

ANALISI DELLA DISTORSIONE DEI DATI NELL'ELABORAZIONE DEI RISULTATI DELLA VALUTAZIONE DELLA DIDATTICA E PROPOSTE ALTERNATIVE

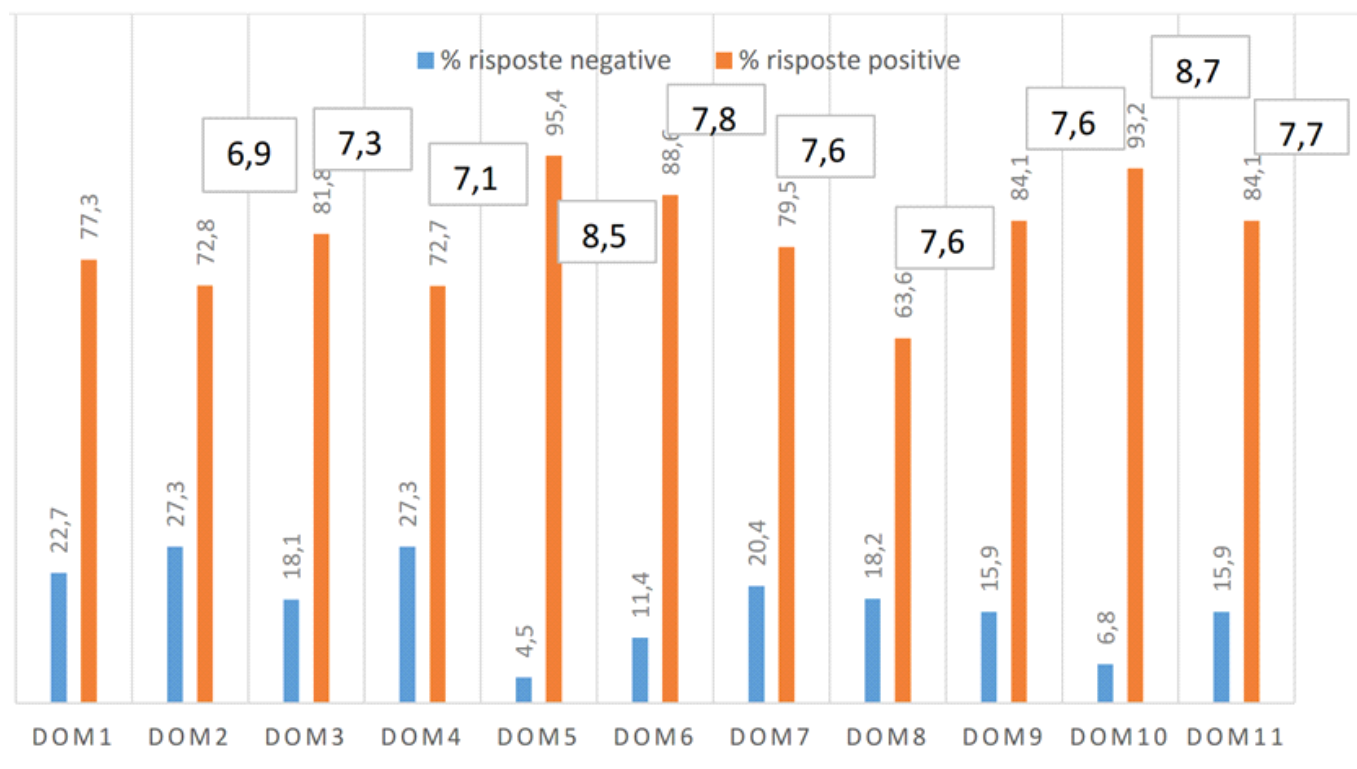
Ho utilizzato alcuni dati della rilevazione delle opinioni degli studenti 2014-15 per mostrare graficamente quanto da alcuni anni provo a spiegare relativamente all'arbitrarietà delle valutazioni applicate dall'ufficio statistica.

Il grafico è stato costruito utilizzando le % di risposta emerse dai questionari e sommando da una parte le valutazioni positive (decisamente sì, più sì che no), dall'altra quelle negative (più no che sì, decisamente no). E questo è quello che emerge.

Sopra ciascuna coppia di colonnine in un riquadro è riportata la media relativa a ciascuna domanda così come calcolata dall'ufficio statistica che, invece di lavorare sulle percentuali, attribuisce un punteggio arbitrario a ciascuna alternativa di risposta.

L'esito distorto della procedura ad oggi vigente è palese. Si guardino le domm. 7, 8 e 9. La media fornita dagli uffici è la stessa (7,6) per tutte e tre le domande; questa cifra però emerge da configurazioni di dati parecchio diverse perché la domanda 8 ha ottenuto una % di risposte positive di 63,6 mentre la dom.9 una ben più elevata: l'84,1%. Ma questa significativa differenza non viene registrata dalla media, che opera appiattendolo tutto.

Inoltre, per fare solo un altro esempio, la dom. 5 ha una media di 8,5 pur avendo ottenuto il 95,4% di risposte positive. Il che mi pare una significativa stortura, oltretutto al ribasso.



Dal punto di vista della correttezza metodologica, una variabile categoriale ordinata, quale è la scala di giudizio proposta nel questionario, non può essere trattata trasformandola in variabile cardinale, attraverso punteggi (qualsiasi essi siano).

La mia opinione è che la procedura metodologicamente corretta di costruire e leggere i dati di valutazione basati su quella batteria di domande sia lavorare sulle percentuali, e non attribuire punteggi arbitrari ex post alle risposte. La lettura tramite le percentuali può essere peraltro effettuata in due versioni: a 4 categorie e a 2 categorie (come fatto nell'esempio sopra illustrato). Ciò rende possibile una prima rappresentazione fedele delle opinioni espresse e una seconda rappresentazione di sintesi. Entrambe sono perfettamente funzionali anche ad una comparazione diacronica. Comunque, l'aggregazione in 2 categorie viene realizzata senza operare insostenibili forzature dal punto di vista semantico, come invece accade con l'assegnazione stipulativa dei punteggi.

Si potrà obiettare che sarebbe più corretto attribuire un peso statistico differente ai giudizi estremi (decisamente sì, decisamente no) rispetto a quelli intermedi; opzione perseguibile a patto di stabilire delle soglie che in definitiva risulterebbero arbitrarie anch'esse, se non supportate da adeguata e condivisa riflessione teorica.

Spero che questo banale esempio possa contribuire a fare chiarezza e contribuire alla ricerca di soluzioni.

Stefania Tusini

con la collaborazione di Rolando Marini

4 ottobre 2016