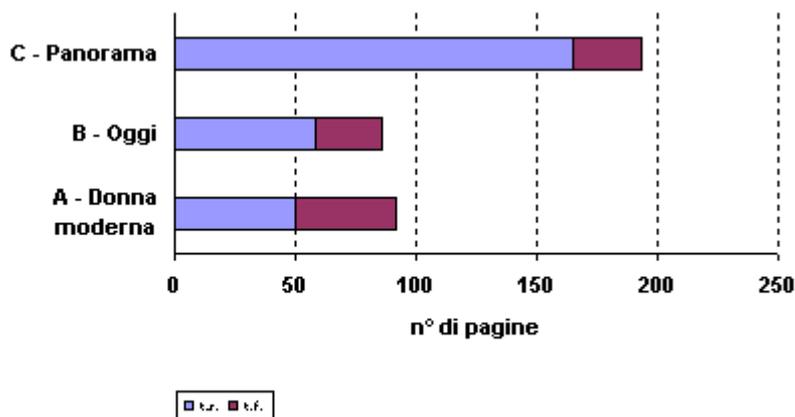
C'è un altro modo per presentare le risultanze statistiche, basato sulla raffigurazione grafica dei valori ottenuti, che integra la comunicazione numerica della dimensione dei fenomeni e arricchisce la nostra capacità di percepirla e di comprenderla.

In particolare, possiamo rappresentare i valori percentuali della tavola 2 con dei *diagrammi a nastri suddivisi*, tali che le varie parti abbiano superficie proporzionale alle rispettive frequenze percentuali.



**Figura 1 – Diagrammi a nastri suddivisi per tre settimanali, costruiti sulla base delle distribuzioni percentuali delle pagine individuate nella tavola 2**

I dati della tavola 3 possono essere raffigurati sempre con dei *diagrammi a nastri suddivisi*, di superficie proporzionale al numero totale di pagine, suddivisi al loro interno, a seconda che il testo sia scritto o fotografico.



**Figura 2 – Diagrammi a nastri suddivisi per tre settimanali secondo i dati della tavola 3**

### Approssimazioni

Se la vostra spanna (cioè la lunghezza della mano aperta e distesa, dall'estremità del mignolo a quella del pollice) è pari a *circa* 18 cm e misurate la lunghezza del tavolo su cui state leggendo, trovandola di 5 spanne e un po', potete misurare *approssimativamente* la lunghezza del tavolo.

Si hanno  $5 \times 18 = 90$  cm a cui vanno aggiunti *un po'* di centimetri che potete valutare, ad esempio, in 4 o 5; dunque una lunghezza totale di cm 94 o 95.

Questa misura approssimata va benissimo se dovete comperare una tovaglia per ricoprire il tavolo, poiché la acquirerete più ampia in modo che ricada lungo i bordi di 20 o 30 centimetri per parte; non va affatto bene se dovete rivestire con un foglio di plastica la superficie del vostro tavolo. Il foglio di plastica deve risultare della stessa lunghezza del tavolo e quindi le misure vanno prese con precisione: forse anche l'errore di 1 mm può far riuscire male il lavoro.

In matematica si opera come nella realtà. Sulla calcolatrice stabiliamo  $1:6 = 0,16666667$  (*attenzione*: la vostra calcolatrice potrebbe visualizzare meno cifre di quelle che abbiamo usato). La macchina ha *approssimato* l'ultima cifra scritta, poiché l'espressione corretta è:

$1/6 = 0,1\overline{6}$ . Per molte applicazioni questo risultato è troppo *lungo* e possiamo abbreviarlo (approssimarlo) per eccesso o per difetto (cioè con un numero *vicino* più grande o, rispettivamente, più piccolo), come nel seguente specchietto.

Approssimazioni	
per difetto per eccesso	a meno di
0,1 0,2	1/10
0,16 0,17	1/100
$0,166 < \frac{1}{6} < 0,167$	1/1000
0,1666 0,1667	1/10.000
0,16666 0,16667	1/100.000

Ecco due casi: i valori seguenti divengono

$3,25 \overline{6} \rightarrow 3,26$   $3,25 \overline{3} \rightarrow 3,25$

quando sia utile effettuare un'approssimazione ai centesimi.

Questo è il motivo per cui può accadere che la somma di più percentuali non dia esattamente 100, ma un valore approssimato, come nel caso della tavola 2.

Sulla destra leggete la precisione dell'approssimazione: a meno di un decimo vuol dire che si commette un errore (in difetto o in eccesso) più piccolo di 0,1 (1/10; un decimo).

Usualmente nei calcoli statistici si opera per approssimazione con la regola che i decimali che terminano per 0, 1, 2, 3, 4 vengono trascurati e quelli che terminano per 5, 6, 7, 8, 9 fanno elevare di una unità la cifra alla loro sinistra.

## Esercitazioni

1. Analizzare la parte dedicata alla pubblicità su due giornali quotidiani dello stesso giorno.
2. Analizzare la parte dedicata agli annunci economici di tre quotidiani nazionali, edizione della domenica.
3. Analizzare la parte dedicata ai titoli della prima pagina di un giornale sportivo e di un quotidiano politico.
4. Verificare le modifiche intervenute nel corso del tempo in uno stesso quotidiano o settimanale (un numero del 1998 e uno del 2001, possibilmente dello stesso mese e giorno all'interno della settimana).
5. Discutere se la classificazione utilizzata in questo capitolo avrebbe dato risultati differenti se invece delle pagine si fossero utilizzate le superfici (espresse in  $\text{cm}^2$ ).
6. Sviluppare e discutere esempi di misure approssimate prese in classe (tempi misurati da più studenti per la durata di uno stesso evento; lunghezza di uno stesso oggetto misurato sia dalla stessa persona più volte, in momenti successivi, sia da più persone).
7. Ricercare sul libro di matematica (ma anche su altri libri) se c'è una parte dedicata alle approssimazioni numeriche e alle approssimazioni relative alle calcolatrici. Osservate che cosa accade calcolando  $3/5$  e  $5/3$  e prendendo il prodotto dei risultati, quale risulta dalla calcolatrice.

### 3. Uomini, donne, sport

#### Che cosa si impara nel capitolo 3

Per approfondire le determinanti di un fenomeno calcoliamo i **rapporti di derivazione**. Per leggere agevolmente dati che evolvono nel tempo – i primati italiani per il salto in alto e per il nuoto – utilizziamo i grafici cartesiani e i **tassi di variazione**.

Abbiamo raccolto presso la Fidal (Federazione italiana di atletica leggera) e presso la Fin (Federazione italiana nuoto) i dati relativi all'evoluzione nel tempo di alcuni primati sportivi maschili e femminili (Tavole 1 e 2).

**Tavola 1 - Primati italiani per il salto in alto senza pedana, maschili e femminili, nel tempo (in metri)**

Maschile				Femminile			
PRIMATI	Primatista	Luogo di	Data di conseguimento	PRIMATI	Primatista	Luogo di	Data di conseguimento
1,70	C. Colombo	Milano	30-6-1889	1,15	E. Ghiringhelli	Montecarlo	16-4-1922
1,70	O. Pizio	Milano	16-9-1900	1,15	B. Pizzini	Montecarlo	16-4-1922
1,725	G. Torretta	Roma	31-3-1906	1,30	A. Sacco	Roma	20-5-1922
1,75	L. Brambilla	S.S. Giovanni	- 4-1906	1,33	A. Sacco	Milano	11-3-1923
1,75	A. Vecchi	Carpi	7-10-1906	1,35	L. Banzi	Montecarlo	4-4-1923
1,75	G. Torretta	Vigevano	28-10-1906	1,35	A. Sacco	Montecarlo	4-4-1923
1,75	G. Torretta	Venezia	8-5-1907	1,37	L. Banzi	Montecarlo	4-4-1923
1,75	E. Brambilla	Vigevano	11-10-1908	1,40	L. Banzi	Milano	17-6-1923
1,75	E. Brambilla	Milano	22-11-1908	1,40	A. Sacco	Milano	20-7-1924
1,75	A. Pedrelli	Bologna	8-12-1909	1,41	A. Sacco	Imola	14-9-1924
1,75	C. Butti	Milano	19-2-1911	1,42	S. Martini	Dalmine	15-7-1928
1,75	A. Pagani	Tivoli	9-7-1911	1,43	O. Valla	Firenze	5-10-1930
1,75	A. Tonini	Roma	8-6-1912	1,45	O. Valla	Bologna	3-5-1931
1,75	C. Butti	Alessandria	18-5-1913	1,48	O. Valla	Bologna	18-6-1931
1,75	G. Tugnoli	Firenze	30-4-1916	1,48	O. Valla	Milano	27-8-1933
1,770	P. Pisati	La Spezia	17-8-1919	1,50	O. Valla	Verona	1-10-1933
1,781	C. Ghiringhelli	Legnano	2-10-1921	1,517	O. Valla	Udine	8-10-1933
1,805	E. Uicich	Roma	22-4-1923	1,54	C. Testoni	Torino	27-9-1936
1,835	G. Corona	Cagliari	13-1-1924	1,56	O. Valla	Bologna	5-9-1937
1,843	G. Palmieri	Napoli	13-6-1926	1,57	P. Paternoster	Roma	25-9-1955
1,85	G. Palmieri	Roma	21-4-1927	1,60	P. Paternoster	Roma	10-5-1956
1,86	G. Palmieri	A. Piceno	17-7-1927	1,61	P. Paternoster	Strasburgo	27-5-1956
1,86	G. Palmieri	Padova	9-6-1929	1,62	P. Paternoster	Napoli	24-6-1956
1,863	A. Tommasi	Verona	27-9-1931	1,63	O. Giardi (1)	Bologna	14-9-57
1,867	A. Tommasi	Verona	30-3-1932	1,64	M. Bortoluzzi	Roma	6-9-1959
1,90	A. Tommasi	Milano	15-5-1932	1,65	M. Bortoluzzi	Carrara	12-6-1960
1,905	A. Tommasi	Firenze	26-6-1932	1,65	M. Bortoluzzi	Roma	11-6-1961
1,915	A. Tommasi	Verona	1-10-1933	1,66	M. Bortoluzzi	Ostia	25-6-1961
1,915	A. Tommasi	Bologna	23-8-1936	1,67	O. Giardi	Bergamo	23-9-1962
1,92	R. Dotti	Bologna	23-7-1938	1,68	A. R. Bellamoli	Milano	28-6-1969
1,93	A. Campagner	Torino	18-6-1939	1,70	L. Bortoli	Padova	14-4-1970
1,95	A. Campagner	Parma	26-5-1940	1,71	S. Simeoni	Padova	9-5-1970
1,96	A. Campagner	Parma	5-10-1941	1,72	S. Simeoni	Roma	16-5-1970
1,98	A. Campagner	Parma	14-6-1942	1,73	S. Simeoni	Roma	14-7-1970
1,99	G. Roveraro	Bologna	24-6-1956	1,73	S. Simeoni	Bucarest	2-8-1970
2,01	G. Roveraro	Lugano	9-9-1956	1,75	S. Simeoni	C. Veneto	6-9-1970
2,02	G. Roveraro	Genova	6-10-1957	1,75	L. Bortoli	Padova	23-5-1971
2,03	W. Zamparelli	Roma	7-4-1962	1,75	S. Massenz	Torino	3-7-1971
2,03	R. Galli	Roma	7-4-1962	1,76	S. Simeoni	Roma	7-7-1971
2,04	W. Zamparelli	Roma	21-4-1962	1,77	S. Massenz	Ancona	17-7-1971
2,04	A. Brandoli	Milano	5-7-1962	1,78	S. Simeoni	Helsinki	12-8-1971
2,05	M. Bogliatto	Alessandria	1-5-1963	1,78	S. Simeoni	Madrid	12-9-1971
2,06	R. Galli	Pisa	1-6-1963	1,80	S. Simeoni	Madrid	12-9-1971
2,08	R. Galli	Pisa	1-6-1963	1,80	S. Simeoni	Londra	5-8-1972
2,09	M. Bogliatto	Porto Alegre	8-9-1963	1,82	S. Simeoni	Monaco	4-9-1972
2,09	M. Bogliatto	Roma	21-8-1965	1,85	S. Simeoni	Monaco	4-9-1972

**Tavola 1 segue - Primati italiani per il salto in alto senza pedana, maschili e femminili, nel tempo (in metri)**

Maschile				Femminile			
PRIMATI	Primatista	Luogo di	Data di conseguimento	PRIMATI	Primatista	Luogo di	Data di conseguimento
2,10	M. Bogliatto	Roma	10-10-1965	1,86	S. Simeoni	Roma	15-9-1973
2,11	E. Azzaro	Sindelfingen	19-6-1966	1,86	S. Simeoni	Roma	8-9-1974
2,12	G. Crosa	Roma	26-5-1968	1,89	S. Simeoni	Roma	8-9-1974
2,12	G. Crosa	Città del Messico	19-10-1968	1,90	S. Simeoni	Avezzano	22-9-1974
2,14	G. Crosa	Città del Messico	19-10-1968	1,90	S. Simeoni	Torino	8-7-1976
2,14	G. Crosa	Città del Messico	20-10-1968	1,91	S. Simeoni	Montreal	28-7-1976
2,15	E. Azzaro	Formia	11-5-1969	1,92	S. Simeoni	S. Sebastiano	12-3-1977
2,16	E. Azzaro	Milano	29-6-1969	1,93	S. Simeoni	Formia	15-5-1977
2,17	E. Azzaro	Atene	16-9-1969	1,95	S. Simeoni	Milano	23-2-1978
2,17	E. Azzaro	Formia	7-5-1970	1,95	S. Simeoni	Formia	18-6-1978
2,17	E. Azzaro	Siracusa	5-7-1970	1,97	S. Simeoni	Kouvola	11-7-1978
2,18	E. Azzaro	Rieti	28-8-1971	1,98	S. Simeoni	Brescia	4-8-1978
2,18	E. Azzaro	Madrid	11-9-1971	2,01	S. Simeoni	Brescia	4-8-1978
2,19	E. Del Forno	Milano	26-6-1973	2,01	S. Simeoni	Praga	31-8-1978
2,20	E. Del Forno	Genova	27-2-1974				
2,21	E. Del Forno	Udine	23-3-1974				
2,20	E. Del Forno	Viareggio	7-8-1974				
2,20	E. Del Forno	Milano	1-5-1975				
2,20	G. Ferrari	F. Modenese	29-5-1975				
2,21	E. Del Forno	Siena	16-7-1975				
2,22	E. Del Forno	Siracusa	8-10-1975				
2,22	R. Bergamo	Milano	8-6-1976				
2,23	R. Fortini	Livorno	13-6-1976				
2,23	O. Raise	Milano	4-2-1978				
2,24	O. Raise	Trinec	3-3-1978				
2,24	R. Bergamo	Roma	25-7-1978				
2,26	B. Bruni	Genova	3-2-1979				
2,25	M. Di Giorgio	N. Gorica	15-4-1979				
2,26	M. Di Giorgio	Udine	20-5-1979				
2,27	O. Raise	Bologna	19-9-1979				
2,27	Di Giorgio	Bologna	19-9-1979				
2,27	B. Bruni	Bologna	19-9-1979				
2,28	P. Borghi	S. L. del Piave	25-5-1980				
2,29	M. Di Giorgio	Pisa	5-7-1980				
2,30	M. Di Giorgio	Udine	15-6-1981				
2,30	L. Toso	Padova	13-6-1988				
2,32	L. Toso	Torino	21-7-1988				
2,33	M. Benvenuti	Verona	12-9-1989				

Fonte: Fidal

(1) Osvalda Giardi (Cus Pisa) saltò 1,63 a Bologna il 14 settembre 1957, ma con l' aiuto di una scarpetta a suola più alta del consentito.

**Tavola 2 - Primati italiani per i 100 metri stile libero, maschili e femminili, nel tempo (in minuti, secondi e decimi di secondo)**

Maschile				Femminile			
PRIMATI (a)	Primatista	Luogo di	Data di conseguimento	PRIMATI (a)	Primatista	Luogo di	Data di conseguimento
1.05.6	E. Polli	Budapest	20-8-26	1.35.4	A. Ippavitz	Milano	9-8-24
1.05.6	E. Polli	Bologna	3-9-27	1.32.8	N. Bravin	Bologna	15-8-27
1.03.4	E. Polli	Roma	1-9-28	1.24.4	M. Schwartz	Roma	2-9-28
1.03.0	S. Banchelli	Bologna	9-8-31	1.19.6	A. Savi	Roma	15-8-29
1.02.1	D. Cappellini	Torino	4-6-33	1.18.0	I. Sulligi	Roma	16-8-31
1.01.2	G. Gambetta	Torino	14-7-35	1.17.2	A. Savi	Bologna	7-9-31
1.00.5	M. Costa	Torino	20-2-37	1.16.2	B. Lokar, T.	Sanremo	25-8-35
1.00.3	W. Lisardi	Genova	19-9-37	1.15.8	G. Ruzzier	Milano	7-9-35
1.00.0	C. Brunelleschi	Firenze	30-4-48	1.15.6	B. Lokar	Genova	14-9-35
59.5	C. Pedersoli	Salsom	19-9-50	1.14.0	B. Lokar	Trieste	2-9-37
58.9	C. Pedersoli	Genova	28-4-51	1.13.7	B. Lokar	Milano	4-6-39
58.2	C. Pedersoli	Torino	21-6-52	1.12.3	B. Lokar	Trieste	28-6-39
57.2	A. Romani	L'Aquila	19-8-55	1.12.0	B. Lokar	Trieste	26-7-39
				1.11.7	B. Lokar	Trieste	8-8-39
				1.10.6	B. Lokar	Trieste	3-9-39
	NUOVO REGOLAMENTO						
58.3	P. Pucci	Melbourne	29-11-56	1.10.6	D. Gamacchio	Napoli	10-6-50
57.0	P. Pucci	Roma	13-7-57	1.09.8	D. Gamacchio-Finc.	Levanto	22-7-50
56.8	P. Pucci	Bologna	26-7-58	1.09.5	D. Gamacchio-Fine.	Genova	9-9-50
56.3	P. Pucci	Budapest	31-8-58	1.09.4	R. Calligari	Torino	21-6-52
56.1	P. Pucci	Budapest	31-8-58	1.08.8	R. Calligaris	Genova	5-7-52
56.1	B. Bianchi	Sanremo	19-7-64				
55.8	P. Boscaini	Tokyo	11-10-64		NUOVO REGOLAMENTO		
55.7	P. Boscaini	Sanremo	18-7-65	1.09.3	S. Valle	Napoli	22-6-56
55.4	P. Boscaini	Milano	7-8-65	1.09.1	S. Valle	Genova	23-8-57
55.2	P. Boscaini	Roma	23-3-66	1.08.7	S. Valle	Genova	7-9-57
55.1	P. Boscaini	Tunisi	11-9-67	1.08.6	S. Valle	Spalato	4-10-58
54.7	P. Boscaini	Milano	20-8-68	1.08.4	M. C. Pacifici	Roma	14-6-59
54.6	R. Pangano	Roma	1-8-70	1.08.3	P. Saini	Genova	7-8-59
54.4	R. Pangano	Siracusa	5-7-72	1.07.7	M. C. Pacifici	Genova	7-8-59
54.3	R. Pangano	Dortmund	14-4-73	1.07.7	P. Saini	Roma	9-4-60
54.06	R. Pangano	Belgrado	5-9-73	1.06.7	P. Saini	Roma	25-4-60
53.87	R. Pangano	Belgrado	9-9-73	1.05.7	P. Saini	Roma	18-6-60
53.6	R. Pangano	Roma	20-7-74	1.05.3	P. Saini	Milano	17-7-60
53.32	R. Pangano	Vienna	18-8-74	1.04.4	P. Saini	Roma	26-8-60
53.08	R. Pangano	Vienna	18-8-74	1.04.0	D. Beneck	Roma	6-4-62
52.59	R. Pangano	Vienna	23-8-74	1.03.5	P. Saini	Mosca	24-5-62
52.5	M. Guarducci	Las Palmas	5-4-75	1.03.3	D. Beneck	Dieren	16-8-64
51.57	M. Guarducci	Montreal	24-7-76	1.03.2	D. Beneck	Tokio	12-10-64
51.35	M. Guarducci	Montreal	24-7-76	1.02.9	D. Beneck	Tokio	12-10-64
51.25	M. Guarducci	Chiavari	8-9-77	1.02.6	D. Beneck	Roma	5-9-65
51.24	M. Guarducci	Rio de Janeiro	9-4-79	1.02.5	D. Beneck	Roma	5-9-65
51.21	M. Guarducci	Warendorf	18-9-82	1.02.4	D. Beneck	Utrecht	21-8-66
50.99	M. Guarducci	Casablanca	7-9-83	1.02.3	M. Sacchi	Sanremo	24-9-68
50.99	M. Guarducci	Casablanca	7 9 1983	1.02.2	N. Calligaris	Barcellona	11-9-70
50.97	G. Lamberti	Roma	31 5 1987	1.02.0	N. Calligaris	Smirne	9-10-71
50.82	G. Lamberti	Catania	5 7 1987	1.01.8	N. Calligaris	Berlino	8-4-72
50,55	G. Lamberti	Strasburgo	20 8 1987	1.01.4	L. Podestà	Milano	4-8-73
50.53	G. Lamberti	Strasburgo	21 8 1987	1.01.38	L. Podestà	Belgrado	8-9-73
50.47	R. Gleria	Firenze	25 3 1988	1.01.36	L. Podestà	Belgrado	9-9-73
50.41	G. Lamberti	Genova	9 7 1989	1.00.5	C. Savi-Scarponi	Pescara	4-6-77
49.48	G. Lamberti	Bonn	17 8 1989	1.00.47	C. Savi-Scarponi	Chiavari	8-9-77
49.24	G. Lamberti	Bonn	17 8 1999	59.97	C. Savi-Scarponi	Chiavari	8-9-77
49.23	L. Vismara	Sydney	16 9 2000	59.59	C. Savi-Scarponi	Roma	12-3-78
				59.55	C. Savi-Scarponi	S. Juan	10-4-78
				59.52	C. Savi-Scarponi	Milano	13-7-78
				59.34	M. Dalla Valle	Milano	20-7-79
				59.25	M. Vallarin	Torino	30-3-80
				59.10	M. Vallarin	Piacenza	15-6-80

**Tavola 2 segue - Primati italiani per i 100 metri stile libero, maschili e femminili, nel tempo (in minuti, secondi e decimi di secondo)**

Maschile			Femminile		
PRIMATI (a)	Primatista	Luogo di Data di conseguimento	PRIMATI (a)	Primatista	Luogo di Data di conseguimento
			57.98	S. Persi	Roma 19-7-83
			57.62	S. Persi	Los Angeles 29-7-1984
			57.24	S. Persi	Los Angeles 29-7-1984
			57.13	S. Persi	Catania 05-7-1987
			57.08	S. Persi	Strasburgo 18-8-1987
			56.97	S. Persi	Bonn 15-8-1999
			56.84	V. Susin	San Donato 10-7-1997
			56.70	C. Chiuso	Roma 01-6-2000
			56.37	L. Striani	Monfalcone 05-8-2000
			56.10	C. Vianini	Sydney 16-9-2000
			55.96	C. Vianini	Sydney 06-9-2000
			55.80	C. Vianini	Fukuoka 24-7-2001
			55.80	eg C. Vianini	Fukuoka 24-7-2001
			55.07	C. Vianini	Genova 06-8-2001

Fonte: Fidal

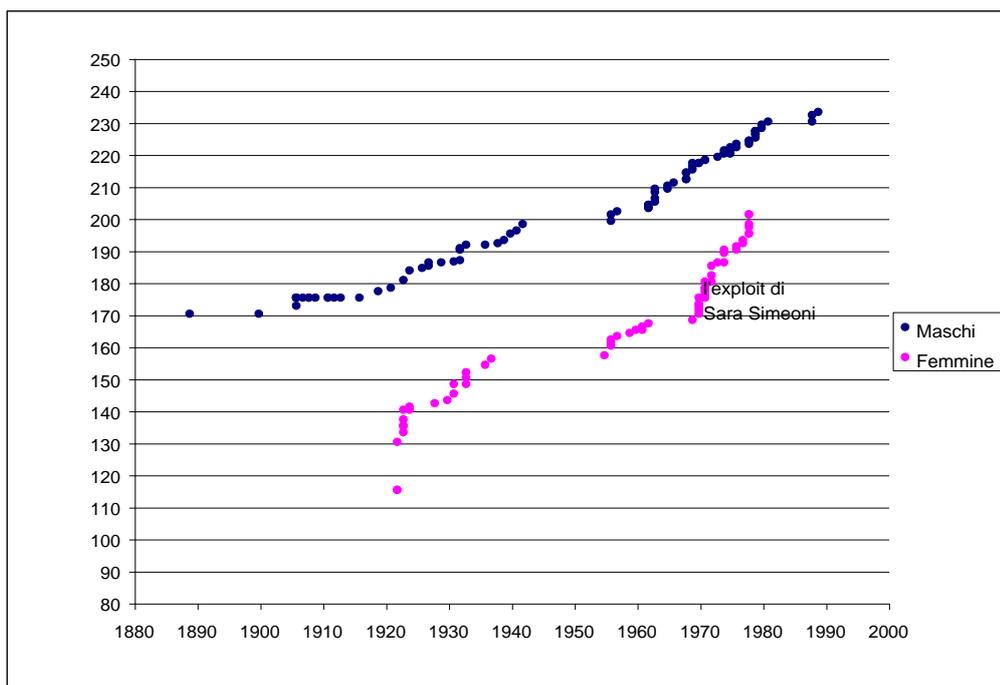
(a) Minuti, secondi e decimi (dal 1973 centesimi) di secondo.

Quanti numeri! Difficile raccapezzarcisi! Pure, contengono numerose informazioni. Cerchiamo di scoprirne qualcuna insieme, con l'aiuto di semplici strumenti statistici: le rappresentazioni grafiche e i rapporti.

Sui dati della tavola 2 occorre fare qualche precisazione: nel 1956 i primati, che sono espressi in tempo misurato in minuti primi, minuti secondi e frazioni di minuto secondo peggiorano, seppure di poco. Da quella data, infatti, è stato autorizzato il più preciso cronometraggio elettronico. Dal 1973, inoltre, sono stati cronometrati anche i centesimi di secondo.

Abbiamo rappresentato i dati delle tavole 1 e 2 nelle figure 1 e 2. Trattandosi di serie temporali, abbiamo utilizzato un grafico cartesiano. Sulle ascisse è descritto il tempo, sulle ordinate la misura del primato. La tendenza è, ovviamente, ad aumentare per il salto in alto e a diminuire per i 100 metri stile libero.

**Figura 1 – Primati italiani, maschili e femminili, per il salto in alto, nel tempo (in centimetri)**



**Figura 2 – Primati italiani, maschili e femminili, per i 100 metri stile libero, nel tempo (in minuti, secondi, frazioni di secondo)**



Ecco alcune prime considerazioni per il salto in alto:

- l'evoluzione del record maschile è più regolare di quello femminile;
- il susseguirsi temporale dei primati presenta alcuni salti. Ciò può derivare dall'emergere di qualche atleta particolarmente dotato (si potrebbero controllare i miglioramenti dei primati personali, per vedere se questo è vero) oppure da una modificazione radicale della tecnica: ad esempio, il passaggio dallo scavalcamento ventrale iniziato da Albritton allo stile Fosbury, scavalcamento dorsale, introdotto intorno al 1968;
- il divario fra record maschile e femminile tende a diminuire nel tempo; a ciò ha contribuito fortemente l'exploit di Sara Simeoni che, fra il 1970 e il 1978, ha portato il primato italiano del salto in alto femminile da 1,71 a 2,01 (+ 30 cm!).

Per dare evidenza alle ultime due considerazioni, nella tavola 3, per alcuni anni, fra il 1923 e il 1981, è stata calcolata la differenza tra i record italiani maschili e femminili. Questa è rimasta intorno a 40-43 cm fino al 1970; poi si è ridotta sotto i 30. Un apprezzamento analogo può essere fatto in base al rapporto fra le due misure (giudicate voi quale metodo sia migliore). Si vede che il primato femminile è stato circa l'80 per cento di quello maschile fino al 1970 ed è poi passato all'86 per cento nel periodo 1975-1981. Il merito, lo si può constatare dai dati della tavola 1, è stato specialmente, come si è già accennato, di Sara Simeoni.

D'altra parte, il recupero delle donne rispetto agli uomini è abbastanza generale, per tutte le specialità sportive. Esso ha avuto luogo negli ultimi 40 anni in Italia. Potete verificarlo voi stessi, chiedendo i dati alla Fidal, o ricorrendo a qualche pubblicazione specializzata.

Il processo di avvicinamento delle performance femminili a quelle maschili è dovuto essenzialmente alla più diffusa pratica sportiva rispetto al passato da parte delle donne. Storicamente, la discriminazione secondo il sesso si è espressa anche nella diversa opportunità di praticare gli sport. Appena qualche decina di anni fa, i genitori non si aspettavano che la propria figlia corresse veloce o che lanciaresse lontano una palla di ferro; così come non si aspettavano che riuscisse bene in scienze.

Chiedete alla Fidal i dati sui tesseramenti maschili e femminili, per specialità, nel tempo e confrontateli fra loro; potete anche rapportarli, rispettivamente, ai dati relativi alle popolazioni maschile e femminile fra i 15 e i 25 anni, alle date dei censimenti più vicini<sup>1</sup>.

**Tavola 3 - Primati italiani maschili e femminili per il salto in alto, in alcuni anni tra il 1925 e il 1980 e confronti (in metri)**

ANNI	Primato (a)		Differenza (1) - (2)	Rapporto (2) : (1) %
	Maschile (1)	Femminile (2)		
1925	1,835	1,410	0,43	76,8
1930	1,860	1,430	0,43	76,9
1935	1,915	1,517	0,39	79,2
1940	1,950 1,560		0,39	80,0
1955	2,010(b)	1,570	0,44	78,1
1960	2,020	1,650	0,37	81,7
1965	2,100	1,670	0,43	79,5
1970	2,170	1,750	0,42	80,6
1975	2,220	1,900	0,32	85,6
1980	2,290	2,100	0,19	91,7

(a) Si considera il primato più elevato stabilito nell'anno.

(b) Relativo al 1956.

### Rapporto di derivazione

Quoziente che si ottiene dal rapporto fra l'intensità di un certo fenomeno e l'intensità di un altro che ne costituisca il presupposto. Qualche volta, il numeratore del rapporto rappresenta l'ammontare della variazione intervenuta, in un certo intervallo di tempo, per il fenomeno che figura al denominatore.

Sono esempi di rapporti di derivazione: i quozienti di natalità (numero di nati in un anno diviso la popolazione all'inizio dell'anno o a metà anno) che forniscono una misura della forza riproduttiva di una popolazione; i quozienti di scolarità (numero di alunni iscritti in un anno scolastico diviso la popolazione; eventualmente, soltanto quella di età corrispondente); i tassi di pratica sportiva della popolazione giovanile.

Guardate ora i risultati delle donne nei 100 metri stile libero (Tavola 2 e specialmente Figura 2) fra il 1924 e il 2001. Il loro record è migliorato di 40 secondi e 33 centesimi (-42,3 per cento). Più o meno nello stesso periodo, fra il 1926 e il 2000, il record maschile si è abbassato di 16 secondi e 37 centesimi (-25,0 per cento). Abbiamo usato, per il confronto, una variazione relativa, o tasso di variazione, che è stato percentualizzato per poter apprezzare più facilmente il risultato. Ad esempio, per le donne, abbiamo sottratto il record del 1924 da quello del 2001 (cioè :  $55,07 - 95,40 = -40,33$  secondi)<sup>2</sup>; abbiamo rapportato il valore della differenza a quello del record iniziale ( $-40,33 \div 95,40 = -0,423$ ) e infine abbiamo moltiplicato per 100 ( $-0,423 \times 100 = -42,3$ ). Noterete che la prima differenza è espressa in secondi, come i record. I due rapporti, invece, non hanno più dimensione (sono numeri puri), poiché esprimono quella stessa differenza in funzione del valore iniziale del record, assumendo quest'ultimo come una vera e propria unità di misura; ad esempio, -42,3 per cento significa che l'ultimo record è diminuito, rispetto a quello preso a riferimento – cioè come unità di misura – di una frazione pari a 423 millesimi, cioè del 42,3 per cento.

<sup>1</sup> Il valore di questi rapporti vi dirà quanti uomini (o donne) praticavano le diverse specialità dell'atletica leggera per ogni coetaneo dello stesso sesso che avrebbe potuto praticarle. Se moltiplicate il valore per 100, avrete quanti le praticavano rispetto a 100 che avrebbero potuto. Quelli qui proposti sono rapporti di derivazione.

<sup>2</sup> Per poter effettuare la sottrazione ambedue i record sono stati espressi in secondi e centesimi di secondo.

Possiamo costruire una tavola analoga alla tavola 3. I dati che si ottengono sono descritti nella tavola 4. La distanza fra i record maschili e femminili era di circa 12 secondi nel 1955 e diventa di circa 7 secondi nel 2000, riducendosi dal 16,81 per cento al 13,67 per cento.

**Tavola 4 - Primati italiani maschili e femminili di nuoto nei 100 metri stile libero, in alcuni anni tra il 1950 e il 2000) e confronti (in secondi e centesimi di secondi)**

ANNI	Primato (a)		Differenza (2) - (1)	Rapporto (2) : (1) %
	Maschile (1)	Femminile (2)		
1950	59,50	69,50	10,00	116,81
1955	57,20	68,80	11,60	120,28
1960	56,10	64,40	8,30	114,80
1965	55,40	62,50	7,10	112,82
1970	54,60	62,20	7,60	113,92
1975	52,50	61,36	8,86	116,88
1980	51,24	58,62	7,38	114,40
1985	50,99	57,24	6,25	112,26
1990	49,48	57,08	7,60	115,36
2000	49,23	55,96	6,73	113,67

(a) Si considera il primato più elevato stabilito nell'anno.

### Tasso di variazione

I tassi di variazione sono molto utili allorché si vogliono confrontare due situazioni. Supponiamo, ad esempio, di voler valutare l'effetto del rincaro del prezzo del petrolio (avvenuto nel 1981) sul consumo di benzina per autoveicoli in Italia (che importa tutto il suo fabbisogno) e negli Stati Uniti d'America (che, invece, sono produttori). Potremmo considerare la variazione del consumo nei due paesi (differenza fra i consumi nel 1982 rispetto al 1981), ma i risultati che otterremmo sarebbero fortemente influenzati dal differente numero di autoveicoli circolanti nei due paesi e anche dalla loro diversa cilindrata media; in altri termini, dal diverso ordine di grandezza dei consumi nei due paesi. Per eliminare l'effetto di questa circostanza, possiamo rapportare la variazione intervenuta fra i due anni al valore iniziale del consumo, eventualmente moltiplicando per 100 il risultato ottenuto. Se il valore del rapporto per l'Italia fosse -0,22 (in termini percentuali -22 per cento) e per gli Usa -0,05 (in termini percentuali -5 per cento) vorrebbe dire che il consumo in Italia, fra il 1981 e il 1982, è diminuito del 22 per cento rispetto al livello iniziale e negli Usa del 5 per cento.

In simboli, il tasso di variazione ( $t$ ) può essere così indicato:

$$t = \frac{I_2 - I_1}{I_1},$$

dove  $I_1$  e  $I_2$  indicano l'ammontare del fenomeno rispettivamente ai tempi 1 e 2.

Tutto ciò si capisce ancora più chiaramente, considerando l'evolversi nel tempo del valore del rapporto fra il record femminile e maschile. Nel 1955, la donna più veloce impiegava circa un quinto in più del tempo del primatista maschile per nuotare i 100 metri stile libero; fra il 1960 e il 2000, circa un settimo in più del tempo, cioè fra il 14 per cento e il 13 per cento in più. Negli ultimi anni sembra che il rapporto tra le performance maschili e femminili si sia stabilizzato.

Il divario fra uomini e donne non è uguale nei vari sport. In alcuni di essi, la conformazione fisica delle donne penalizza maggiormente i loro risultati. Gli esperti sostengono che la minore potenza muscolare femminile incida relativamente di più nelle specialità in cui conti lo scatto; meno quando occorra resistenza. Per verificarlo, abbiamo raccolto i dati dei primati maschili e femminili, per alcuni anni fra il 1980 e il 2000, oltre che per i 100 metri stile libero, anche per i 200, i 400, gli 800 e i 1500.

**Tavola 5 - Confronto tra i primati italiani di nuoto maschili e femminili nei 100, 200, 400, 800, e 1500 metri stile libero, in alcuni anni fra il 1980 e il 2000 (a)**

ANNI	Primati femminili in percentuale di quelli maschili				
	100 mt.	200 mt.	400 mt.	800 mt.	1500 mt.
1980	114,4	111,3	108,4	104,6	104,7
1983	114,6	110,9	109,4	105,8	-
1985	113,7	111,7	109,2	105,8	106,5
1987	113,0	110,4	107,7	-	-
1989	115,7	-	-	106,6	-
2000	113,7	112,6	-	-	-

(a) Si considera il primato più elevato stabilito nell'anno. Il simbolo "-" indica che il confronto non è possibile

Questa volta, se vogliamo confrontare l'evoluzione dei risultati femminili rispetto a quelli maschili nelle diverse specialità, non possiamo basarci sulle differenze fra i record, ma dobbiamo ricorrere ai rapporti fra gli stessi. Infatti, le differenze risulterebbero dell'ordine di grandezza dei record. Ad esempio nel 1985, per i 100 metri stile libero maschili, il record è di 51 secondi e per i 1500 maschili è di 15 minuti e mezzo. I rapporti, invece, risultano adimensionali e consentono un corretto confronto fra le situazioni considerate. Abbiamo scritto il valore dei rapporti fra primati femminili e maschili nella tavola 5. Si vede immediatamente che i primati femminili degli 800 e 1500 metri quasi eguagliano quelli maschili. C'è ancora, forse, da erodere qualcosa sulle distanze brevi (100, 200 e 400 metri).

Emerge con chiarezza, comunque, che le donne ottengono risultati comparativamente migliori nelle gare di resistenza, come afferma la medicina sportiva.

## Esercitazioni

1. Ripetete l'indagine proposta in questo capitolo per altri sport che vi interessano, eventualmente soltanto per la vostra regione o provincia. In particolare, provate a verificare, per la corsa, se le donne conseguano risultati relativamente migliori, in confronto agli uomini, nelle gare di resistenza rispetto a quelle di velocità.
2. Organizzate una gara di salto in alto con i compagni della vostra e di altre classi, tenendo distinte le misure dei maschi da quelle delle femmine. Effettuate ciascuno tre prove e considerate la migliore. Rapportate questa misura alla statura di chi l'ha conseguita. Il valore del rapporto darà un'idea dell'agilità individuale (quale frazione della propria statura si riesce a saltare). Lo stesso rapporto potete stabilirlo fra la somma di tutti i risultati ottenuti nel salto in alto e la somma dei valori delle stature, ottenendo una misura dell'agilità media dell'intero gruppo. Se aggregate per sottogruppi (i maschi; le femmine; i dodicenni; i tredicenni; gli alti; i bassi, avendo stabilito, per questi ultimi due gruppi, il valore di separazione), potete fare interessanti confronti.
3. Approfondite la questione della discriminazione secondo il sesso, sperimentata storicamente nel nostro come in altri paesi, in relazione a diversi aspetti del vivere sociale. Consultate l'Annuario statistico italiano; il rapporto tra femmine e maschi nel nostro paese è risultato, al censimento del 2001, pari a 1,06: in altri termini, ogni 100 maschi viventi, c'erano circa 106 femmine. Quanti erano i deputati di sesso femminile rispetto a 100 colleghi maschi? E gli insegnanti universitari? E i chirurghi? E ...
4. I tempi vengono misurati nei vari sport anche in centesimi di minuto secondo. C'è spazio per riflessioni sull'evoluzione degli strumenti di misura del tempo e sulle approssimazioni che si commettono nella lettura dei valori rilevati. L'insegnante di matematica e quello di scienze vi possono aiutare a confrontare l'espressione decimale e sessagesimale che si usa per esprimere il tempo. Provate anche a valutare le misure decimali o centesimali mediante frazioni semplici ( $1/2$ ;  $1/3$ ;  $1/4$ ) di secondo. Per esempio, 27 centesimi sono di più o di meno di  $1/3$  di secondo? E di  $1/4$  di secondo?

## 4. Quanti eravamo, quanti siamo, quanti saremo

### Che cosa si impara nel capitolo 4

Per conoscere le caratteristiche e l'evoluzione della popolazione italiana attraverso un lungo arco di tempo utilizziamo il **tasso di incremento medio annuo composto**, il **tasso di natalità**, il **tasso di mortalità**, **l'indice di vecchiaia**, il **tasso di incremento naturale della popolazione**, i **rapporti di coesistenza**.

Tu fai parte dei 56.996.000 residenti in Italia a ottobre del 2001. Ne fanno parte anche i tuoi fratelli, se ne hai, e i tuoi genitori. Ti sei mai chiesto quanta sia la gente che vive nella tua regione, nella tua città? Forse no. Eppure, dal suo ammontare dipendono molti altri fenomeni, perché le persone producono, consumano, utilizzano servizi pubblici (ad esempio, come te, le scuole).

Se leggi queste righe nell'anno scolastico 2003-2004 sarai nato, quasi certamente, fra il 1988 e il 1991. Bene, nel 1988 sono nati in tutto, nel nostro Paese, 579.055 bambini. Nei tre anni successivi, rispettivamente, 569.242, 579.343, 571.324. Più recentemente, nel 1999 ne sono nati 537.242 e nel 2000 543.039. Venti anni fa nel 1983, sono state registrate 614.006 nascite. Quaranta anni fa, nel 1963, i nati erano stati 978.110<sup>3</sup>. Le cose sono cambiate parecchio!

Cerchiamo di saperne di più. Cominciamo con la popolazione: dall'unità d'Italia a oggi, è passata da 22.176.000 a 56.996.000 unità (Tavola 1).

**Tavola 1 - Popolazione italiana residente alle date dei censimenti generali, riportata ai confini attuali - Anni 1861-2001 (migliaia di unità)**

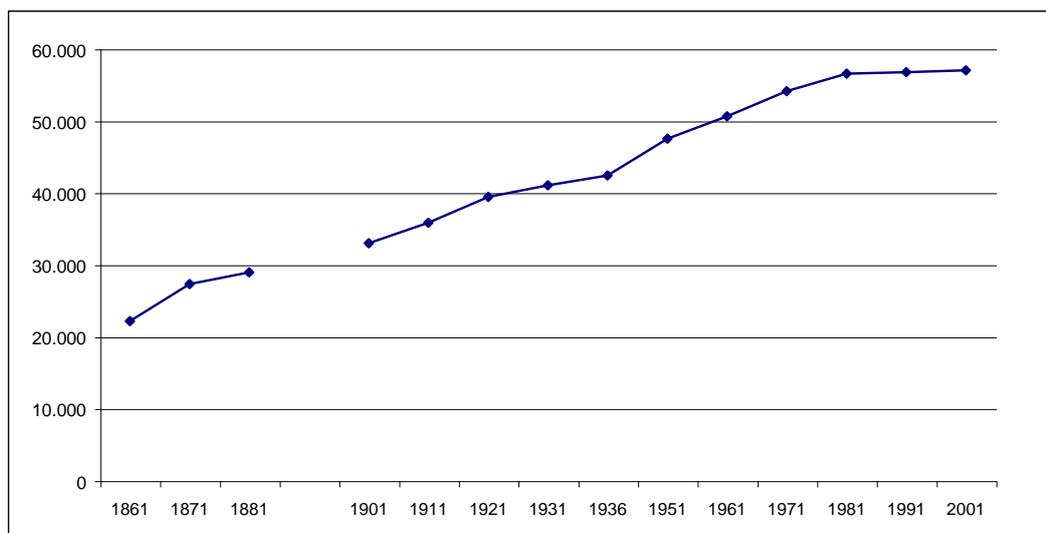
ANNI	Popolazione residente
1861	22.176
1871	27.300
1881	28.952
1901	32.963
1911	35.842
1921	39.397
1931	41.043
1936	42.398
1951	47.516
1961	50.624
1971	54.137
1981	56.557
1991	56.778

Fonte: Istituto nazionale di statistica. Censimenti della popolazione.  
<http://www.istat.it/Censimenti/index.htm>

I dati possono essere rappresentati graficamente. Si può usare un diagramma cartesiano, descrivendo sulle ascisse gli anni e sulle ordinate l'ammontare della popolazione (Figura 1).

<sup>3</sup> Fonte: Istituto nazionale di statistica, 2002. Database "Health for all". Per i nati del 1963: Istituto nazionale di statistica, 1964. Popolazione e circoscrizioni amministrative dei comuni

Figura 1 – Evoluzione della popolazione italiana. Anni 1861-2001 (migliaia di unità)



La popolazione italiana è costantemente cresciuta dall'unità a oggi; però, nell'ultimo periodo, sempre più lentamente. Possiamo capirlo meglio se consideriamo la variazione relativa che si è avuta in media ogni anno negli intervalli fra le rilevazioni censuarie. La variazione relativa è data dal rapporto fra la variazione assoluta, differenza fra l'ammontare della popolazione alla fine del periodo considerato e quello all'inizio, e l'ammontare iniziale. Se il periodo è pluriennale, la variazione relativa media annua (o tasso medio annuo di variazione) si ottiene dividendo ancora per il numero di anni. Poiché, nel nostro caso, le variazioni sono state sempre in aumento, si parlerà di tasso di incremento medio annuo che qualifichiamo come semplice. In simboli:

$$\frac{P_{t+n} - P_t}{P_t} : \left( \frac{1}{n} \right) = \frac{P_{t+n} - P_t}{P_t} : n = \left( \frac{P_{t+n}}{P_t} - 1 \right) : n.$$

In generale, il tasso viene moltiplicato per 100 o per 1.000, esprimendolo perciò con riferimento a 100 o a 1.000 abitanti.

Ad esempio, per il decennio 1991-2001, otteniamo:

$$\text{tasso di incremento medio annuo semplice} = \left( \frac{56996}{56778} - 1 \right) : 10 = 0,00384 \times 1000 = 0,384 \text{ unità per ogni 1.000 abitanti.}$$

Quando si ha a che fare con fenomeni di accumulo, come la popolazione, si preferisce calcolare il tasso di incremento medio annuo composto, attraverso la relazione:

$$\sqrt[n]{\frac{P_{t+n}}{P_t}} - 1 = \sqrt[n]{\frac{P_{t+n}}{P_t}} - 1.$$

I tassi di incremento medio annuo composto calcolati sui dati della tavola 1 sono esposti nella tavola 2. Per il decennio 1991-2001, il tasso risulta, ad esempio:

$$\text{tasso di incremento medio annuo composto} = \sqrt[10]{\frac{56996}{56778}} - 1 = 0,00383 \times 1000 = 0,383 \text{ unità per ogni 1.000 abitanti.}$$

In definitiva, la popolazione è aumentata fra il 1991 e il 2001 di circa 0,4 unità per ogni 1.000 abitanti, ossia circa 4 per ogni 10.000 abitanti, in media per ciascun anno.

Il riferimento potrebbe essere fatto a 100, a 1.000, a 2.500, a 5.000 o a 10.000 abitanti (sono naturalmente più comodi i riferimenti alle potenze di dieci), così come la velocità di un autoveicolo, generalmente espressa in chilometri per ora (per esempio 85 km/h), potrebbe esserlo in chilometri per minuto  $(85 \div 60 = 1,4167 \text{ km/min})$  e anche in metri per secondo  $(5.000 \div (60 \times 60) = 23,6 \text{ m/sec})$ .

Per il decennio considerato, dunque, il tasso semplice e quello composto praticamente coincidono. Per il decennio 1861-1871, invece, i due tassi risultano diversi e pari, rispettivamente, a 23,1 per mille e 21,0 per mille. La popolazione si accresceva in media ogni anno di oltre 20 unità ogni 1.000 abitanti, cioè più del 2 per cento: un aumento molto consistente.

**Tavola 2 - Tassi di incremento medio annuo composto della popolazione italiana fra le date dei censimenti - Anni 1861-2001 (per 1.000 abitanti all'inizio dei periodi considerati)**

PERIODI	Tasso di incremento	PERIODI	Tasso di incremento
1861-1871	21,0	1936-1951	7,4
1871-1881	5,9	1951-1961	6,4
1881-1901	6,8	1961-1971	6,7
1901-1911	8,1	1971-1981	4,4
1911-1921	9,1	1981-1991	0,4
1921-1931	4,4	1991-2001	0,4
1931-1936	6,5		

Fonte: Istituto nazionale di statistica. Censimenti della popolazione.  
<http://www.istat.it/Censimenti/index.htm>

È stato utile calcolare il tasso medio annuo, poiché i periodi considerati sono di ampiezza differente: in genere di dieci anni, ma anche di cinque, quindici e venti. Se non avessimo calcolato la media annua, la valutazione dell'evoluzione della popolazione nel tempo ne sarebbe risultata falsata.

### Tasso di incremento medio annuo della popolazione

Il tasso di incremento medio annuo (composto) della popolazione, che possiamo indicare con  $k$ , risulta, in simboli:

$$k = \left( \sqrt[t+n-t]{\frac{P_{t+n}}{P_t}} - 1 \right) \times 1000 = \left( \sqrt[n]{\frac{P_{t+n}}{P_t}} - 1 \right) \times 1000,$$

dove  $P_{t+n}$  è la popolazione nell'anno di calendario finale ( $t+n$ ) del periodo considerato,  $P_t$  la popolazione nell'anno iniziale del periodo ( $t$ ) e  $n$  è appunto il numero di anni che sono trascorsi.

Come siamo arrivati al tasso di incremento precedente? Proviamo a seguire il percorso che ci ha portato a quel risultato.

Quando abbiamo a che fare con dati che evolvono nel tempo, calcoliamo in generale la variazione, cioè la differenza fra il dato registrato alla fine di un periodo di osservazione e quello registrato all'inizio:

$$P_{t+n} - P_t.$$

La differenza può essere positiva o negativa, secondo che il fenomeno sia aumentato oppure diminuito tra il tempo  $t$  e il tempo  $t+n$ . Nel primo caso, la variazione viene detta anche incremento; nel secondo caso, decremento.

Se vogliamo fare confronti fra l'andamento di due o più fenomeni in uno stesso periodo di tempo o tra gli andamenti dello stesso fenomeno in differenti periodi, può essere conveniente riportare la variazione al valore iniziale, ottenendo la variazione relativa:

$$\frac{P_{t+n} - P_t}{P_t}.$$

Facendo in questo modo, abbiamo eliminato l'influenza dell'unità di misura o dell'ordine di grandezza nei quali il fenomeno sotto osservazione è espresso, e il confronto con altri fenomeni o situazioni, trattati nello stesso modo e per i quali esso abbia significato, può essere stabilito correttamente. In generale, la variazione relativa così ottenuta, che viene definita tasso di variazione, è moltiplicata per 100 o per 1.000. Se si tratta della popolazione, il tasso di variazione sarà quindi riferito a 100 o 1.000 abitanti.

In qualche caso le variazioni, assolute o relative che siano, corrispondono a periodi di tempo di differente durata: talvolta un anno, altre volte un quinquennio o un decennio. In questi casi, le variazioni, assolute o relative, non sono più confrontabili tra loro, se non riferendole tutte a una stessa durata temporale, ad esempio a un anno. È quello che normalmente si fa per dati di popolazione. Se  $t$  è un anno di calendario e  $t+n$  è pure un anno di calendario  $n$  anni dopo il precedente, possiamo calcolare il tasso di variazione medio annuo, dividendolo per  $t+n-t=n$ . Questo tasso si definisce semplice poiché, per ottenerlo, si calcola la media aritmetica semplice dei tassi relativi ai singoli anni, tutti riferiti alla popolazione all'inizio del periodo, cioè:

$$\left[ \frac{P_{t+1} - P_t}{P_t} + \frac{P_{t+2} - P_{t+1}}{P_t} + \dots + \frac{P_{t+n} - P_{t+n-1}}{P_t} \right] : n.$$

La precedente relazione è chiaramente equivalente a:

$$\frac{P_{t+n} - P_t}{P_t} : n.$$

Tuttavia, nel caso della popolazione e di altri fenomeni di accumulo, come ad esempio una somma investita in banca la quale frutta ogni anno degli interessi che si aggiungono al capitale iniziale e fruttano essi stessi interessi successivamente, viene utilizzato più frequentemente il tasso medio annuo composto. Si prende spunto proprio da quanto si fa in matematica finanziaria per il tasso di interesse. Vediamo come.

Se investiamo una somma  $S_0$  all'inizio di un certo anno, diciamo appunto anno zero, a un tasso di interesse  $i_0$ , alla fine dell'anno essa avrà fruttato un interesse pari a  $S_0 \times i_0$ . Se non ritiriamo l'interesse maturato e lo lasciamo a frutto, la somma all'inizio dell'anno 1 sarà ora pari a:

$$S_1 = S_0 + S_0 \times i_0 = S_0(1+i_0)$$

e frutterà alla fine dell'anno un interesse pari a  $S_1 \times i_1$ . Quindi, alla fine dell'anno, disporremo di una somma pari a:

$$S_1 + S_1 \times i_1 = S_1(1+i_1) = S_0(1+i_0)(1+i_1).$$

Se investiamo in questo modo per  $n$  anni, a tassi di interesse che variano di anno in anno da quelli già considerati  $i_0$  e  $i_1$  a  $i_2$  ..... fino a  $i_n$ , al termine del periodo avremo accumulato una somma pari a:

$$S_n = S_0(1+i_0)(1+i_1)\dots(1+i_n).$$

Ci possiamo ora chiedere a quale tasso medio annuo di interesse abbiamo investito, nell'intero periodo di  $n$  anni considerato, la somma iniziale. Chiamiamo questo tasso medio annuo semplicemente  $i$ . Poiché vale l'uguaglianza:

$$S_n = S_0(1+i_0)(1+i_1)\dots(1+i_n) = S_0(1+i)(1+i)\dots(1+i) = S_0(1+i)^n,$$

si può calcolare con semplicità:

$$i = \left( \sqrt[n]{\frac{S_n}{S_0}} - 1 \right).$$

Il tasso  $i$  può essere moltiplicato per 100, e quindi espresso con riferimento a 100 lire investite.

Trasferiamo ora il risultato precedente al problema del calcolo del tasso medio annuo composto di incremento (in generale di variazione) della popolazione fra due date. Vale la seguente analoga uguaglianza:

$$P_{t+n} = P_t(1+k_t)(1+k_{t+1})\dots(1+k_{t+n}) = P_t(1+k)(1+k)\dots(1+k) = P_t(1+k)^n,$$

dove  $t$  è l'anno di calendario iniziale (per la somma investita l'avevamo chiamato anno 0) e  $t+n$  quello finale del periodo di osservazione (per la somma investita, anno  $n$ );  $k_{t+i}$  ( $i=0,1,\dots,n$ ) è il tasso di incremento che si registra nell'anno  $i$ -esimo;  $k$  è appunto il tasso medio annuo ricercato. È ora facile calcolare il tasso medio annuo composto di incremento (in generale di variazione) della popolazione:

$$k = \left( \sqrt[t+n-t]{\frac{P_{t+n}}{P_t}} - 1 \right) = \left( \sqrt[n]{\frac{P_{t+n}}{P_t}} - 1 \right),$$

che eventualmente possiamo moltiplicare per 1.000, esprimendolo così per 1.000 abitanti.

Svolgiamo ora qualche considerazione sui risultati che abbiamo ottenuto:

nel periodo compreso fra il 1861 e il 1871, la popolazione è aumentata in media di 21 abitanti ogni 1.000, per ciascun anno;

nel decennio 1911-1921 si è registrato un incremento medio annuo del 9,1 per mille, poco meno dell'1 per cento;

nel decennio successivo 1921-1931, esso è stato del 4,4 per mille;

nel quinquennio 1931-1936, la campagna a favore dell'incremento demografico (fatti raccontare, da tua nonna o dall'insegnante di storia, del premio che veniva dato alle madri che avevano un quinto figlio) fece aumentare il tasso al 6,5 per mille;

dal 1971, inizia una rapida caduta del tasso di incremento: fra il 1971 e il 1981 la popolazione si accresce in media, per ciascun anno, di poco più di 4,4 unità ogni 1.000 abitanti; nel decennio 1981-1991, di 0,4 unità ogni 1.000 abitanti (oltre dieci volte di meno) e lo stesso incremento si registra nel decennio successivo, 1991-2001.

Da che cosa dipende se la popolazione aumenta o diminuisce fra un anno e l'altro? Proviamo a riflettere. Dipende dal numero di coloro che nascono e dal numero di coloro che muoiono. I primi la faranno aumentare; i secondi diminuire. Dipende anche dal numero di coloro che si spostano all'estero per periodi così lunghi da giustificare la loro cancellazione dai registri anagrafici dei comuni in cui risiedono e di coloro che, invece, rientrano dall'estero o, immigrati da altri paesi, sono presenti non occasionalmente oppure ottengono la residenza nel nostro. Trascuriamo questa seconda componente e concentriamo l'attenzione sui nati e sui morti. Fino a quando i primi superano i secondi, la popolazione cresce; nel caso contrario, diminuisce. La differenza fra nascite ( $N$ ) e morti ( $M$ ) in un dato anno (periodo) ci dice di quanto aumenta (se  $N$  è maggiore di  $M$ ) o diminuisce (se  $N$  è minore di  $M$ ) la popolazione fra l'inizio e la fine dell'anno (periodo) considerato.

Se dividiamo le nascite (i nati vivi) e le morti per la popolazione all'inizio dell'anno, otteniamo due tassi, che converrà moltiplicare per mille per aderire all'uso corrente dei demografi (coloro che studiano la popolazione). Il primo rapporto si chiama *tasso di natalità*:

$$n = \frac{N}{P} \times 1.000$$

e il secondo *tasso di mortalità*<sup>4</sup>:

$$m = \frac{M}{P} \times 1.000.$$

La loro differenza:

$$s = n - m = \frac{N}{P} \times 1.000 - \frac{M}{P} \times 1.000 = \frac{N - M}{P} \times 1.000$$

viene detta *tasso di incremento naturale della popolazione* (non soltanto tasso di incremento, perché non è considerato il saldo delle migrazioni con l'estero).

Qual è il significato di  $s$ ,  $n$  e  $m$ ? Il tasso  $s$  ci dice di quanto è aumentata la popolazione nell'anno (periodo) preso in considerazione ogni mille abitanti all'inizio dell'anno (periodo). Se questi erano un milione e  $s$  è risultato pari a quattro per mille, alla fine dell'anno (periodo) avremo quindi una popolazione di

$$1.000.000 + 1.000.000 \times \frac{4}{1.000} = 1.004.000.$$

Analogamente,  $n$  ci dice quanti bambini sono nati vivi nell'anno ogni mille abitanti iniziali e  $m$  quante persone sono morte nell'anno sempre ogni mille abitanti iniziali.

Nella tavola 3 sono scritti i tassi di natalità, di mortalità e di incremento naturale per mille abitanti, registrati in Italia fra il 1951 e il 2000. Nel grafico della figura 2 sono illustrati, con evidenza annuale, gli andamenti dei due tassi componenti e della loro differenza.

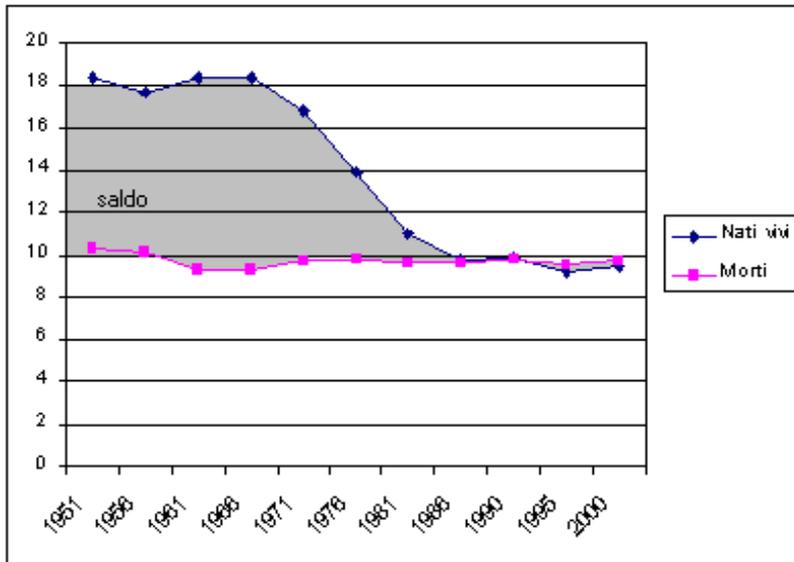
**Tavola 3 - Nati vivi, morti e saldo naturale della popolazione italiana in alcuni anni fra il 1951 e il 2000 (per 1.000 abitanti)**

ANNI	Nati vivi	Morti	Differenza: saldo naturale
1951	18,4	10,3	8,1
1956	17,7	10,1	7,6
1961	18,4	9,3	9,1
1966	18,4	9,3	9,1
1971	16,8	9,7	7,1
1976	13,9	9,8	4,1
1981	11,0	9,6	1,4
1986	9,8	9,6	0,3
1990	9,9	9,8	0,1
1995	9,2	9,5	-0,3
2000	9,4	9,7	-0,3

Fonte (per gli anni dal 1955 al 1999): Istituto nazionale di statistica. Annuario statistico italiano; (per l'anno 2000) Istituto nazionale di statistica. Popolazione e statistiche demografiche. Indicatori strutturali 2001. <http://demo.istat.it>

<sup>4</sup> Si tratta, in ambedue i casi, di rapporti di derivazione.

Figura 2 - Nati vivi, morti e saldo naturale della popolazione italiana - Anni 1951 – 2000 (per 1.000 abitanti)



Commentiamo brevemente i dati.

Dal 1951 al 2000 i nati sono diminuiti da poco meno di 2 ogni 100 abitanti (18,4 per mille) a 0,9 per 100 (9,4 per mille). È diminuita anche la mortalità: da 10,3 morti per 1.000 abitanti nel 1951 a 9,7 per 1.000 abitanti nel 2000. L'andamento dei due tassi fa sì che, da saldi naturali positivi di 7-9 per 1.000 abitanti (cioè ogni anno la popolazione aumentava quasi dell'1 per cento), si sia passati a meno di 1 ogni 1.000 nel 1986, fino ad arrivare al valore negativo di -0,3 per 1.000 nel 1995. La caduta è risultata drastica dopo il 1970.

Il grafico della figura 2 è molto eloquente. Il tasso di incremento naturale (eccedenza dei nati sui morti ogni 1.000 abitanti) si assottiglia sempre più. Nel 1990 le due spezzate si sono incrociate e questo vuole dire che la popolazione ha cominciato a diminuire.

Si può fare un'altra considerazione. Se ci sono meno nati e meno morti, la popolazione, nel suo complesso, è destinata a invecchiare; poiché diventeranno proporzionalmente più numerosi gli anziani e meno numerosi i giovani. Vediamo con quale intensità ciò è avvenuto in Italia nel corso del tempo. Possiamo calcolare il rapporto fra l'ammontare della popolazione con 60 anni e più e quello della popolazione con meno di 15 anni, eventualmente moltiplicando il risultato per 100, per agevolarne l'interpretazione. Si tratta di un rapporto di coesistenza, che i demografi chiamano *indice di vecchiaia*:

$$r = \frac{\text{popolazione con 60 anni e più}}{\text{popolazione con meno di 15 anni}} \times 100.$$

### Rapporti di coesistenza

Il rapporto è utile se si vuole mettere in evidenza lo squilibrio fra la quantità posta al numeratore e quella posta al denominatore. Si pensi al rapporto fra il numero dei matrimoni celebrati in un anno con rito civile e rispettivamente religioso, eventualmente moltiplicato per 100. Esso ci dice quanti matrimoni del primo tipo sono stati celebrati ogni 100 del secondo tipo.

Al censimento del 1981, gli anziani risultavano 9.847.600 e i giovani 12.127.600, cosicché l'indice di vecchiaia valeva:

$$v = \frac{9.847.600}{12.127.600} \times 100 = 81,2.$$

Ciò significa che, per ogni 100 giovani (con meno di 15 anni), c'erano, a quella data, 81,2 anziani (con 60 anni e più).

La tavola 4 illustra l'andamento dell'indice di vecchiaia nel tempo; per gli anni 1901, 1951 e 1981 essi sono calcolati sui dati di censimento; per il 1991 e il 2001 sui dati registrati presso le anagrafi. La tavola illustra anche i valori regionali più elevati e più bassi.

**Tavola 4 - Indici di vecchiaia per la popolazione italiana (rapporto percentuale della popolazione con 65 anni e più e la popolazione di 0-14 anni) - Anni 1901, 1951, 1981, 1991, 2001**

DESCRIZIONE	1901	1951	1981	1991(a)	2001(b)
Italia in complesso	27,8	46,4	81,2	94,3	127,1
Valore regionale più elevato	34,5 (Abruzzo e Molise)	92,7 (Piemonte)	156,5 (Liguria)	200,0 (Liguria)	238,4 (Liguria)
Valore regionale più basso	21,7 (Sicilia)	29,3 (Calabria)	48,5 (Campania)	50,5 (Campania)	72,9 (Campania)

Fonte: Istat

(a) Istituto nazionale di statistica. Database Health for all. Un sistema informativo territoriale su sanità e salute.

Settembre 2002. <http://www.istat.it/Banche-dat/index.htm>.

(b) Istituto nazionale di statistica. Popolazione e statistiche demografiche. Indicatori strutturali 2001. <http://demo.istat.it>.

Quante informazioni in una tavola così piccola. Consideriamola insieme:

nel 1901 c'erano soltanto 28 anziani ogni 100 giovani; nel 1951 erano quasi 50 ogni 100; nel 1981 un po' più di 80 ogni 100, fino ad arrivare al 2001 allorché gli anziani risultano molti di più dei giovani;

nel 1901 c'era una forte omogeneità regionale; al massimo, in Abruzzo e Molise, si registravano 35 anziani per 100 giovani; al minimo, in Sicilia, 22 per cento;

nel 1951 c'era già una forte caratterizzazione regionale che si è accentuata sempre più in seguito e che ha visto la popolazione delle regioni settentrionali invecchiare assai più rapidamente rispetto a quella delle regioni meridionali; nel 1951, in Piemonte, il numero degli anziani era quasi uguale a quello dei giovani. Nel 1981, in Liguria, gli anziani erano oltre una volta e mezzo i giovani. Nel 2001, sempre in Liguria, gli anziani sono diventati più del doppio dei giovani.

I livelli minimi (registrati sempre nelle regioni meridionali durante tutto il periodo) sono risultati più stabili di quelli massimi. In effetti le modificazioni sociali e demografiche sono state più lente nel Mezzogiorno.

Fin qui, abbiamo illustrato dati. Vengono molte domande. È bene o male che diminuiscano le nascite? Quali conseguenze avrà il progressivo invecchiamento della popolazione? Provate a discuterne con i vostri compagni; con il vostro insegnante di storia; con i vostri fratelli più grandi e con i genitori.

## Esercitazioni

1. Provate a soddisfare altre curiosità riguardanti la popolazione: in Italia sono più i maschi o le femmine? Quanti sono i laureati ogni 100 abitanti con più di 22 anni? Quanti lavorano e cercano lavoro ogni 100 abitanti con 15 anni e più? Quest'ultimo rapporto, chiamato tasso di attività è uguale per maschi e femmine? Basterà sfogliare l'Annuario statistico italiano. Si può richiederlo alla più vicina sede regionale dell'Istat, oppure consultarlo on line sul sito dell'Istat ([www.istat.it](http://www.istat.it)).
2. Ecco, per le regioni italiane, la popolazione per classi d'età, nel 2001. Procedi a qualche confronto delle distribuzioni regionali secondo l'età, usando valori percentuali (poni il totale di ciascuna regione uguale a 100), in modo che la tua valutazione non sia disturbata dal diverso ammontare complessivo della popolazione delle varie regioni: puoi confrontare grossolanamente, anche senza questo accorgimento metodologico, le situazioni di Piemonte, Veneto ed Emilia-Romagna che hanno quasi la stessa popolazione complessiva.

**Tavola 2.2 - Popolazione residente per classe di età, regione e sesso al 1° gennaio 2001**

REGIONI	Meno di 1	1-4	5-9	10-14	15-24	25-44	45-64	65 e più	Totale
MASCHI E FEMMINE									
Piemonte	35.773	138.626	170.383	170.503	415.032	1.305.472	1.165.264	888.678	4.289.731
Valle d'Aosta	1.153	4.280	5.010	5.015	11.554	38.911	31.873	22.793	120.589
Lombardia	85.209	326.068	394.189	389.439	934.198	2.949.302	2.423.780	1.619.529	9.121.714
Trentino-Alto Adige	10.540	41.796	50.263	48.108	103.606	304.050	226.698	158.062	943.123
<i>Bolzano-Bozen</i>	<i>5.452</i>	<i>22.000</i>	<i>26.490</i>	<i>25.410</i>	<i>53.971</i>	<i>152.714</i>	<i>107.249</i>	<i>71.978</i>	<i>465.264</i>
<i>Trento</i>	<i>5.088</i>	<i>19.796</i>	<i>23.773</i>	<i>22.698</i>	<i>49.635</i>	<i>151.336</i>	<i>119.449</i>	<i>86.084</i>	<i>477.859</i>
Veneto	43.059	166.298	200.361	197.521	475.496	1.477.235	1.165.214	815.669	4.540.853
Friuli-Venezia Giulia	9.552	36.163	44.420	44.030	109.272	364.032	328.878	252.247	1.188.594
Liguria	11.200	44.767	57.064	56.827	132.806	469.023	444.311	405.018	1.621.016
Emilia-Romagna	33.896	126.327	151.773	146.673	362.773	1.246.472	1.053.039	887.710	4.008.663
Toscana	28.283	108.785	136.067	139.386	339.595	1.067.056	945.536	782.896	3.547.604
Umbria	6.774	26.305	34.003	35.706	86.958	245.991	216.973	187.772	840.482
Marche	12.471	49.057	63.135	65.214	156.693	436.606	370.086	315.933	1.469.195
Lazio	48.362	194.716	251.247	253.945	580.743	1.681.729	1.371.327	920.233	5.302.302
Abruzzo	10.938	44.163	60.963	64.864	152.113	381.973	309.647	256.622	1.281.283
Molise	2.621	11.239	15.953	17.332	39.950	95.324	76.428	68.330	327.177
Campania	68.081	274.150	370.591	387.930	837.596	1.769.656	1.271.571	802.669	5.782.244
Puglia	42.505	171.317	232.698	249.346	579.750	1.235.319	945.184	630.489	4.086.608
Basilicata	5.648	22.744	32.222	35.902	81.007	181.754	135.818	109.712	604.807
Calabria	19.680	81.558	116.787	131.643	290.504	607.695	454.299	341.122	2.043.288
Sicilia	52.859	217.636	304.920	323.528	691.379	1.496.915	1.155.206	834.257	5.076.700
Sardegna	13.764	54.688	77.293	88.599	220.605	528.193	408.708	256.194	1.648.044
<b>ITALIA</b>	<b>542.368</b>	<b>2.140.683</b>	<b>2.769.342</b>	<b>2.851.511</b>	<b>6.601.630</b>	<b>17.882.708</b>	<b>14.499.840</b>	<b>10.555.935</b>	<b>57.844.017</b>
<b>NORD</b>	<b>230.382</b>	<b>884.325</b>	<b>1.073.463</b>	<b>1.058.116</b>	<b>2.544.737</b>	<b>8.154.497</b>	<b>6.839.057</b>	<b>5.049.706</b>	<b>25.834.283</b>
<b>CENTRO</b>	<b>95.890</b>	<b>378.863</b>	<b>484.452</b>	<b>494.251</b>	<b>1.163.989</b>	<b>3.431.382</b>	<b>2.903.922</b>	<b>2.206.834</b>	<b>11.159.583</b>
<b>MEZZOGIORNO</b>	<b>216.096</b>	<b>877.495</b>	<b>1.211.427</b>	<b>1.299.144</b>	<b>2.892.904</b>	<b>6.296.829</b>	<b>4.756.861</b>	<b>3.299.395</b>	<b>20.850.151</b>

Fonte: Annuario statistico italiano 2002

Rifletti sulle diversità e sulle conseguenze:

- una popolazione mediamente più anziana pone rilevanti problemi di sanità e assistenza;
- se ci sono pochi giovani, chi sostituirà fra qualche decina di anni le classi non più abili al lavoro?

Discuti di questi problemi con i tuoi compagni di classe e insegnanti.

## 5. Il campionato di calcio italiano di serie A

### Che cosa si impara nel capitolo 5

Si può vedere come è cambiato il gioco del calcio - partite vinte/perse/pareggiate e goal fatti/subiti - attraverso la tabella a doppia entrata o di contingenza, il diagramma ad albero, la frequenza, la distribuzione di frequenza, la seriazione, la media aritmetica ponderata.

Il campionato di calcio italiano, oltre ad appassionare migliaia di tifosi, ha anche molti altri risvolti a cominciare da quello economico, importantissimo perché mette in moto una grande quantità di soldi: introiti delle partite (per la serie A nel 2002: 1 miliardo e 126 milioni di euro<sup>5</sup>), acquisto e gestione dei giocatori a costi, spesso, elevatissimi, il Totocalcio, trasmissioni radiotelevisive, sponsorizzazioni pubblicitarie, turismo, eccetera.

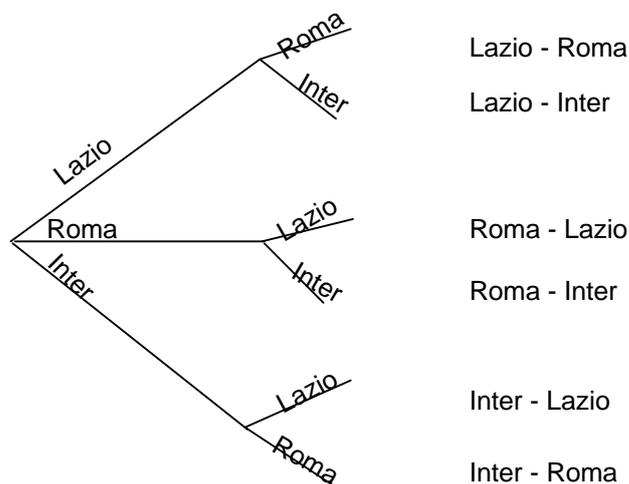
Vogliamo gettare uno sguardo ai campionati della serie A degli ultimi quarant'anni, iniziando dal 1961-62. Non prenderemo in considerazione tutte le annate, ma soltanto alcune: 1961-1962, 1971-1972; 1981-1982; 1991-1992, 2001-2002. E ancora, ci interesseremo soltanto ad alcuni aspetti per i quali è più semplice disporre della documentazione statistica.

Una prima questione riguarda l'esito delle partite nel loro complesso. Tutti sanno che nel campionato, ciascuna squadra incontra due volte tutte le altre squadre: per esempio, secondo i risultati del campionato 2001-2002:

Torino - Roma	0-1
Roma - Torino	1-0

La Roma vinse giocando sia in casa sia fuori casa. Il campionato di serie A attualmente viene disputato da 18 squadre e quindi abbiamo  $17 \times 18 = 306$  partite. Come abbiamo ottenuto questa cifra?

Costruiamo un albero, cioè un diagramma come il seguente; per capirlo e per semplicità della figura, pensiamo per un momento a un minicampionato in cui le squadre siano soltanto tre: Lazio, Roma, Inter. Allora:



dove sui primi tre rami (quelli di sinistra) abbiamo scritto i nomi di tutte le squadre del nostro immaginario campionato. A partire da ogni ramo facciamo sorgere una biforcazione poiché ogni squadra deve incontrare le altre due. I risultati possiamo leggerli sulla destra, sono in numero di  $3 \times 2 = 6$ , e includono tutti i possibili incontri in casa e fuori casa.

Ritorniamo ora al caso del campionato reale con 18 squadre: poiché ognuna deve incontrare le altre 17, allora - in analogia col minicampionato - il numero complessivo di gare

<sup>5</sup> Ufficio studi Federcalcio.

deve esser pari a  $18 \times 17 = 306$ . Si può anche stabilire direttamente, dal momento che ogni settimana si disputano 9 partite, che il campionato prevede  $306 \div 9 = 34$  giornate di gioco.

#### Quadro indice del campionato 2001-2002

CAMPIONATO DI SERIE A 2001-2002	ATALANTA	BOLOGNA	BRESCIA	CHIEVO	FIorentina	INTER	JUVENTUS	LAZIO	LECCE	MILAN	PARMA	PERUGIA	PIACENZA	ROMA	TORINO	UDINESE	VENEZIA	VERONA
ATALANTA		2-2	0-0	1-2	2-0	2-4	0-2	0-1	2-1	1-1	4-1	2-1	1-1	1-1	1-1	1-5	1-0	1-0
BOLOGNA	1-0		2-1	3-1	3-2	2-1	0-0	2-0	4-3	2-0	1-0	2-1	1-2	1-3	1-0	0-1	1-1	2-1
BRESCIA	3-3	3-0		2-2	3-0	1-3	0-4	1-1	1-1	0-0	1-4	3-0	2-2	0-0	1-2	2-0	3-2	0-0
CHIEVO	2-1	2-0	1-1		2-2	2-2	1-3	3-1	2-1	1-1	1-0	2-0	4-2	0-3	3-0	1-2	1-1	2-1
FIorentina	3-1	1-1	1-0	0-2		0-1	1-1	0-1	1-2	1-1	1-2	1-3	1-3	2-2	0-0	0-0	3-1	0-2
INTER	1-2	1-0	2-1	1-2	2-0		2-2	0-0	2-0	2-4	2-0	4-1	3-1	3-1	0-0	3-2	2-1	3-0
JUVENTUS	3-0	2-1	5-0	3-2	2-1	0-0		1-1	3-0	1-0	3-1	2-0	2-0	0-2	3-3	3-0	4-0	1-0
LAZIO	2-0	2-2	5-0	1-1	3-0	4-2	1-0		1-0	1-1	0-0	5-0	1-1	1-5	0-0	2-0	4-2	5-4
LECCE	0-2	1-0	1-3	2-3	4-1	1-2	0-0	1-2		0-1	1-1	2-3	0-0	1-1	1-1	1-2	2-1	1-1
MILAN	0-0	0-0	0-0	3-2	5-2	0-1	1-1	2-0	3-0		3-1	1-1	0-0	0-0	2-1	2-3	1-1	2-1
PARMA	1-1	2-1	1-0	0-0	2-0	2-2	1-0	1-0	1-1	0-1		2-1	2-2	1-2	0-1	2-0	2-1	2-2
PERUGIA	2-0	1-0	1-1	2-2	2-0	0-2	0-4	0-0	2-1	3-1	2-1		1-0	0-0	2-0	1-2	2-0	3-1
PIACENZA	1-2	2-0	0-1	2-2	3-0	2-3	0-1	1-0	1-2	0-1	2-3	2-0		2-0	3-1	1-2	5-0	3-0
ROMA	3-1	3-1	0-0	5-0	2-1	0-0	0-0	2-0	5-1	1-0	3-1	1-0	2-0		1-0	1-1	1-0	3-2
TORINO	1-2	1-1	1-3	2-2	1-0	0-1	2-2	1-0	1-1	1-0	1-0	1-0	1-1	0-1		3-1	1-2	5-1
UDINESE	1-3	0-1	3-2	1-2	1-2	1-1	0-2	1-4	0-1	1-2	3-2	0-0	1-1	1-1	2-2		1-0	2-1
VENEZIA	0-1	0-1	1-2	0-0	2-0	1-1	1-2	0-0	1-1	1-4	3-4	0-2	2-3	2-2	1-1	2-1		0-1
VERONA	3-1	0-1	2-0	3-2	1-2	0-3	2-2	3-1	2-1	1-2	1-0	1-1	1-0	1-1	0-1	1-0	1-0	

Ritorniamo al punto. Per ogni squadra si ha, alla fine del campionato, il risultato di tutte le partite disputate, come nello schema indicato, che si legge così: la squadra scritta sulla fiancata sinistra gioca in casa, quella sulla prima riga in alto fuori casa. Per esempio, leggiamo nel triangolo in alto: Inter – Lazio 0-0 e in quello in basso Lazio - Inter 4-2. Possiamo ora condensare, restringere tutte queste informazioni e registrare ognuna delle 306 partite a seconda dei goal fatti dalla squadra che gioca in casa e di quelli messi a segno dalla squadra ospite. Per esempio, nello schema della tavola 1 (detta tabella a doppia entrata), il risultato Lazio – Inter (4-2) va registrato nella casella che si trova all'incrocio della quinta riga (corrispondente a 4 goal fatti dalla squadra che gioca in casa) e della terza colonna (dove sono registrate le partite in cui la squadra che gioca fuori casa mette a segno 2 goal). In questa casella troviamo 3 frequenze. Cioè 3 delle 306 partite del campionato di serie A, per l'anno 2001-2002, sono terminate sul punteggio di 4 a 2.

La completa classificazione del campionato di serie A secondo l'esito delle partite come appena spiegato, viene mostrata, per gli anni che abbiamo preso in considerazione, nelle tavole da 1 a 5.

**Tavola 1 - Partite classificate secondo il loro esito per i goal fatti dalla squadra che giocava in casa e per i goal fatti dalla squadra avversaria, che giocava fuori casa. Risultati del campionato di calcio dell'anno 2001-2002, serie A**

		Goal effettuati dalla squadra ospite						
		0	1	2	3	4	5	Totale
Goal effettuati dalla squadra che gioca in casa	0	28	20	8	2	2	0	60
	1	33	39	25	8	3	2	110
	2	27	25	18	6	2	0	78
	3	12	17	9	2	1	0	41
	4	1	3	3	1	0	0	8
	5	5	2	1	0	1	0	9
<b>Totale</b>		<b>106</b>	<b>106</b>	<b>64</b>	<b>19</b>	<b>9</b>	<b>2</b>	<b>306</b>

**Tavola 2 - Partite classificate secondo il loro esito per i goal fatti dalla squadra che giocava in casa e per i goal fatti dalla squadra avversaria, che giocava fuori casa. Risultati del campionato di calcio dell'anno 1991-1992, serie A**

		Goal effettuati dalla squadra ospite						Totale
		0	1	2	3	4	5	
Goal effettuati dalla squadra che gioca in casa	0	38	23	13	0	1	0	75
	1	41	56	13	10	3	0	123
	2	29	19	10	3	0	0	61
	3	6	11	4	4	1	0	26
	4	7	5	1	0	1	0	14
	5	2	2	2	0	0	0	6
<b>Totale</b>		<b>123</b>	<b>116</b>	<b>43</b>	<b>17</b>	<b>6</b>	<b>0</b>	<b>305</b>

**Tavola 3 - Partite classificate secondo il loro esito per i goal fatti dalla squadra che giocava in casa e per i goal fatti dalla squadra avversaria, che giocava fuori casa. Risultati del campionato di calcio dell'anno 1981-1982, serie A**

		Goal effettuati dalla squadra ospite							Totale
		0	1	2	3	4	5	6	
Goal effettuati dalla squadra che gioca in casa	0	40	26	7	1	1	0	0	75
	1	38	37	7	3	1	1	0	87
	2	16	25	10	1	0	0	0	52
	3	6	3	7	0	0	0	0	16
	4	2	3	3	0	0	0	0	8
	5	0	0	0	0	0	0	0	0
	6	0	1	0	0	0	0	0	1
<b>Totale</b>		<b>102</b>	<b>95</b>	<b>34</b>	<b>5</b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>239</b>

**Tavola 4 - Partite classificate secondo il loro esito per i goal fatti dalla squadra che giocava in casa e per i goal fatti dalla squadra avversaria, che giocava fuori casa. Risultati del campionato di calcio dell'anno 1971-1972, serie A**

		Goal effettuati dalla squadra ospite							Totale
		0	1	2	3	4	5	6	
Goal effettuati dalla squadra che gioca in casa	0	41	13	8	2	2	0	0	66
	1	44	33	8	4	1	0	1	91
	2	19	28	1	3	1	0	0	52
	3	5	9	1	0	0	0	0	15
	4	3	1	1	0	1	0	0	6
	5	0	0	0	0	0	0	0	0
	6	0	0	1	0	0	0	0	1
<b>Totale</b>		<b>112</b>	<b>84</b>	<b>20</b>	<b>9</b>	<b>5</b>	<b>0</b>	<b>1</b>	<b>231</b>

**Tavola 5 - Partite classificate secondo il loro esito per i goal fatti dalla squadra che giocava in casa e per i goal fatti dalla squadra avversaria, che giocava fuori casa. Risultati del campionato di calcio dell'anno 1961-1962, serie A**

Goal effettuati dalla squadra ospite

		0	1	2	3	4	5	6	Totale
Goal effettuati dalla squadra che gioca in casa	0	29	26	11	7	0	0	0	73
	1	42	39	10	10	0	0	0	101
	2	24	25	2	7	3	1	0	62
	3	12	14	2	2	0	0	0	30
	4	5	8	3	2	0	0	0	18
	5	0	3	4	0	0	0	0	7
	6	1	0	0	0	1	0	0	2
<b>Totale</b>		<b>113</b>	<b>115</b>	<b>32</b>	<b>28</b>	<b>4</b>	<b>1</b>	<b>0</b>	<b>293</b>

Lo scopo della costruzione di queste classificazioni consiste nel valutare se nell'arco di circa 40 anni si siano registrate modificazioni nell'esito delle partite giocate nel campionato di serie A.

Osserviamo che tutte le tavole presentate possiedono una comune caratteristica: nella diagonale da nord-ovest a sud-est cadono le frequenze, cioè il numero delle partite che sono terminate in parità. Per esempio, nella tavola 1, leggiamo che 28 partite sono terminate 0 a 0; 39 col punteggio 1 a 1; 18 partite con il punteggio di 2 a 2; soltanto 2 partite con il punteggio di 3 a 3 e infine zero col risultato di 4 a 4 e 5 a 5. Ancora, nel triangolo inferiore, rispetto alla diagonale appena individuata, cadono tutte le partite che sono terminate con la vittoria della squadra che giocava in casa, mentre nel triangolo superiore vengono a trovarsi le partite in cui la squadra ospite ha riportato la vittoria.

Potete anche rilevare l'andamento delle partite secondo il risultato di parità, vittoria o sconfitta della squadra che giocava in casa. Vale la pena, a fini comparativi, dei vari anni fra loro, raccogliere questi risultati nella tavola 6.

**Tavola 6 - Esiti delle partite di serie A negli anni sottoindicati e relative percentuali**

ESITI PARTITE	1961-1962		1971-1972		1981-1982		1991-1992		2001-2002	
	N	%	N	%	N	%	N	%	N	%
In parità	72	24,6%	76	32,9%	87	36,4%	109	35,7%	87	28,4%
Vince la squadra che gioca in casa	146	49,8%	112	48,5%	104	43,5%	129	42,3%	140	45,8%
Vince la squadra che gioca fuori casa	75	25,6%	43	18,6%	48	20,1%	67	22,0%	79	25,8%
<b>Totale partite</b>	<b>293</b>	<b>100%</b>	<b>231</b>	<b>100,0%</b>	<b>239</b>	<b>100,0%</b>	<b>305</b>	<b>100,0%</b>	<b>306</b>	<b>100,0%</b>

Poiché nei campionati dei differenti anni il numero complessivo delle partite giocate non è sempre uguale e dipende dal numero di squadre che formano il girone, è necessario calcolare i valori percentuali delle partite finite in parità oppure vinte o perse dalle squadre che giocavano in casa. In questo modo i confronti vengono a essere facilitati.

Il calcolo è molto semplice e lo illustriamo qui di seguito con riferimento ai risultati della serie A per l'anno 2001-2002. Poiché 87 partite sono terminate in parità su 306 giocate, quante sarebbero se ne fossero state giocate 100? A questa domanda si risponde formalmente impostando la seguente proporzione:

$$87 \div 306 = X \div 100$$

ovvero, altrimenti scritta

$$\frac{87}{306} = \frac{X}{100}$$

da cui si ricava facilmente, moltiplicando entrambi i membri per 100, che  $X = \frac{87}{306} \times 100$  e quindi circa uguale a  $X \approx 28,4\%$  (che si può anche scrivere in forma decimale:  $X \approx 0,284$ ).

In modo analogo determiniamo la percentuale di partite con la vittoria della squadra che gioca in casa:

$$140 \div 306 = X \div 100$$

ovvero 
$$\frac{140}{306} = \frac{X}{100} \quad X \approx 45,8\%$$

e la percentuale di partite vinte dalla squadra ospite, e quindi perse dalla squadra che giocava in casa:

$$79 \div 306 = X \div 100$$

ovvero

$$\frac{79}{306} = \frac{X}{100} \quad X \approx 25,8\%.$$

Vi rendete conto anche, senza molte difficoltà, che la somma delle tre percentuali deve risultare pari a 100, poiché

$$\frac{87}{306} \times 100 + \frac{140}{306} \times 100 + \frac{79}{306} \times 100 = 100 \times \frac{87 + 140 + 79}{306} = 100 \times 1 = 100.$$

Vediamo ora di leggere e interpretare i valori delle percentuali appena calcolate.

Le partite terminate in parità, nel corso del tempo, sono andate aumentando passando dal 25 per cento del campionato 1961-1962 al 36 per cento del campionato 1991-1992, a riscontro di un gioco sempre più chiuso e difensivo. Successivamente, anche per l'introduzione della regola dei tre punti per le partite vinte, invece di due, a metà degli anni Novanta, il numero delle partite terminate con un pareggio è consistentemente diminuito (28 per cento). Considerazioni analoghe si possono effettuare leggendo le percentuali delle partite vinte in casa e fuori casa.

C'è ancora un altro modo di guardare ai risultati raccolti nelle tavole da 1 a 5. Possiamo calcolare il numero medio di goal per partita fatti dalle squadre che giocavano in casa, da quelle che giocavano fuori casa e il numero medio di goal per partita che, naturalmente, risulta essere la somma dei due precedenti valori.

Vediamo come calcolare queste medie con riferimento alla tavola 1 (campionato 2001-2002). Leggiamo nella tavola i totali marginali per riga e per colonna che stanno a indicarci, i primi, la distribuzione di tutte le partite a seconda dei goal messi a segno dalla squadra che giocava in casa quale che sia il numero dei goal messi a segno dalla squadra contraria; i secondi, invece, la distribuzione delle partite a seconda dei goal ottenuti dalla squadra che giocava fuori casa, senza tener conto dei goal fatti dalla squadra che giocava in casa. Così, per esempio, 60 partite finirono con zero reti per la squadra che giocava in casa; 110 finirono con 1 rete per la squadra che giocava in casa; 78 partite finirono con 2 reti e così via. Cambiando il ruolo delle squadre possiamo interpretare, in via del tutto analoga, i totali che leggiamo nell'ultima riga (totali marginali di colonna).

Veniamo ora a calcolare le medie, e, in particolare, stabiliamo il numero medio di goal per partita messo a segno dalle squadre che giocavano in casa. Questo valore lo possiamo calcolare così:

$$\frac{0 \times 60 + 1 \times 110 + 2 \times 78 + 3 \times 41 + 4 \times 8 + 5 \times 9}{60 + 110 + 78 + 41 + 8 + 9} = \frac{466}{306} = 1,523,$$

cioè come media aritmetica ponderata del numero di goal pesato con le partite che hanno registrato quel particolare risultato. Al numeratore si individua così il numero complessivo di goal (466) segnato dalle squadre che giocavano in casa, che, diviso per il numero delle partite (306), dà il numero medio di goal a partita (1,523); non dovrebbero sussistere dubbi sull'interpretazione delle frazioni di goal!

In modo del tutto analogo calcoliamo, come segue, il numero medio di goal a partita segnato dalle squadre che giocavano fuori casa:

$$\frac{0 \times 106 + 1 \times 106 + 2 \times 64 + 3 \times 19 + 4 \times 9 + 5 \times 2}{106 + 106 + 64 + 19 + 9 + 2} = \frac{337}{306} = 1,101$$

Naturalmente, queste medie le possiamo calcolare per tutti i campionati qui considerati. I risultati sono sintetizzati nella tavola 7.

**Tavola 7 - Goal fatti in casa, fuori casa e in totale per i campionati di serie A sottoindicati. Numero medio di goal per partita, fatti in casa, fuori casa e in totale**

PARTITE	Goal				
	1961-1962	1971-1972	1981-1982	1991-1992	2001-2002
In casa	434	270	277	409	466
Fuori casa	284	177	191	277	337
Totale	718	447	468	686	803
Partite giocate	293	231	239	305	306
	NUMERO MEDIO DI GOAL PER PARTITA				
In casa	1,481	1,169	1,159	1,340	1,523
Fuori casa	0,969	0,766	0,799	0,908	1,101
Totale	2,450	1,935	1,958	2,249	2,624

Se soffermiamo lo sguardo sulle cifre dei valori medi che, fra quelle della tavola 7, appaiono come le più interessanti, ci accorgiamo facilmente di una caratteristica. Non soltanto il numero medio di goal segnati in casa risulta sempre superiore a quello delle partite giocate in trasferta, ma anche in questo caso l'introduzione della regola dei tre punti per ogni partita vinta sembra generare un gioco più vivace e ricco, come risulterebbe dall'elevarsi delle medie dei goal passando dal penultimo all'ultimo dei periodi qui considerati.

### **Frequenza, distribuzione di frequenza, seriazione, media aritmetica e ponderata**

La **frequenza** è il numero di volte in cui un carattere presenta una sua data modalità. Ad esempio, se il carattere è il numero dei rigori assegnati a favore delle diverse squadre in un campionato e le sue modalità sono: zero, uno, due...cinque, le frequenze rappresentano il numero delle squadre che hanno usufruito, appunto, di zero, uno, due, ...cinque rigori a favore.

In una **distribuzione di frequenza**, una variabile quantitativa articolata in modalità viene presentata assieme alle frequenze di casi che corrispondono a ciascuna di esse; questa organizzazione dei dati viene denominata anche **seriazione**. La **media aritmetica ponderata** viene definita come:

$$M = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i}$$

dove  $x_i$  è la modalità  $i$ -esima – eventualmente, può trattarsi del valore centrale della  $i$ -esima classe – e  $f_i$  la frequenza associata a essa. Si tratta di una estensione della media aritmetica semplice, nel caso in cui alcuni valori siano, o possano ritenersi, ai fini del calcolo, uguali fra loro (cioè compaiano più volte, con frequenza  $f_i$ ).

Si diceva poco sopra che un'analisi anno per anno, più minuziosa, di tutti i campionati potrebbe portarci a una descrizione più chiara delle tendenze; e questo è certamente vero. Tuttavia, tenete ben presente che un maggior dettaglio non significa automaticamente una maggior facilità nel capire le cause dei cambiamenti. Per comprendere e interpretare questi risultati occorre una buona conoscenza del gioco e del mondo del calcio. La statistica, in questo caso, ci ha aiutato soltanto a mettere un po' di ordine fra il grosso mare dei dati, a fare un sunto delle osservazioni.

Il gioco del calcio si presta a molte considerazioni di carattere quantitativo; voi potete esaminare altri aspetti come le autoreticoli, i rigori realizzati rispetto a quelli concessi.

Per quanto riguarda i rigori, potete verificare se la regola dei rigori, adottata nei campionati del mondo di calcio e in altri tornei, secondo la quale in caso di parità anche dopo i tempi supplementari si decide con una serie di dieci rigori, cioè cinque per ogni squadra (si prosegue poi a oltranza in eventuale caso di parità) ha una sua conferma nella realtà e sia stata pensata avendo in mente un risultato statistico (da alcuni studi risulta che circa sette rigori su dieci finiscono in rete). Il numero di rigori misura anche l'andamento del gioco: calcolando la percentuale di rigori per ogni 100 partite si ha un'informazione sulla correttezza o meno del gioco (la percentuale si ottiene dividendo il numero di rigori concessi per il numero di partite giocate, moltiplicato 100). Un carattere collegato alla scorrettezza del gioco e alle valutazioni arbitrali, più o meno severe, riguarda le espulsioni di giocatori dal campo. Possiamo prendere ad esempio le espulsioni che si sono avute nel 1973-1974 e nel 1983-1984; conosciamo le distribuzioni di frequenza del numero di squadre che, nel corso del campionato, hanno avuto un certo numero di giocatori espulsi:

**Tavola 8 – Giocatori espulsi della serie A nei campionati 1973-1974, 1983-1984**

CAMPIONATI	Numero di giocatori espulsi						Totale
	0	1	2	3	4	5	
1973-74	4	2	5	5	-	-	16
1983-84	1	2	7	3	-	3	16

Da queste distribuzioni si rileva che un numero ristretto di squadre, 16 in ambedue i campionati, ha avuto giocatori che si sono comportati in modo particolarmente scorretto; ma per capire meglio come vanno le cose occorrerebbe un'analisi più particolareggiata (vedere, per esempio, di quali squadre si sia trattato e se le espulsioni si siano concentrate in qualche particolare partita eccetera) e l'attenzione dovrebbe essere estesa a molti più campionati.

## Esercitazioni

1. Si possono dedicare pochi minuti del lunedì mattina a rilevare i dati del campionato di serie A (o anche di quello in cui milita la squadra della provincia in cui abitate) al fine di compilare una tavola come la tavola 1. A turno, due o tre ragazzi/e possono rilevare dati sui rigori, sulle espulsioni o altro. Questo materiale dovrebbe essere raccolto per tutto il campionato ed elaborato, magari nell'anno scolastico successivo.
2. Con riferimento ai dati delle tavole 1-5, calcolate le medie dei goal fatti in casa e fuori casa. In questo modo si possono controllare i risultati della tavola 7.
3. Rappresentate su un diagramma cartesiano il numero medio di goal fatti dalla squadra che gioca in casa oppure da quella ospite, in ordinata, in relazione agli anni di campionato, in ascissa.
4. Costruite le percentuali di squadre della serie A per numero di autoreticoli in alcuni campionati.
5. Disegnate i diagrammi a torta per i rigori realizzati e non realizzati per la serie A in alcuni campionati.

## 6. La pesca sportiva in acque dolci

### Che cosa si impara nel capitolo 6

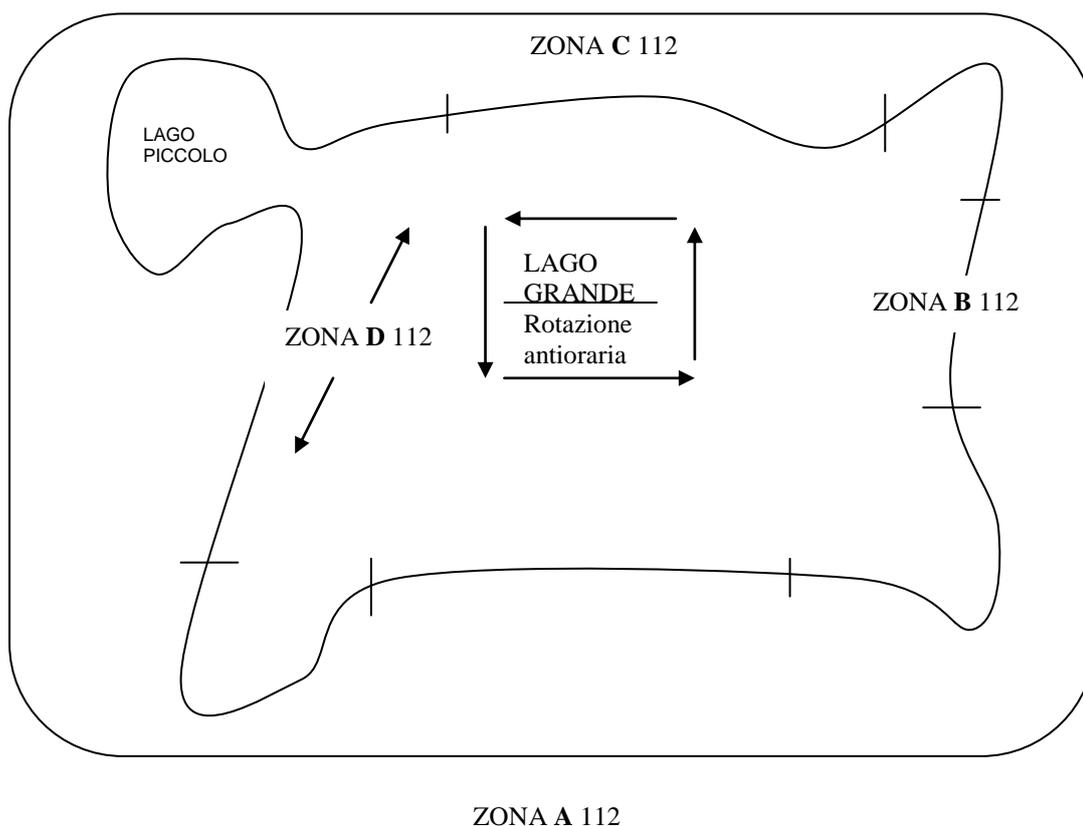
I risultati di una gara di pesca fra alcuni concorrenti si possono rappresentare attraverso , il **diagramma ramo-foglia**, la **mediana**, i **quartili**, le **frequenze cumulate** e **retrocumulate**, le **distribuzioni di frequenze cumulate** e **retrocumulate**, il **diagramma a scatola (box plot)**. Si impara a calcolare la **varianza** e lo **scostamento quadratico medio**.

Nel 1986 si è svolta al lago di Salasco – un bacino artificiale nei pressi di Vercelli – una gara di pesca. Alla gara partecipano 448 pescatori.

I giudici di gara registrano per ogni concorrente il numero di trote catturate al momento in cui il pesce viene portato a terra e controllano la regolarità della gara. I concorrenti iniziano a pescare su un tratto della riva ma poi, nel corso della gara, vengono fatti spostare (ruotare nel caso presente) per eliminare le eventuali differenze di pescosità di parti del lago rispetto ad altre. Alla fine della gara si pesano le trote catturate da ciascun pescatore e si compila una classifica per singolo pescatore e per ogni squadra, come si vede nel prospetto della figura 2. Si seguono le seguenti regole:

- ogni trota catturata, indipendentemente dal peso, vale mille punti;
- il peso complessivo dei pesci pescati, registrato in grammi, si aggiunge al punteggio precedente.

Figura 1 – Piantina del lago con le postazioni di pesca e il senso della rotazione



**Figura 2 – Esempio di modulo di registrazione**

N. di gara	Concorrente	Società	N. Catture	Punti peso				Totale punti	Classifica
22	Concorrente A	Società a	14	1	4	0	0	18.060	1
					4	0	6		
24	Concorrente B	Società b	12	1	2	0	0	15.310	2
					3	3	1		
100	Concorrente C	Società c	11	1	1	0	0	13.980	3
					2	9	8		
88	Concorrente D	Società d	10	1	0	0	0	12.870	4
					2	8	7		
80	Concorrente E	Società e	9	9	0	0	0	11.700	5
					2	7	0		
14	Concorrente F	Società f	9	9	0	0	0	11.690	6
					2	6	9		
110	Concorrente G	Società g	9	9	0	0	0	11.565	7
					2	5	6		
67	Concorrente H	Società h	9	9	0	0	0	11.555	8
					2	5	5		
70	Concorrente I	Società i	9	9	0	0	0	11.450	9
					2	4	5		
77	Concorrente L	Società l	9	9	0	0	0	11.210	10
					2	2	1		
Ora di esposizione 13.16		Il Direttore di gara			Il Comm. sportivo federale				

A questo punto, ci si trova con 45 fogli, come quello riportato qui sopra (Figura 2) e altri 48 per le registrazioni delle squadre; un bel po' di carta in cui si legge la classifica che permetterà di assegnare le coppe ai vincitori. Per concludere la gara può bastare; ma per i dirigenti sportivi che vogliono riflettere sui risultati della singola gara e compararli con quelli di altre, e per le autorità locali che badano al controllo della fauna ittica, pensate che sia utile avere i dati in quella forma? Non crediamo; proviamo a vedere che cosa si può fare.

Intanto, poiché le sponde del lago di Salasco sono state suddivise in quattro zone, su cui i pescatori sono stati ripartiti per iniziare la gara (Figura 1), possiamo costruire le seriazioni per numero di catture, a seconda della zona in cui è iniziata la gara (Tavola 1).

E già possiamo iniziare a osservare alcune cose interessanti che nei fogli non si vedevano.

Su ognuna delle quattro zone iniziali c'era lo stesso numero di pescatori (112 è il totale comune): perché? Ma perché le squadre erano formate da quattro persone e per non favorire nessuna squadra, come potrebbe accadere se una zona fosse ricchissima di pesci e un'altra completamente spopolata, gli organizzatori hanno assegnato a ognuna delle quattro zone un differente componente della stessa squadra. (Attenzione, la situazione estrema che abbiamo supposto circa la pescosità è poco verosimile che si verifichi in realtà: però è bene usare la massima prudenza, per evitare effetti indesiderati).

Inoltre, ogni pescatore è stato fatto successivamente ruotare in senso antiorario lungo le sponde del lago nel corso della gara durata un'ora e quindici minuti. In questo modo ogni concorrente ha pescato in tutti e quattro i settori e in due posti diversi di ciascun settore. Abbiamo comunque elaborato le distribuzioni di frequenza per settori iniziali di pesca in mancanza del dettaglio dei pesci catturati da ciascun pescatore nei successivi settori (sarebbe interessante saggiare che cosa accade nella pesca col trascorrere del tempo, per capire se le regole di gara sono adatte alle situazioni reali), e nella convinzione che dopo un po' di tempo dall'inizio della gara muti la distribuzione dei pesci nel lago.

**Tavola 1 – Pescatori per zona in cui hanno iniziato a pescare e per catture di trote effettuate. Media, varianza e scostamento quadratico medio per ciascuna della quattro zone**

CATTURE	Zone			
	A	B	C	D
0	17	6	12	10
1	6	10	11	9
2	10	9	19	9
3	12	17	13	25
4	12	20	14	11
5	15	9	15	12
6	14	8	3	8
7	8	9	8	7
8	8	9	7	11
9	1	4	6	4
10	5	6	1	3
11	2	4	1	3
12	-	-	1	-
13	1	1	-	-
14	-	-	1	-
15	1	-	-	-
<b>Totale</b>	112	112	112	112
Media	4,446	4,821	4,018	4,375
Varianza	10,104	9,129	8,838	8,252
Scostamento quadratico medio	3,179	3,021	2,973	2,873

E ancora possiamo aggiungere che, poiché tutti i totali sono uguali, non c'è bisogno per la comparazione di calcolare le frequenze relative; vediamo subito così che nella zona iniziale A – quella da cui è venuto fuori il vincitore con 15 catture – vi è stato il maggior numero di pescatori sfortunati: ben 17 sono rimasti a bocca (ad amo) asciutta.

Possiamo poi vedere che i pescatori molto bravi o molto fortunati (o tutt'e due), cioè quelli che hanno preso almeno 11 pesci (vale a dire: o 11 o 12 o 13 o...) sono – c'era da aspettarselo – pochini (quattro di quelli che hanno iniziato nella zona A; cinque in B; tre in C; tre in D; che in percentuale fa:

$$\frac{4}{112} \times 100 \approx 3,6\% \text{ in A;}$$

$$\frac{5}{112} \times 100 \approx 4,5\% \text{ in B;}$$

$$\frac{3}{112} \times 100 \approx 2,7\% \text{ in C e D).}$$

Bene, già ci siamo resi conto di alcune cose. Possiamo sintetizzare di più i nostri dati? Possiamo renderci conto di quanto variano; che varino lo vedete direttamente sulla tavola, dal momento che alcuni hanno preso 0, altri 1,..., altri 15 trote. Possiamo tentare di misurare questa variazione.

Cominciamo a calcolare la media aritmetica ponderata dei pesci catturati nelle quattro zone e iniziamo dalla zona A:

$$M = \frac{(0 \times 17 + 1 \times 6 + 2 \times 10 + 3 \times 12 + 4 \times 12 + 5 \times 15 + 6 \times 14 + 7 \times 8 + 8 \times 8 + 9 \times 1 + 10 \times 5 + 11 \times 2 + 12 \times 0 + 13 \times 1 + 14 \times 0 + 15 \times 1)}{112} = \frac{498}{112} = 4,446$$

Le altre medie le trovate scritte in fondo alla tavola 1 e il loro computo ricalca da vicino quello che abbiamo qui effettuato; sarà bene che per esercizio ne controlliate i valori.

Vediamo ora qual è il significato di questa media. Al numeratore, composto dalla somma di prodotti, trovate il numero di pesci pescati da tutti i pescatori del settore A: si tratta di 498 trote. Se ora dividiamo questo valore per 112, per il numero di pescatori dunque, otteniamo il numero di pesci catturati per pescatore, secondo il significato esplicito dell'operazione di divisione (in parti uguali). O ancora meglio: individuiamo il numero (medio) di pesci pescati per persona supponendo che ogni pescatore abbia pescato la stessa quantità di pesci. Si tratta di un numero ipotetico e da noi costruito, però! Fate attenzione: è il numero

che viene fuori da una situazione che immaginiamo per nostra comodità e, in particolare, per cogliere l'ordine di grandezza del numero di pesci che, grosso modo, viene pescato da ciascun pescatore. Di questo fatto vi accorgete subito perché trovate un valore un po' strano: 4,446; cioè non un numero intero come sarebbe capitato se effettivamente tutti i pescatori avessero pescato un numero uguale di pesci. Che senso dare ai decimi di pesce (ai 44 centesimi, ai 446 millesimi,...) che dobbiamo aggiungere al valore intero 4? Beh, intanto e ovviamente, che la media è più grande di 4 ma più piccola di 5 pesci a testa; poi, se proprio vi dà fastidio il numero decimale, possiamo dire che 10 pescatori avrebbero pescato 44 pesci (ovvero, 100 pescatori 444 pesci, 1000 pescatori 4446 pesci,...).

Se siamo d'accordo sul significato da accordare a questa media andiamo avanti e compariamo i valori delle quattro medie per le quattro zone iniziali di pesca. Ma prima osserviamo, effettuando la somma dei numeratori delle quattro medie, che sono state catturate nel complesso:

$$498 + 540 + 450 + 490 = 1978 \text{ trote.}$$

Chiariamo che, se si divide 1978 per il totale dei pescatori ( $4 \times 112 = 448$ ), si ottiene il numero medio di pesci pescati, senza considerare la zona iniziale di gara e che questa quantità deve risultare – dal momento che tutte e quattro le medie in basso nella tavola 1 sono superiori a 4 e inferiori a 5 – circa ... (tentate voi una stima). Infatti:

$$\frac{1978}{448} = 4,415.$$

Rileviamo anche che tutte e quattro le medie sono numericamente differenti tra loro, però questo criterio puramente algebrico per stabilire se le medie differiscano fra loro non è soddisfacente: fra le medie relative ai pescatori che hanno iniziato nella zona A e a quelli che hanno iniziato nella zona D intercorre solo 0,071, mentre fra quelle di B e C 0,803; quantità questa più di 11 volte maggiore della precedente. Non è affatto semplice dire se effettivamente la media della zona B sia superiore a quella della zona C, cioè se le due zone risultino realmente pescose in modo diverso, perché la differenza riscontrata potrebbe essere stata prodotta da fatti casuali. Immaginiamo, ad esempio, che nella zona B il giorno precedente la gara, dei ragazzi abbiano gettato gli avanzi del vicino ristorante; allora sarebbe verosimile che più pesci si siano diretti in quella zona per trovarvi cibo; se lo sapessimo potremmo tenerne conto nell'interpretazione, ma non lo sappiamo! Potremmo fare molte altre supposizioni, ma sarebbero tutte inutili e non ci farebbero fare passi avanti.

Il problema del confronto fra due o più medie ha interessato gli statistici che lo hanno dovuto affrontare, assieme a quello dell'induzione sperimentale, cioè di come risalire da un caso particolare alla situazione generale. Possiamo a ogni modo notare che i nostri dati indicano che la zona C risulta quella in cui si è registrato il valore medio di pesci pescati più basso per i pescatori che hanno iniziato dalla zona C e, di contro, la media risulta più elevata per i pescatori che hanno iniziato dalla zona B.

Ancora, sarà utile interessarci alla variabilità del numero di pesci pescati nelle quattro zone, perché se è vero che i pescatori si distribuiscono sui valori da zero a quindici catture, mostrando così la variabilità del fenomeno, non lo fanno in modo uguale qualsiasi sia la zona dalla quale hanno iniziato a pescare (le stesse medie sono differenti). Se misuriamo la variabilità in modo sintetico, calcolando le varianze per ciascuna zona (scritte in basso nella tavola 1), ci accorgiamo ancora una volta di avere valori differenti, ma non molto diversi fra loro. Più o meno sembra esserci la stessa variabilità secondo i territori iniziali di pesca.

Consideriamo ora un'altra caratteristica che si desume facilmente dalle registrazioni della gara. Poiché per ogni concorrente abbiamo sia il peso complessivo delle trote pescate sia il loro numero, possiamo calcolare il peso medio, per ciascun pescatore, delle trote da lui catturate:

$$\text{peso medio in grammi} = \frac{\text{peso complessivo in grammi}}{\text{numero delle trote pescate}}.$$

Disponiamo così di tanti pesi medi quanti sono i pescatori che hanno catturato almeno un pesce. Questi valori li possiamo raccogliere nei seguenti diagrammi ramo-foglia per ciascuna delle quattro zone iniziali di pesca (Figura 3).

Figura 3 - Diagrammi ramo-foglia del peso medio dei pesci pescati da ciascun pescatore per ognuna delle quattro zone iniziali di gara. Parametri riassuntivi di ciascuna esperienza

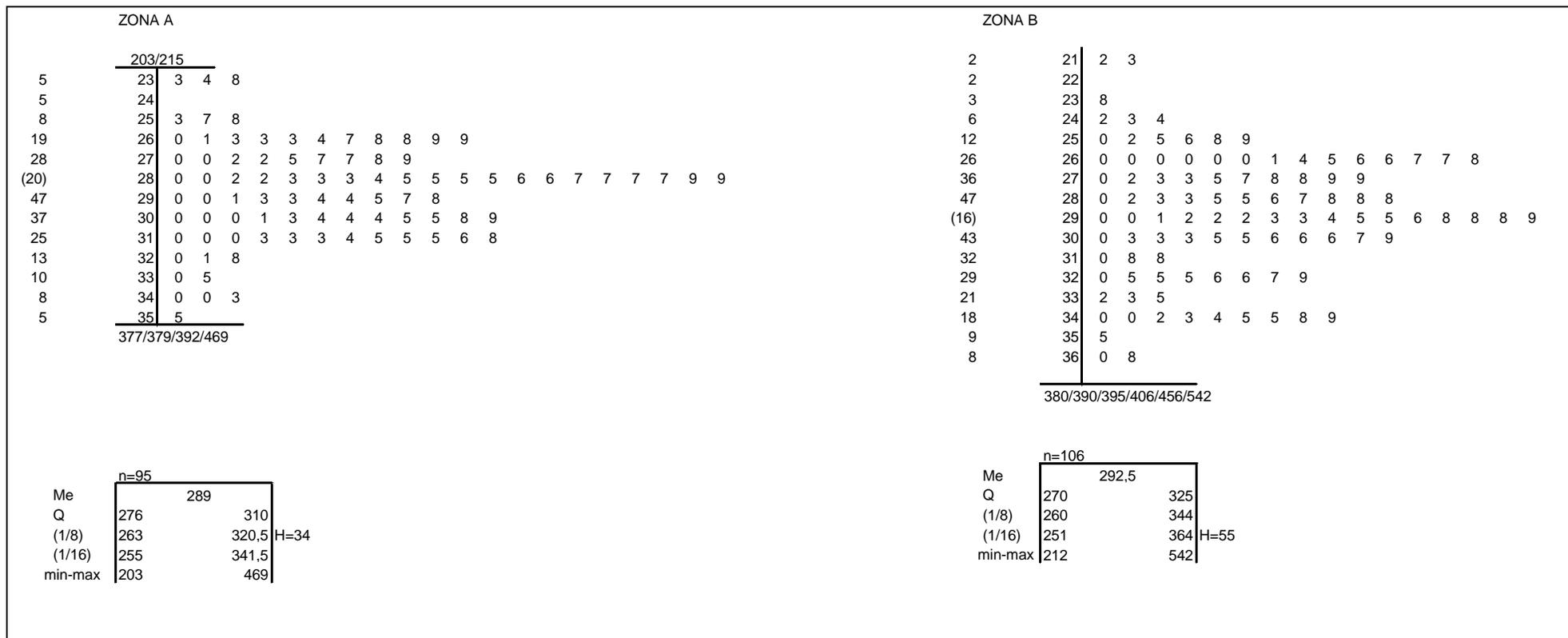
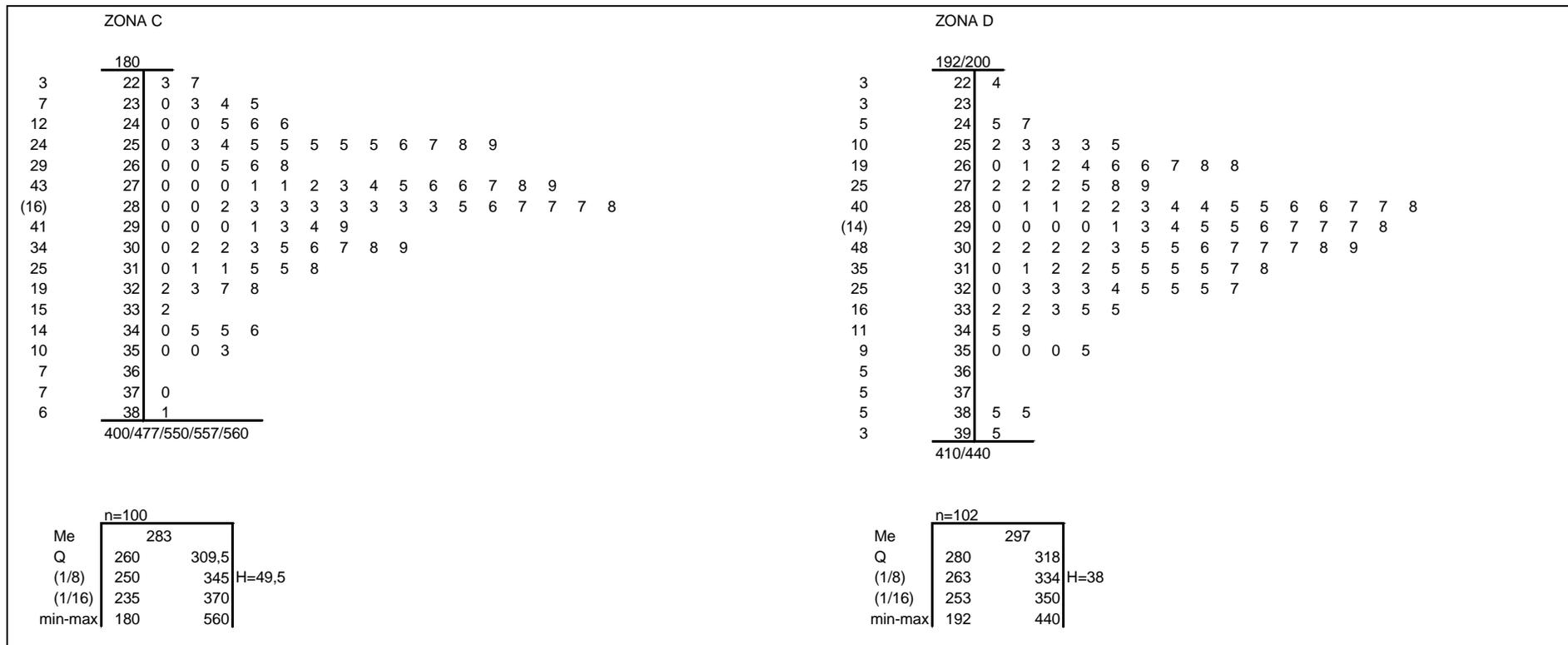




Figura 3 segue - Diagrammi ramo-foglia del peso medio dei pesci pescati da ciascun pescatore per ognuna delle quattro zone iniziali di gara. Parametri riassuntivi di ciascuna esperienza



Qualche parola di commento può aiutarci a chiarire l'impianto di queste costruzioni. Le osservazioni sono state suddivise così: il valore di grammi 327, per esempio, viene ripartito in due gruppi di cifre, le prime due  $\boxed{32}$  formano il ramo e la terza  $\boxed{7}$  la foglia. In ciascuno dei diagrammi – poi – i valori molto bassi e molto alti sono stati scritti separatamente in cima e in basso dei rami considerati. Sul lato sinistro, si trovano le frequenze cumulate e retrocumulate sino al ramo che contiene la mediana (si tratta della loro somma progressiva a partire dalla prima e, rispettivamente, dall'ultima). Al di sotto di ciascun diagramma trova posto il riassunto composto dal numero delle osservazioni ( $n$ ), dalla mediana ( $Me$ ), dal primo e terzo quartile ( $Q$ ), dall'ottavo e dal sedicesimo inferiore e superiore (rispettivamente  $(1/8)$ ,  $(1/16)$ ), dal massimo e dal minimo dei valori osservati e dalla differenza interquartile, cioè la differenza fra il terzo e il primo quartile ( $H=Q_3-Q_1$ ).

Vediamo rapidamente, per i dati della zona A, come si calcolano questi parametri. La mediana è quel valore che si trova al centro delle osservazioni ordinate, cioè quel valore tale che le osservazioni maggiori di esso sono tante quante le minori: siccome  $n=95$ , l'ordine della mediana è pari a  $\left(\frac{95+1}{2}\right)=48$ , cioè il 48° valore (un peso pari a 289 grammi). Il primo quartile ha ordine:

$$\frac{\text{ordine della mediana} + 1}{2} = \frac{48 + 1}{2} = \frac{49}{2} = 24,5$$

vale a dire che occorre prendere la semisomma del valore di posto 24° e del valore di posto 25°, poiché il primo quartile lo possiamo considerare come la mediana della prima metà dei dati. Vediamo subito che si ha:

$$Q_1 = \frac{275 + 277}{2} = 276 \text{ grammi.}$$

Il terzo quartile, in modo analogo, lo troviamo nella seconda metà dei dati più elevati e vale, considerando i posti 24° e 25° dall'alto

$$Q_3 = \frac{310 + 310}{2} = 310 \text{ grammi.}$$

L'ottavo inferiore ha ordine:

$$\frac{\text{valore intero dell'ordine del primo quartile} + 1}{2} = \frac{\text{intero di } (24,5) + 1}{2} = \frac{24 + 1}{2} = \frac{25}{2} = 12,5.$$

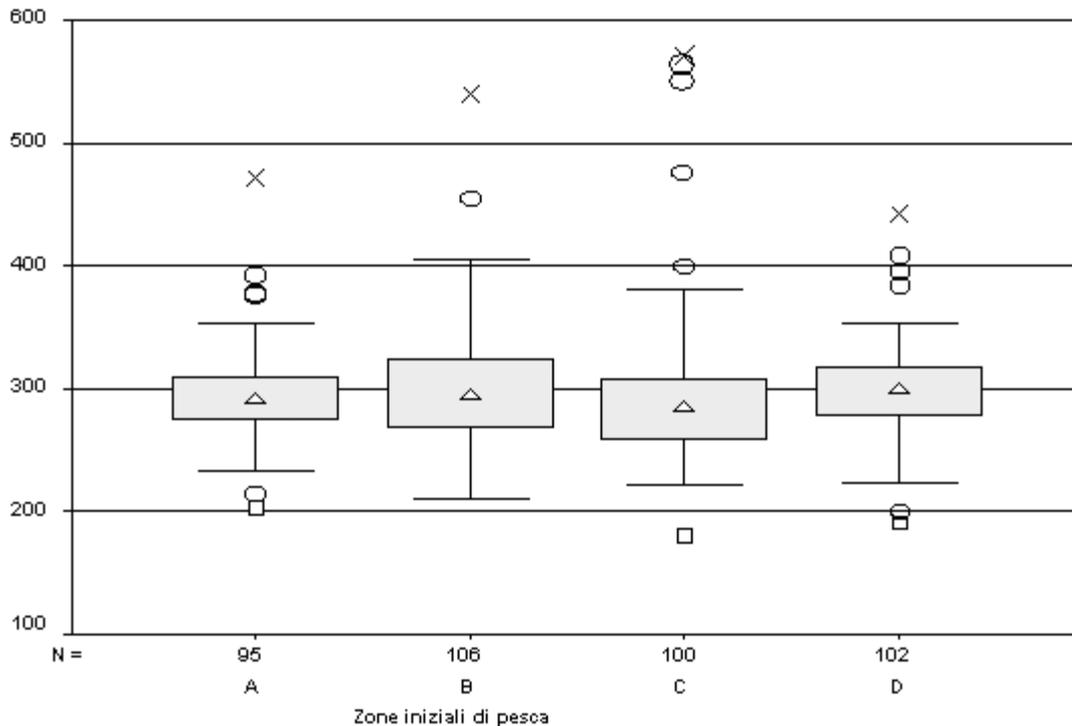
Dunque, anche in questo caso, dobbiamo prendere la semisomma del 12° e 13° termine. Risultano:

$$\text{ottavo inferiore} = \frac{263 + 263}{2} = 263 \text{ grammi}$$

$$\text{ottavo superiore} = \frac{321 + 320}{2} = 320,5 \text{ grammi.}$$

In modo analogo si procede per individuare i due sedicesimi che hanno ordine  $(12+1)/2=6,5$  e risultano quindi dalla semisomma del 6° e 7° termine. Infine, l'intervallo interquartile, che ci dà un'informazione sulla variabilità, in quanto fra i due quartili cade il 50 per cento delle osservazioni, si ottiene come differenza fra  $Q_3$  e  $Q_1$ ; risulta  $H=310-276=34$  grammi.

Figura 4 – Diagramma a scatola o box plot relativo ai dati della figura 3 sul peso medio dei pesci per pescatore, a seconda della zona iniziale di pesca (in grammi)



Vediamo ora cosa possono dirci questi risultati, la cui lettura viene facilitata dal diagramma a scatola della figura 4. Spieghiamo il significato e la costruzione dei diagrammi a scatola. Con un triangolo viene individuata nella figura, per mezzo della scala stabilita per le ordinate, la mediana per ciascuno dei gruppi, e questa mediana risulta contenuta nella scatola i cui lati lunghi corrispondono al 1° e al 3° quartile: dunque la scatola racchiude il 50 per cento delle osservazioni e la sua ampiezza ci dà un'importante informazione sulla loro variabilità (notate che questa ampiezza corrisponde all'intervallo interquartile  $H$ ). Con tratto continuo, a partire dai bordi della scatola, vengono tracciati gli intervalli:

$$\begin{array}{c}
 Q_3 \quad Q_3 + 1,5H \\
 \text{-----} \\
 Q_1 - 1,5H \quad \text{-----} \quad Q_1
 \end{array}$$

al di sopra e al di sotto di essa. Il significato di questi segmenti consiste nell'individuazione di intervalli in cui le osservazioni non si scostano molto dal nucleo centrale, cioè dal 50 per cento che sta al centro, delle osservazioni stesse. Con delle croci e dei quadrati sono poi indicati i valori massimo e minimo i quali cadono in generale al di là degli intervalli appena definiti e si discostano abbastanza dal resto delle osservazioni.

Notiamo subito, sia dal diagramma a scatola sia dai riassunti numerici delle varie zone, che i quattro valori mediani sono molto prossimi l'uno all'altro e che anche la variabilità, misurata dal valore  $H$  e visivamente dall'ampiezza della scatola, non differisce di molto nelle quattro zone. Dunque le trote del lago di Salasco hanno un peso che si aggira intorno a 290 grammi. Si tratta di un'informazione piuttosto approssimativa e grezza, ma che torna utile se non si sa nulla circa la fauna ittica del lago. Infatti il peso di un pesce è in funzione della sua età.

Possiamo ancora cogliere che i casi che cadono al di là dell'intervallo

$Q_1 - 1,5H$  |-----|  $Q_3 + 1,5H$  sono piuttosto pochi e che, in particolare, sono relativamente più numerosi quelli che eccedono il limite superiore dell'intervallo, di quelli che non superano il limite inferiore. Nella zona B, per esempio, due valori sono molto alti, mentre

nessun valore cade al di sotto del limite minimo  $Q_1-1,5H$ , come si può controllare sul diagramma a scatola della figura 4, e sui dati del diagramma ramo-foglia. Quando si va ad analizzare più da vicino questi valori lontani ci si rende conto di una particolarità: essi derivano da casi in cui i pesci pescati sono stati pochi, cioè il peso medio è stato calcolato su uno o due soli pesci pescati. Infatti, è possibile che si verifichino una o due catture di pesci molto piccoli e molto grandi, ma è molto poco probabile che un pescatore peschi cinque o addirittura dieci pesci tutti piccoli o tutti grossissimi.

Soffermiamoci su un'ultima questione. La gara si svolge sia in modo individuale sia a squadre. I risultati, di conseguenza, vengono esplicitati per ogni singolo pescatore, ma anche per squadra, per cui si dispone pure del numero di pesci pescati per ogni squadra e del peso complessivo del pescato, che concorrono a stabilire la classifica a squadre. Nella gara considerata i risultati per le 108 squadre (che non siano 112, come ci si aspetterebbe per quanto detto in precedenza, è dovuto alla circostanza che venivano ammessi alcuni pescatori indipendenti i quali, pur aggregati in gruppi di quattro, non formavano squadra), relativi al numero di pesci catturati da ciascuna squadra e al peso totale, sono stati raccolti nei diagrammi ramo-foglia seguenti:

**Figura 5 - Diagramma ramo-foglia dei pesci pescati per squadra**

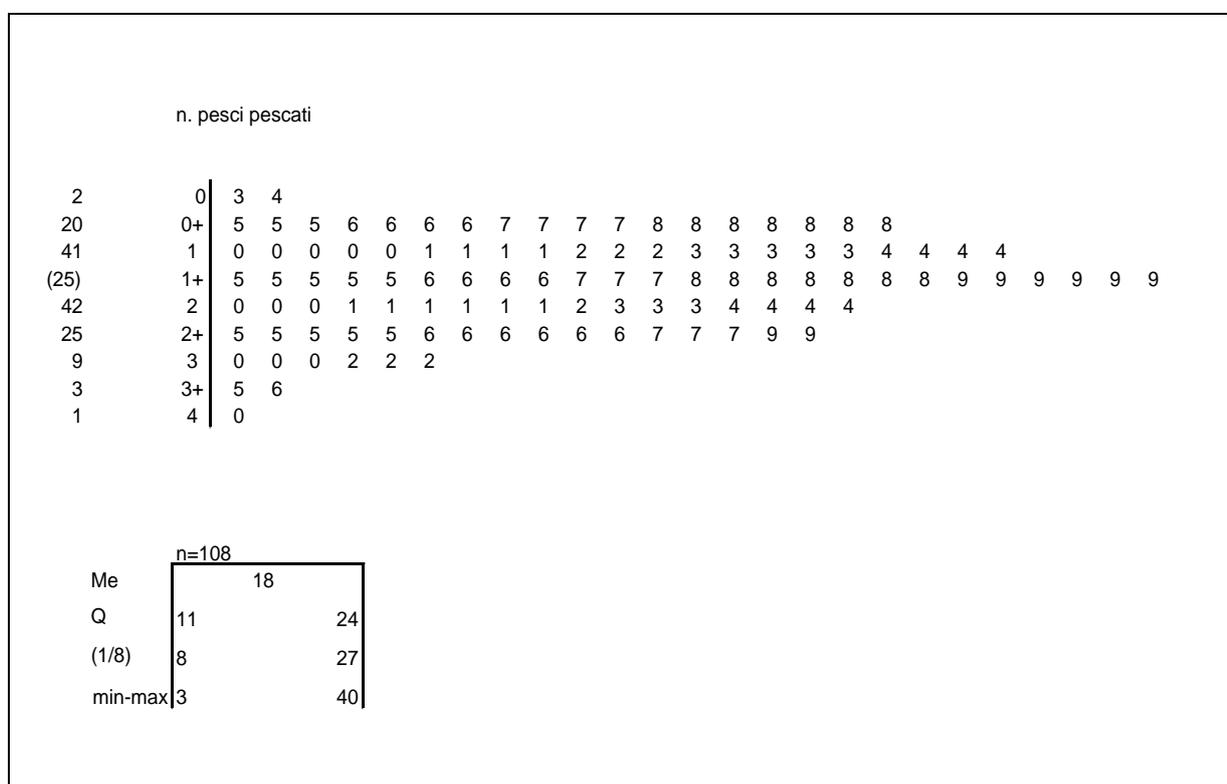
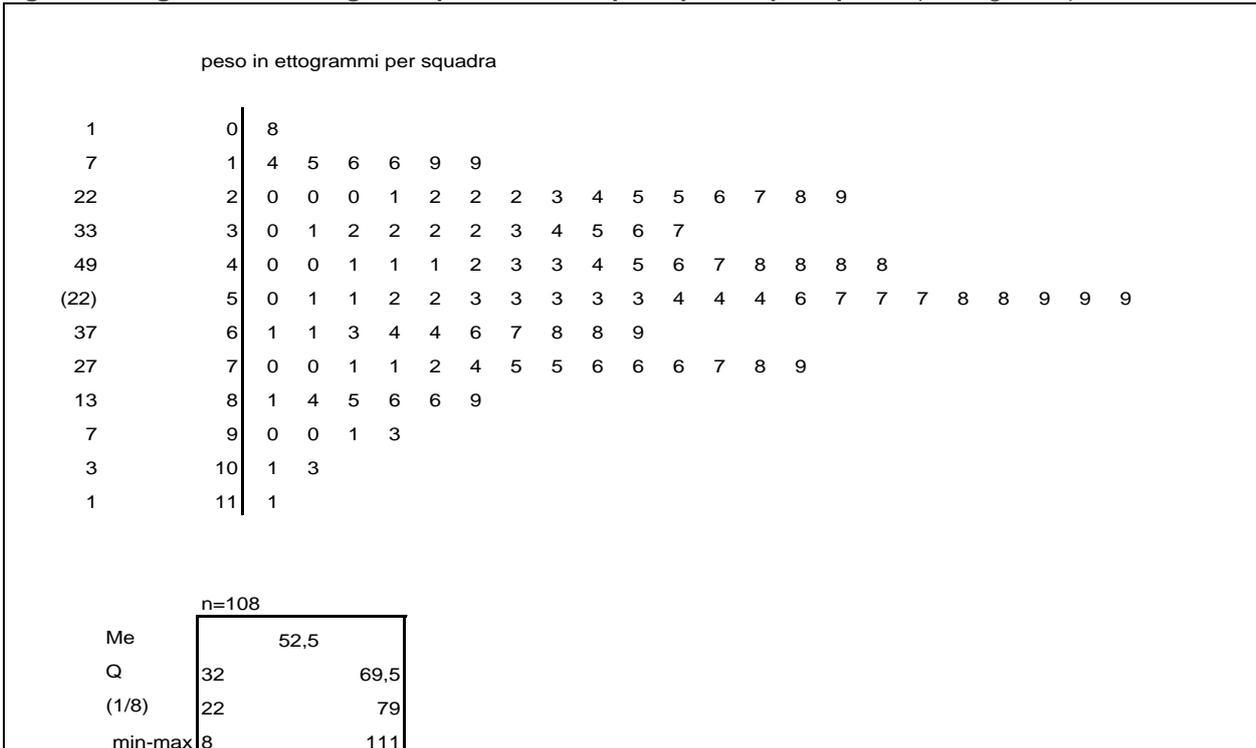


Figura 6 - Diagramma ramo-foglia del peso totale dei pesci pescati per squadra (in ettogrammi)



Da questi diagrammi risulta visivamente abbastanza chiaro come le distribuzioni di frequenza sia del numero di catture, sia del peso del pescato, per ogni squadra, appaiono meno asimmetriche di quelle per singolo pescatore. Il peso tipico di una singola trota può ottenersi dividendo il peso mediano del pescato, 5250 grammi, per il numero di catture 18, che dà 292 grammi (valore del tutto compatibile con quello trovato per altra via).

### Alcune definizioni

Lo **scostamento quadratico medio** è una misura della variabilità delle osservazioni ed è pari alla radice quadrata della varianza. Quest'ultima quantità si calcola come rapporto fra la somma degli scostamenti elevati al quadrato di tutte le osservazioni dalla loro media aritmetica e il numero delle osservazioni stesse. Ovviamente, vale zero se tutte le quantità sono fra loro uguali, ovvero se non c'è variabilità. Esempio: date le quantità 1, 5, 7, 11, 16 la media aritmetica vale  $40/5=8$ .

Gli scarti da tale media valgono

$$1-8=-7$$

$$5-8=-3$$

$$7-8=-1$$

$$11-8=3$$

$$16-8=8$$

e quindi la varianza vale:  $\frac{(-7)^2 + (-3)^2 + (-1)^2 + 3^2 + 8^2}{5} = 26,4$ .

Nel caso che i dati siano forniti in seriazioni allora il calcolo della varianza si modifica leggermente, dovendo far intervenire le frequenze.

La formula, in questo caso, è  $\text{var} = \frac{\sum_{i=1}^k (x_i - m)^2 f_i}{\sum_{i=1}^k f_i}$ ,

dove, come di consueto,  $x_i$  costituiscono i valori centrali delle  $k$  classi in cui la seriazione è organizzata e  $f_i$  sono le frequenze associate.

La **mediana** divide l'insieme delle osservazioni ordinate in due gruppi ugualmente numerosi: se le osservazioni sono in numero dispari allora la mediana coincide con l'osservazione intermedia (ad esempio, se  $n=5$ , la mediana coincide con la terza osservazione); se sono pari, allora risulta pari alla semisomma delle due osservazioni intermedie (ad esempio, se  $n=6$ , la mediana risulta pari alla semisomma della 3<sup>a</sup> e 4<sup>a</sup> osservazione).

I **quartili** (indicati con  $Q_1$  – il primo – e  $Q_3$  – il terzo) sono le mediane delle due parti in cui la mediana generale suddivide l'insieme delle osservazioni ordinate. La loro differenza ( $H=Q_3-Q_1$ ), detta **intervallo interquartile**, indica la variabilità della serie di valori poiché il 50 per cento delle osservazioni cade al suo interno e, dunque, se esso è molto o poco esteso possiamo trarne un'indicazione circa la dispersione dei dati.

I **diagrammi ramo-foglia** costituiscono una comoda forma di presentazione e analisi dei dati quantitativi sin quando essi non sono troppo numerosi (sino a un massimo di 200-300). I vantaggi di questa forma semigrafica di presentazione dei dati possono essere così riassunti:

- portano ad accertare visivamente se vi è simmetria o meno nella disposizione delle osservazioni;
- permettono di individuare immediatamente il massimo e il minimo valore della serie e danno un'idea della dispersione dei valori delle serie;
- rendono agevole il compito di mettere in luce l'esistenza di eventuali poche osservazioni molto lontane dal nucleo più consistente dei valori; circostanza questa che ci spinge a saperne di più sulla loro origine: si tratta di errori di registrazione, oppure di osservazioni che risultano spurie per vari motivi?
- facilitano la possibilità di registrare l'esistenza di intervalli vuoti fra i dati oppure di particolari addensamenti.

Il diagramma ramo-foglia sostituisce anche la rappresentazione grafica del fenomeno, effettuata mediante istogramma (basta ruotarlo di  $90^\circ$  in senso antiorario).

Il **diagramma a scatola** (box plot) viene costruito assumendo come limiti della scatola i due quartili e inserendo nella scatola, con un triangolo, la posizione della mediana. Vengono poi tracciate delle code sino ai punti  $Q_3+1,5H$  e  $Q_1-1,5H$ , per individuare degli intervalli non discosti dal nucleo centrale delle osservazioni (composto dal 50 per cento di esse) che cadono nella scatola. Si possono così mettere a confronto più esperienze ed evidenziare eventuali osservazioni sospette che cadono al di fuori delle code.

## Esercitazioni

- Calcolate, con i dati della tavola 1, il numero medio di catture nelle zone A, B, C, D. Calcolate la varianza e lo scarto quadratico medio per ciascuna seriazione. Discutete con quante cifre decimali è utile assumere la media.
- Commentate l'affermazione, derivante dal caso presentato, che "affinché i dati siano ricchi di notizie e reale fonte di informazione devono essere, di norma, raccolti per uno scopo specifico".
- Può essere utile esercizio di statistica, ma anche di geografia, prendere in considerazione – se c'è un lago nei dintorni – il problema della rilevazione della fauna ittica. Si può chiedere alle associazioni sportive, alla prefettura, alle organizzazioni ambientaliste, se ci sono statistiche della fauna del lago. Se sono disponibili, potete impostare un lavoro di elaborazione dipendente dal genere e dalla quantità di dati disponibili. Provate a metterli in relazione al possibile inquinamento.