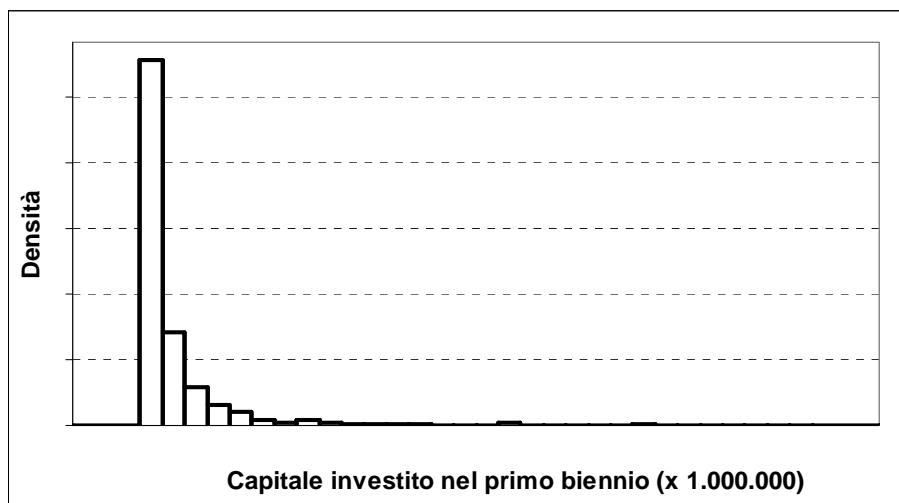


Istogramma: la scelta delle classi

Quando si rappresenta un carattere quantitativo continuo tramite istogramma, uno dei problemi da affrontare è la scelta delle classi da utilizzare per la rappresentazione. Oltre al numero, bisogna anche decidere gli estremi di tali classi, e infine anche se vale la pena rappresentarle tutte o meno. L'idea è di scegliere un numero di classi abbastanza elevato da permettere di comprendere la forma della distribuzione, ma, allo stesso tempo, abbastanza contenuto da consentire una rappresentazione sintetica.

Nel grafico che segue, a scopo esemplificativo, si è deciso di utilizzare per un primo tentativo 30 classi di uguale ampiezza per rappresentare un carattere continuo.

Estr. inf. delle classi	Estr. sup. delle classi	Freq. assolute	Freq. relative	Densità
0.0014	58.74	653	0.651697	0.011095
58.74	117.48	167	0.166667	0.002837
117.48	176.22	69	0.068862	0.001172
176.22	234.96	37	0.036926	0.000629
234.96	293.69	25	0.024950	0.000425
293.69	352.43	9	0.008982	0.000153
352.43	411.17	5	0.004990	0.000085
411.17	469.91	11	0.010978	0.000187
469.91	528.65	5	0.004990	0.000085
528.65	587.39	3	0.002994	0.000051
587.39	646.12	3	0.002994	0.000051
646.12	704.86	2	0.001996	0.000034
704.86	763.60	3	0.002994	0.000051
763.60	822.34	0	0	0
822.34	881.08	0	0	0
881.08	939.82	0	0	0
939.82	998.55	5	0.004990	0.000085
998.55	1057.29	0	0	0
1057.29	1116.03	0	0	0
1116.03	1174.77	0	0	0
1174.77	1233.51	0	0	0
1233.51	1292.25	0	0	0
1292.25	1350.98	2	0.001996	0.000034
1350.98	1409.72	1	0.000998	0.000017
1409.72	1468.46	1	0.000998	0.000017
1468.46	1527.20	0	0	0
1527.20	1585.94	0	0	0
1585.94	1644.68	0	0	0
1644.68	1703.42	0	0	0
1703.42	1762.15	1	0.000998	0.000017



Notiamo che la prima classe, modale, rappresenta circa il 65% della popolazione considerata. Quindi potrebbe essere utile procedere raggruppando i dati in classi di intervallo più sensate, riunendo la lunga coda destra in un'unica classe, senza perdita di informazioni, e disaggregando la prima classe.

Una scelta quindi più sensata potrebbe essere la seguente, dove l'ultima classe è però stata esclusa dalla rappresentazione.

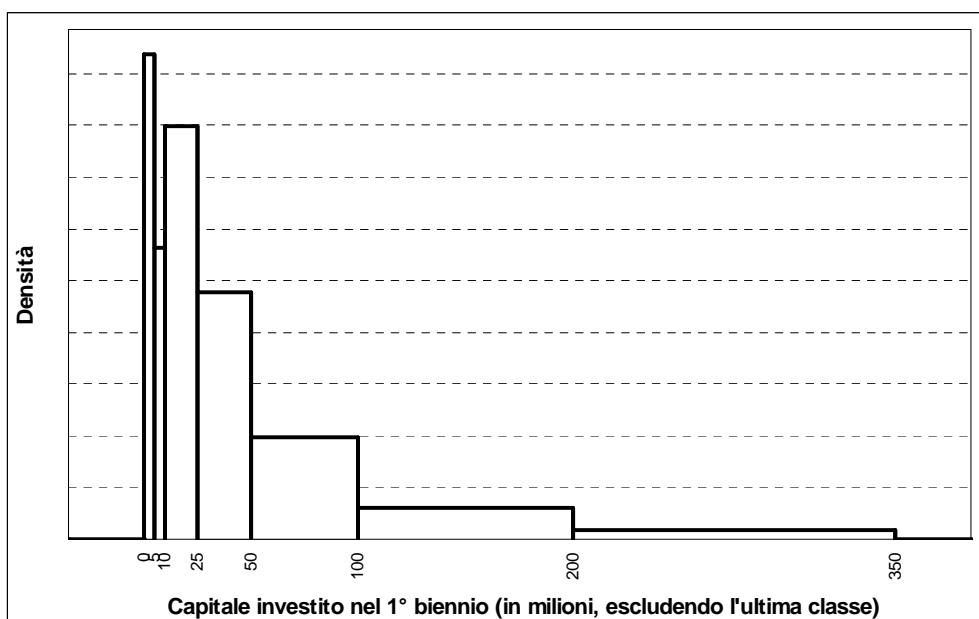


Tabella delle frequenze relative all'istogramma (classi pre-assegnate)

<i>Estr. inferiori</i>	<i>Estr. superiori</i>	<i>Fr. assolute</i>	<i>Fr. relative</i>	<i>Densità</i>
0	5	90	0.08982036	0.0179641
5	10	54	0.05389222	0.0107784
10	25	230	0.22954092	0.0153027
25	50	229	0.22854291	0.0091417
50	100	189	0.18862275	0.0037725
100	200	116	0.11576846	0.0011577
200	350	51	0.0508982	0.0003393
350	1763	43	0.04291417	3.037E-05

In alternativa, un modo semplice ma efficace per comprendere la forma della distribuzione consiste nel considerare, come estremi delle classi, i decili della distribuzione. In questo modo, già abbiamo un'idea della forma del carattere, e possiamo in seguito decidere di raggruppare o scomporre alcune classi di particolare interesse.

Qui di seguito, ecco l'istogramma risultante selezionando, come estremi delle classi, i decili della distribuzione, in cui però l'ultima classe è stata scorporata in due e la seconda parte non è stata rappresentata.

Tabella delle frequenze relative all'istogramma (decili)

<i>Estr. inferiori</i>	<i>Estr. superiori</i>	<i>Fr. assolute</i>	<i>Fr. relative</i>	<i>Densità</i>
0	6	105	0.104790419	0.01746507
6	13	93	0.092814371	0.013259196
13	20	108	0.107784431	0.015397776
20	26	88	0.087824351	0.014637392
26	35	99	0.098802395	0.010978044
35	50	110	0.109780439	0.007318696
50	70	98	0.097804391	0.00489022
70	103	102	0.101796407	0.00308474
103	190	97	0.096806387	0.001112717
190	350	59	0.058882236	0.000368014
350	1763	43	0.042914172	3.0371E-05

