Statistica I

Esercitazione 3: variabilità, istogrammi, boxplot, simmetria, curtosi

Tommaso Rigon

Università Milano-Bicocca

La saggezza della folla è affidabile?

- Nel 1907 lo scienziato ed inventore **Francis Galton**, cugino di Charles Darwin, scrisse una lettera alla prestigiosa rivista scientifica **Nature**, intitolata "*Vox Populi*".
- Francis Galton era stato ad una mostra di bestiame. Era stato indetto un **concorso** per indovinare il peso della carne, dopo la macellazione, di un **grande bue**.
- La partecipazione costava 6 penny e Galton riuscì a procurarsi ben $n = 787$ dei biglietti acquistati. Calcolò la **media** delle varie stime, ovvero 547 kg.
- Il valore medio era notevolmente vicino al **peso reale** di 543 kg, sebbene la maggior parte dei concorrenti avesse fornito una stima molto meno precisa.
- Questo metodo per prendere decisioni è spesso chiamato "**saggezza della folla**".
- Siamo interessati a verificare **se** il fenomeno della saggezza della folla sia replicabile, tramite un esperimento.

Galton, F. (1907), Vox populi, *Nature*, 1949(75)

VOX POPULI.

IN these democratic days, any investigation into the trustworthiness and peculiarities of popular judgments is of interest. The material about to be discussed refers to a small matter, but is much to the point.

A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors bought stamped and numbered cards, for 6*d*. each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. These afforded excellent material. The judgments were unbiassed by passion and uninfluenced by oratory and the like. The sixpenny fee deterred practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge justly was probably much the same in either case.

After weeding thirteen cards out of the collection, as being defective or illegible, there remained 787 for discussion. I arrayed them in order of the magnitudes of the estimates, and converted the *cwt.*, *quarters*, and *lbs.* in which they were made, into lbs., under which form they will be treated.

Distribution of the estimates of the dressed weight of a particular living ox, made by 787 different persons.

Degrees of the length of Array 0°—100°	Estimates in lbs.	Centiles: Observed deviates from 1207 lbs.	Centiles: Normal p.e =37	Excess of Observed over Normal
5	1074	−133	−90	+43
10	1109	− 98	−70	+28
15	1126	− 81	−57	+24
20	1148	− 59	−46	+13
q_1 25	1162	− 45	−37	+ 8
30	1174	− 33	−29	+ 4
35	1181	− 26	−21	+ 5
40	1188	− 19	−14	+ 5
45	1197	− 10	− 7	+ 3
m 50	1207	0	0	0
55	1214	+ 7	+ 7	0
60	1219	+ 12	+14	− 2
65	1225	+ 18	+21	− 3
70	1230	+ 23	+29	− 6
q_3 75	1236	+ 29	+37	− 8
80	1243	+ 36	+46	−10
85	1254	+ 47	+57	−10
90	1267	+ 52	+70	−18
95	1293	+ 86	+90	− 4

q_1, q_3, the first and third quartiles, stand at 25° and 75° respectively.
m, the median or middlemost value, stands at 50°.
The dressed weight proved to be 1198 lbs.

- **Articolo**: `https://galton.org/essays/1900-1911/galton-1907-vox-populi.pdf`

La bottiglia con le biglie di vetro

- Ho riempito una bottiglia con delle **biglie** e ho chiesto alla classe quante fossero.

Informazioni aggiuntive

- Ho quindi posto nuovamente la domanda agli studenti, fornendo però le seguenti **informazioni aggiuntive**. Gli studenti potevano **rivedere** la loro stima.

Informazioni aggiuntive

- La formula per il calcolo del **volume di un cilindro** è:

$$(\text{Volume}) = (\text{Area di base}) \times (\text{Altezza}),$$

  dove l'area di base, ovvero l'**area del cerchio**, è pari a:

$$(\text{Area di base}) = \pi(\text{Diametro}/2)^2.$$

- La bottiglia è stata agitata varie volte durante il riempimento, per ridurre il più possibile gli spazi vuoti tra le biglie.

- La bottiglia contiene approssimativamente **1 litro**. Le biglie di vetro sono tutte uguali tra di loro e hanno diametro di **16 mm**, ovvero **2.144 ml** ciascuna.

- È inoltre **noto** (`https://www.nature.com/articles/nature06981`) che l'impacchettamento casuale di sfere ha una **densità** compresa tra il 55% ed il 64%.