

TOL DAY 2006

PROGETTO TOL GAME: ANALISI DEI DATI

Report intervento
Milano, 19 Maggio 2006
Anna Maria Paganoni

Nell'ambito del progetto TOL GAME sono stati somministrati in 25 classi di licei scientifici 3 questionari (in 3 differenti giornate) ciascuno costituito da 10 domande a risposta multipla (con 5 possibili risposte, di cui una sola corretta). Le domande sono raggruppabili per argomento nel seguente modo:

Algebra:	7 domande
Aritmetica:	5 domande
Geometria:	6 domande
Logica:	7 domande
Trigonometria:	5 domande

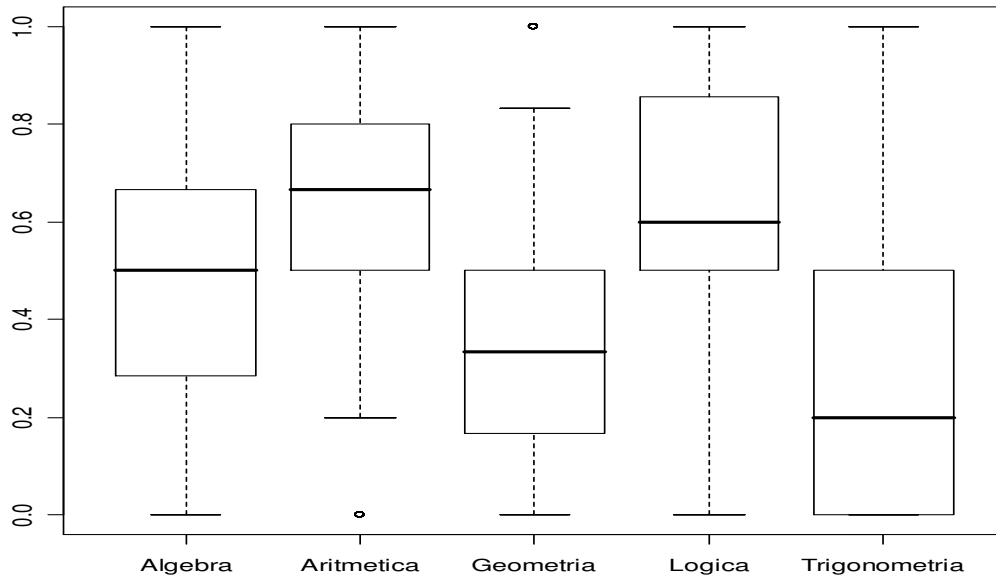
Il totale di unità statistiche (studenti) cui è stato somministrato almeno un questionario è di 827.

Scopo dell'analisi descrittiva dei dati eseguita e sintetizzata in questo intervento è confrontare la distribuzione delle risposte degli studenti nei 5 diversi settori.

Dall'analisi sono stati tolti (all'interno di ogni diverso argomento) gli assenti durante il giorno di somministrazione del questionario.

Nel grafico seguente sono riportati i box-plot delle percentuali di risposte corrette suddivisi per settore.

Percentuali di risposte corrette



Questa prima analisi descrittiva di confronto evidenzia alcune differenze tra le cinque distribuzioni. In particolare si osserva come la mediana della percentuale di risposte corrette in Trigonometria e in Geometria sia inferiore rispetto alle altre materie. Si segnala un problema nel caso dell'Aritmetica in cui la numerosità del campione è inferiore (636) dato che le domande di Aritmetica erano presente solo nel secondo e nel terzo questionario. I ragazzi che hanno partecipato anche ai successivi questionari e non solo al primo potrebbero essere una sottopopolazione selezionata del campione esaminato (forse i più bravi e motivati) e questo potrebbe, in parte, spiegare la superiorità della distribuzione di risposte corrette in Aritmetica che si osserva nel confronto tra i boxplot.

Si vuole ora eseguire un'analisi descrittiva di confronto che associ ad ogni unità statistica (ragazzo) un vettore (p_1, p_2, p_3) dove:

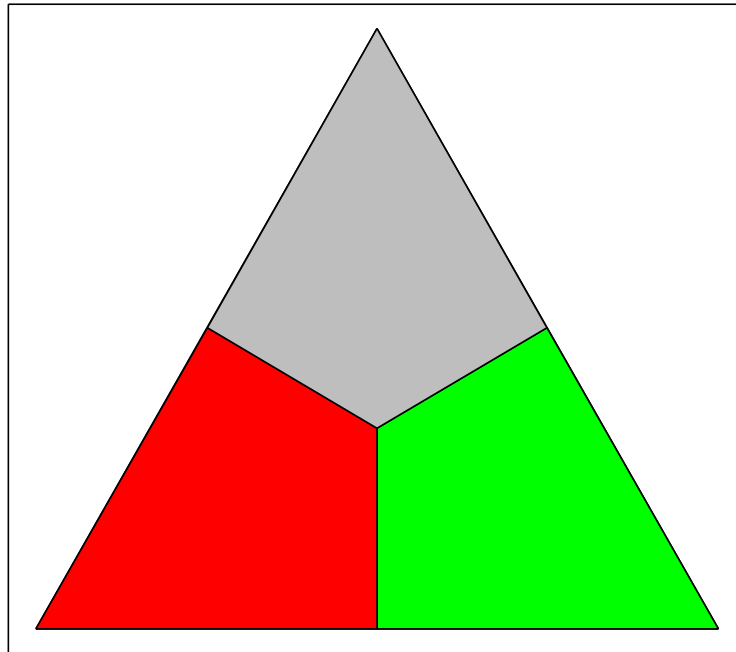
p_1 = percentuale di risposte corrette

p_2 = percentuale di risposte errate

p_3 = percentuale di non risposte

infatti si ricorda che anche se presente un ragazzo può anche scegliere di non rispondere ad una o più domande.

Dato che $p_1 + p_2 + p_3 = 1$ ogni punto appartiene alla faccia superiore del semplice tridimensionale di cui si può dare una rappresentazione visiva bidimensionale con il seguente triangolo:



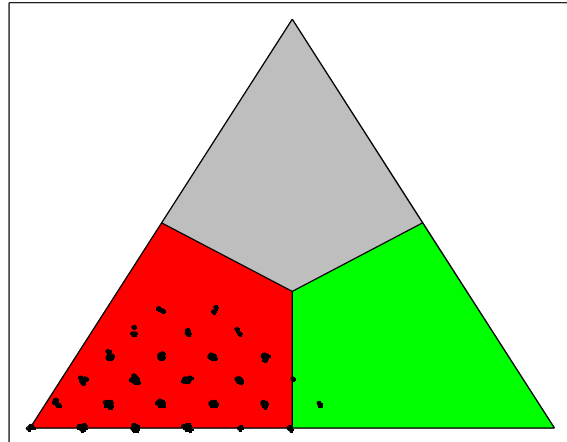
In questa rappresentazione tutti i punti che giacciono nella regione verde sono quelli per cui $\text{Max}\{p_1, p_2, p_3\} = p_1$ (prevalgono le risposte corrette), i punti che giacciono nella regione rossa sono quelli per cui $\text{Max}\{p_1, p_2, p_3\} = p_2$ (prevalgono le risposte errate), ed infine quelli che giacciono nella regione grigia sono quelli per cui $\text{Max}\{p_1, p_2, p_3\} = p_3$ (prevalgono le non risposte).

La percentuale di unità statistiche che cadono nella regione verde può quindi essere una misura della preparazione degli studenti e nel seguito verrà denominata IP (indice di preparazione).

Si osservi inoltre che sul lato superiore sinistro del triangolo giacciono tutti i punti per cui $p_1 = 0$ (coloro che non hanno mai dato una risposta corretta), sul lato superiore destro del triangolo giacciono tutti i punti per cui $p_2 = 0$ (coloro che non hanno mai dato una risposta errata) ed infine sul lato inferiore del triangolo giacciono tutti i punti per cui $p_3 = 0$ (coloro che hanno sempre dato una risposta, corretta o meno).

Per meglio valutare tale strumento descrittivo segue un esempio di simulazione di 200 studenti che rispondessero in modo completamente casuale scegliendo con ugual probabilità o di non rispondere oppure una delle cinque possibili risposte. Tale simulazione viene visualizzata nel seguente modo:

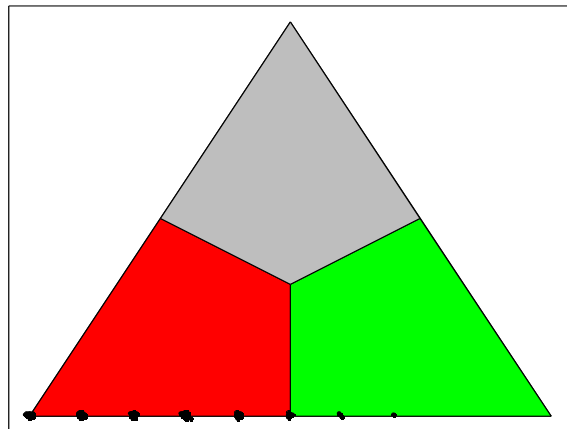
Casuale con non risposta



Si osservi che la quasi totalità della popolazione giace nella zona rossa.

Analogamente un esempio di simulazione di 200 studenti che rispondessero in modo completamente casuale scegliendo con ugual probabilità una delle cinque possibili risposte (ma comunque dando sempre una risposta) viene rappresentato nel seguente modo:

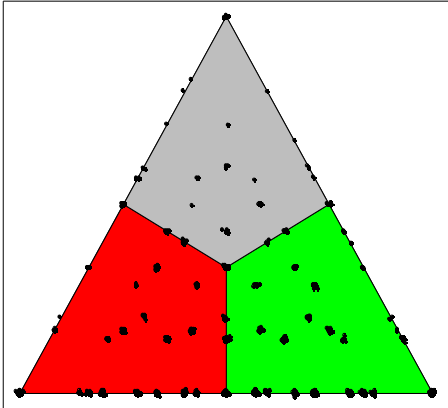
Casuale senza non risposta



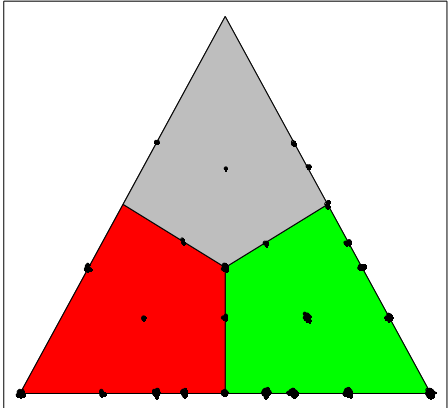
Si osservi che la quasi totalità della popolazione giace nella zona rossa e comunque sul lato inferiore del triangolo.

I triangoli (con relativi indici di preparazione) che visualizzano le distribuzioni delle risposte nelle cinque materie affrontate nell'indagine nelle scuole sono i seguenti:

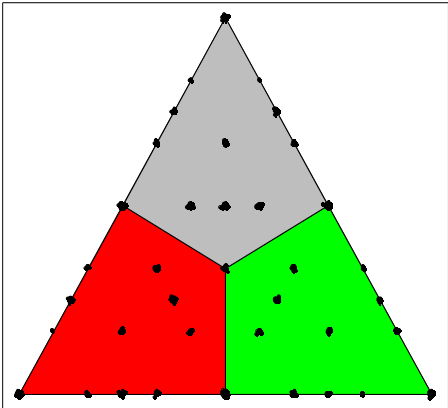
Algebra



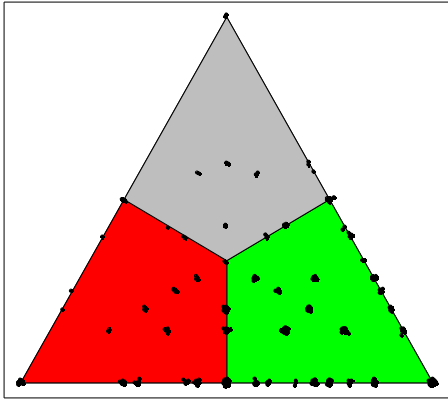
Aritmetica



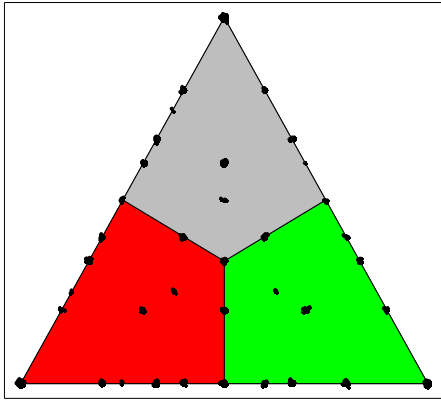
Geometria



Logica



Trigonometria



IP

- Algebra: 0.61
- Aritmetica: 0.82
- Geometria: 0.45
- Logica: 0.86
- Trigonometria: 0.33