

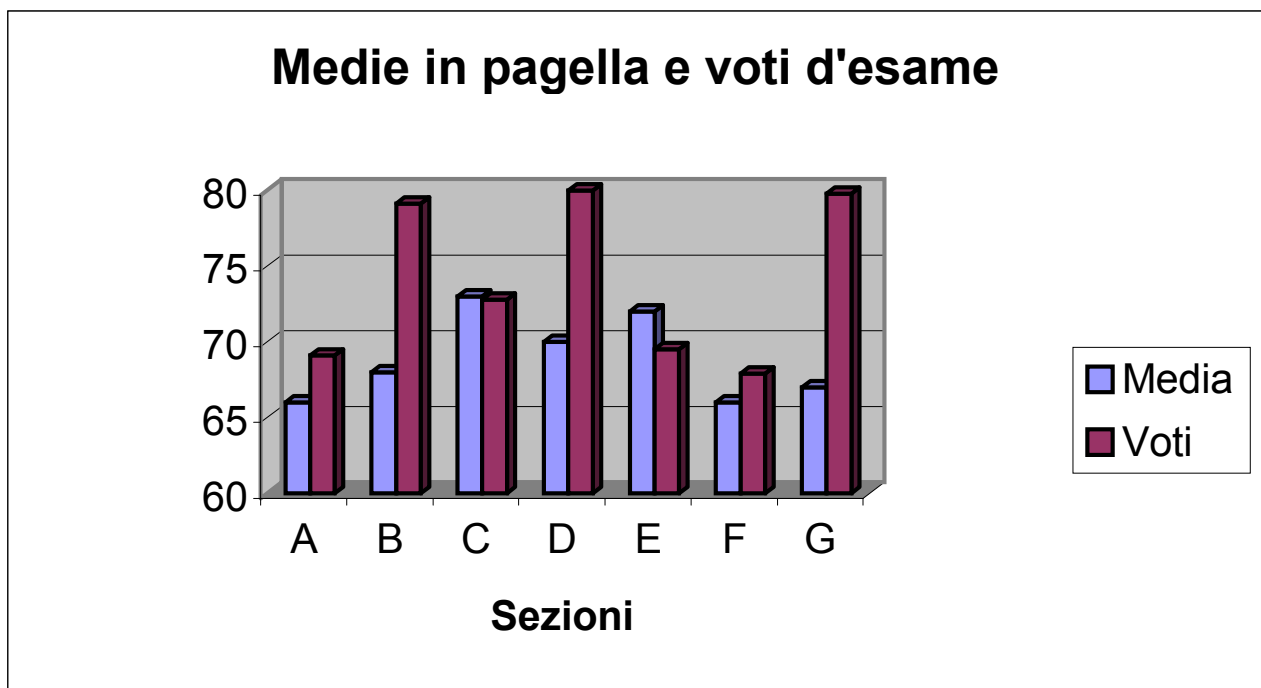
Liceo Scientifico Statale
Galileo Ferraris
Varese

Esame di Stato 1999

Le commissioni e il caso

Analisi della correlazione tra la media dei voti e il voto d'esame

Prof. Michele Impedovo



Introduzione

Sono stato commissario interno agli esami di Stato del 1999. Durante lo svolgimento degli esami si è osservata una certa difformità nello svolgimento e nell'organizzazione degli esami stessi da parte delle diverse commissioni. Uno sguardo anche distratto ai tabelloni degli esiti finali ha confermato questa impressione: salta all'occhio che le commissioni delle sezioni B e D e della sezione G hanno operato e valutato con generosità sensibilmente maggiore rispetto alle commissioni delle sezioni A e C e delle sezioni E e F. Un primo indizio è il numero dei bocciati: : nessuno studente bocciato nelle classi B, D e G, 10 studenti bocciati nelle altre. Ecco nel dettaglio la tabella con il numero di bocciati per classe.

V A	V B	V C	V D	V E	V F	V G
1	0	2	0	2	5	0

Incuriosito da questa disomogeneità dei risultati, che non trova apparenti giustificazioni nella preparazione media degli studenti (non si evidenziano differenze significative nella preparazione generale delle diverse classi: la media in pagella va dal 6.6 per la V F al 7.3 per la V C) ho raccolto i dati in questione (media dei voti in pagella allo scrutinio finale e voti d'esame) e ho provato ad analizzare la vicenda da un punto di vista più razionale.

I risultati dell'analisi statistica sono sorprendenti, e confermano che le diverse commissioni hanno valutato in modi estremamente differenti: è inequivocabile che gli studenti delle sezioni B, D e G hanno tratto vantaggio dall'essere capitati con quelle commissioni.

L'analisi statistica

I dati iniziali, divisi per classe, sono semplicemente:

- M = media (in decimi) dei voti di ciascuno studente allo scrutinio finale di quinta, quella stessa media dei voti che è servita ai consigli di classe per determinare il credito scolastico, come risulta dalle schede dei candidati.
- V = voto finale (in centesimi) all'esame di Stato, come risulta dai tabelloni.

Per ogni classe sono stati calcolati i seguenti parametri statistici:

1. Media di M

Si tratta del *valore medio* di ogni singola classe, cioè della somma delle medie divisa per il numero di studenti della classe. Come si è già detto, non emergono differenze significative. La classe V C risulta la classe migliore; come vedremo i voti dell'esame non rispecchiano in alcun modo questo dato iniziale.

V A	V B	V C	V D	V E	V F	V G
6.6	6.8	7.3	7.0	7.2	6.6	6.7

2. Media di V

V A	V B	V C	V D	V E	V F	V G
69.1	79.1	72.8	80.0	69.5	67.9	79.8

Si possono già osservare, in media, vistose differenze. Partiamo dall'ipotesi iniziale che al voto in decimi di M possa corrispondere, in prima istanza, un voto d'esame in centesimi pari a $10M$ (per esempio, chi si presenta con una media di 6.8 dovrebbe avere 68 come voto finale); non è un'ipotesi del tutto fondata dato che, si sa, gli insegnanti sono restii ad usare la scala dei voti fino al 10. Tuttavia è ragionevole, e ci serve per avere un riferimento omogeneo di confronto: come si sono

comportate le commissioni rispetto a questa ipotesi? La tabella successiva risponde a questa domanda.

3. Scarto medio di V rispetto alla retta $y=10x$

V A	V B	V C	V D	V E	V F	V G
+ 3	+ 10.7	+ 0.2	+ 10.0	- 2.5	+ 2.2	+ 12.4

Come si vede, le difformità sono addirittura impressionanti. Le classi B, D, G (in media) ottengono rispettivamente 10.7, 10.0, 12.4 punti in più rispetto alla media moltiplicata per 10. Salta all'occhio inoltre che in tutte le classi (tranne la V E) le commissioni aggiungono punti, a dimostrazione del fatto che le commissioni d'esame *pensano* in centesimi assai più di quanto i consigli di classe *pensino* in decimi. Il risultato della V E è il peggiore, si tolgono in media 2.5 punti (una spiegazione potrebbe essere che la classe è stata sopravvalutata dal consiglio di classe, o che comunque ha fornito in sede d'esame prestazioni inferiori a quelle preventivate).

È difficile pensare che tali risultati medi siano dovuti al comportamento in sede d'esame degli studenti: si tratta di un dato che caratterizza fortemente l'operato della commissione. In altri termini: nella sezione E il candidato medio ha avuto 2.5 punti in meno rispetto alla propria media; nella sezione G 12.4 punti in più.

4. Grafico a dispersione dei punti (M_i, V_i)

Ogni studente è rappresentato mediante un punto sul piano cartesiano. In ascissa la media dei voti, in ordinata il voto d'esame: il grafico per punti è uno degli strumenti più potenti del matematico. Molto più che una tabella di dati, un grafico visualizza con forte evidenza l'andamento collettivo dei dati. Si noterà che in tutte le classi i punti sono disposti *all'incirca* lungo una retta (i matematici dicono che c'è una sensibile *correlazione lineare*): mancano punti in alto a sinistra (corrisponderebbero a studenti con una media bassa che hanno ottenuto un voto alto) e mancano punti in basso a destra (gli studenti con media alta non hanno ottenuto voti bassi). Questa distribuzione dei voti fa pensare ad un andamento lineare: all'aumento della media corrisponde un aumento in proporzione (lineare) del voto d'esame. Una volta ipotizzato un modello lineare per la distribuzione dei punti, la statistica consente di determinare l'equazione della *miglior retta* che approssima quei punti (è la retta *dei minimi quadrati*): in un certo senso è la retta che si adatta meglio di qualunque altra ai punti dati, e può riassumere da sola l'intero esame. La prossima tabella illustra queste equazioni.

5. Retta di regressione lineare $V=a+bM(x-6)$

Determinare l'equazione di questa retta ci consente, d'ora in poi, di parlare di *comportamento medio* della commissione in una data sezione. Possiamo idealmente cancellare tutti i punti, e prendere questa retta come rappresentante della valutazione media operata. Per esempio, la retta di regressione per la classe V A è

$$\text{Voto} = 62.6 + 10.8 (\text{Media} - 6).$$

Questo significa che la commissione mediamente ha assegnato 62.6 punti a chi aveva 6 come media dei voti allo scrutinio, e per ogni punto in più di media allo scrutinio ha dato 10.8 punti in più come voto d'esame. Il numero 62.6 ci dice, in sostanza, che la commissione è *partita* da 62.6 come livello di sufficienza. Il numero 10.8 (è la *pendenza* della retta di regressione) ci dà un'idea qualitativa dell'incremento dei voti al crescere della media. La retta di regressione ci dice in sostanza, per una certa media di scrutinio, quale voto d'esame è stato conseguito in media, rispetto alla valutazione complessiva della classe. Nella tabella seguente sono riportati per ogni classe, rispettivamente il voto assegnato alla media 6 e l'incremento per ogni punto in più (la pendenza).

V A	V B	V C	V D	V E	V F	V G
62.6, 10.8	66.7, 14.8	51.9, 16.6	64.6, 15.2	54.7, 12.3	57.9, 17.3	69.2, 14.4

Come si vede, ancora una volta le sezioni B, D, e in particolare la G eccellono. La differenza è vistosa: in V G la commissione è partita addirittura da 69.2, mentre in C soltanto da 51.9: questo significa che, mediamente, il candidato che si è seduto davanti alla commissioni B/D o G con una media sufficiente poteva contare su un voto minimo di 69.2. Chi si è seduto davanti alla commissione A/C partendo da una media sufficiente ha dovuto guadagnarsi in proprio oltre 8 punti per raggiungere il 60.

Per fare solo un esempio uno studente che si presentava all'esame con una media del 7 e che ha preso 68.5 nella sezione C avrebbe preso 83.6 nella sezione G.

6. Coefficiente di correlazione lineare tra M e V

Il coefficiente di correlazione lineare tra due grandezze ci dice quanto è buona l'ipotesi che esse si sviluppino all'incirca lungo una retta, cioè all'aumento dell'una corrisponda un aumento proporzionale dell'altra. Nella tabella seguente sono riportati i coefficienti di correlazione espressi in percentuale.

V A	V B	V C	V D	V E	V F	V G
94%	92%	88%	97%	84%	91%	95%

Come si vede sono tutti molto alti. Stupisce invece il dato della sezione C e della sezione E, in cui evidentemente lo sviluppo dei voti non è stato proprio lineare. In particolare in E, anche osservando il grafico della retta di regressione, sembra di capire che la commissione ha valutato in modo diverso: ha penalizzato i voti bassi e supervalutato quelli alti.

Ma complessivamente il modello lineare è altamente credibile. Questo significa che complessivamente per il candidato medio non ci sono state sorprese: gli studenti migliori hanno conseguito voti proporzionalmente più elevati. L'unica sorpresa è costituita dalla commissione assegnata.

7. Confronto tra i grafici delle retta di regressione e della retta $V=10M$

Come abbiamo visto un livello di confronto omogeneo per valutare l'operato delle commissioni consiste nel moltiplicare per 10 la media M di ciascuno studente. Possiamo confrontare dunque l'operato medio della commissione, sintetizzato dalla retta di regressione, con la retta $V=10M$: questa retta assegna 60 alla media 6, e per ogni punto in più di media assegna 10 centesimi in più.

I corrispondenti grafici sono nelle pagine allegate, divise per classe; la retta $10M$ è tratteggiata, la retta di regressione è a tratto continuo. Anche dal punto di vista grafico emerge una sensibile differenza tra la posizione reciproca delle due rette nelle diverse classi.

Nelle classi B, D, G la retta di regressione sta sopra la retta $10M$ in tutto l'intervallo che ci interessa: questo significa che tutti gli studenti di queste classi conseguono un voto d'esame (in centesimi) superiore alla loro media (in decimi): tutti i punti stanno sopra la retta $10M$.

Nelle altre classi accadono cose diverse. In A 4 studenti su 25 prendono un voto inferiore alla media; in C 11 su 25, in F 11 su 26, in E addirittura 16 su 25.

Sono particolarmente significativi i grafici delle classi C ed E. Il punto di intersezione tra la retta di regressione e la retta $10M$ ha ascissa rispettivamente 7.2 e 8.3: questo significa che gli studenti con media inferiore a 7.2 in C e 8.3 in E (in entrambi i casi la gran parte della classe) sono stati penalizzati, cioè hanno avuto un voto inferiore alla loro media moltiplicata per 10. Per contro, i (pochi) studenti migliori hanno avuto valutazioni molto alte.

Questo dato è implacabile: nelle sezioni B, D, G le commissioni hanno valutato gli studenti con punteggi più alti di quanto abbiano fatto i loro insegnanti, nelle altre sezioni con punteggi parecchio

più bassi. In particolare la sezione E è stata abbondantemente sottovalutata rispetto ai valori fissati dal consiglio di classe.

8. Scarto quadratico medio di V rispetto alla retta di regressione lineare

La statistica, tra le altre cose, si occupa di *sintetizzare* i dati, quando questi sono in grande quantità. Uno dei concetti essenziali, come tutti sanno, è quello di *media*: per esempio possiamo sintetizzare l'esame dell'intero Liceo G. Ferraris dicendo che la media dei voti d'esame è stata 74 (non male, quindi); un altro concetto importante è quello di *dispersione*. La media è sempre 74 sia nel caso in cui tutti gli studenti prendano 74 (dispersione=0) sia nel caso in cui la metà di essi prenda 100, e l'altra metà 48. Con quale dispersione i voti si sono distribuiti intorno alla retta di regressione? Cioè: i voti reali quanto lontano si sono collocati dal voto "medio" assegnato dalla legge lineare? Lo strumento statistico in questione è lo *scarto quadratico medio*, riportato per ogni classe nella seguente tabella. Per esempio, nella classe V A, dove la media dei voti finali è stata 69.1, lo scarto quadratico medio è stato 4.6: ciò significa che gli alunni di questa classe ha preso voti che sono mediamente di 4.6 punti maggiori o minori del voto stabilito dalla retta di regressione. Di seguito è riportata la tabella relativa.

V A	V B	V C	V D	V E	V F	V G
4.6	4.8	6.6	3.6	7.4	6.2	4.2

Lo scarto rispetto alla retta di regressione è minimo in D, con 3.6 (circa il 5% della media dei voti) e massimo in E con 7.4 (circa il 10% della media dei voti). In D c'è stata una variabilità piuttosto bassa dei voti, insomma non ci sono stati exploit: stabilito un metro di misura, gli studenti si sono comportati in modo omogeneo rispetto alla loro media iniziale dei voti in pagella. In E invece c'è stata una (relativamente) forte variabilità: in particolare uno studente presentatosi all'esame con media 7.4 ha preso 60. In ogni caso la variabilità generale è piuttosto limitata: questo dimostra ancora una volta l'inutilità dell'esame: il candidato affronta l'esame sapendo che, mediamente, il suo voto finale non potrà essere molto diverso (al massimo 5 punti in più o in meno) rispetto all'andamento scolastico e alla conseguente media dei voti.

Conclusioni

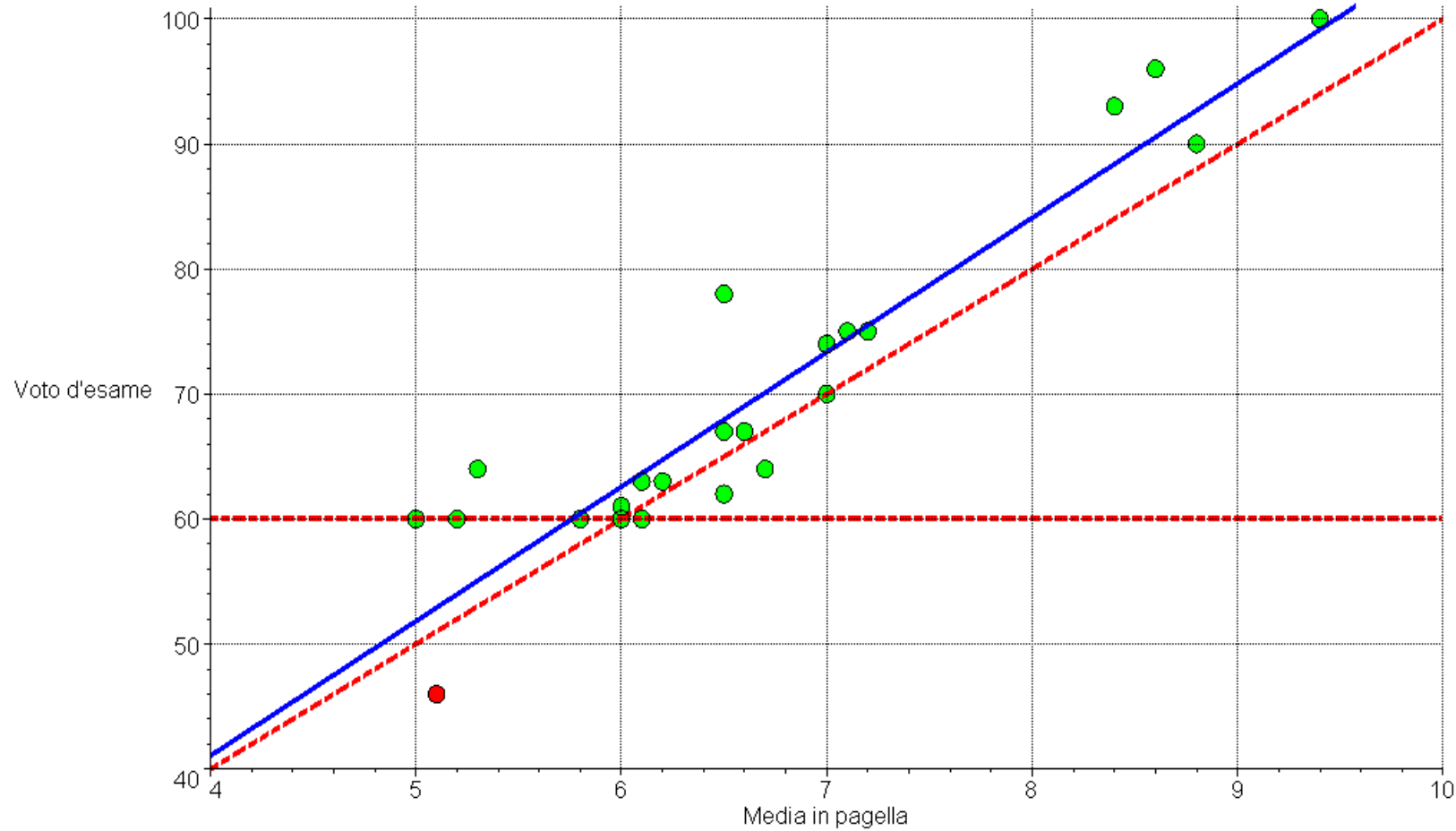
Spiace concludere che le difformità di comportamento delle diverse commissioni sono state vistose e non giustificate. Avendo fatto parte di una commissione (quella della sezione E) che mi è sembrata tutto sommato equilibrata e serena, e registrando che la sezione E non è stata certo favorita nelle valutazioni, si possono trarre le seguenti parziali conclusioni: nessuna delle commissioni ha svolto un "buon lavoro", nel senso che in generale (forse tranne che nella sezione F) non sono state seguite le indicazioni dei consigli di classe: in alcune sezioni sono stati sopravvalutati gli studenti, in altre sottovalutati.

Nelle sezioni B, D, G delle due l'una: o la commissione ha letteralmente regalato punti, oppure i consigli di classe hanno fortemente sottovalutato i propri studenti. Nelle sezioni A, C, E, F delle due l'una: o la commissione ha letteralmente rubato punti, oppure i consigli di classe hanno fortemente sopravvalutato i propri studenti.

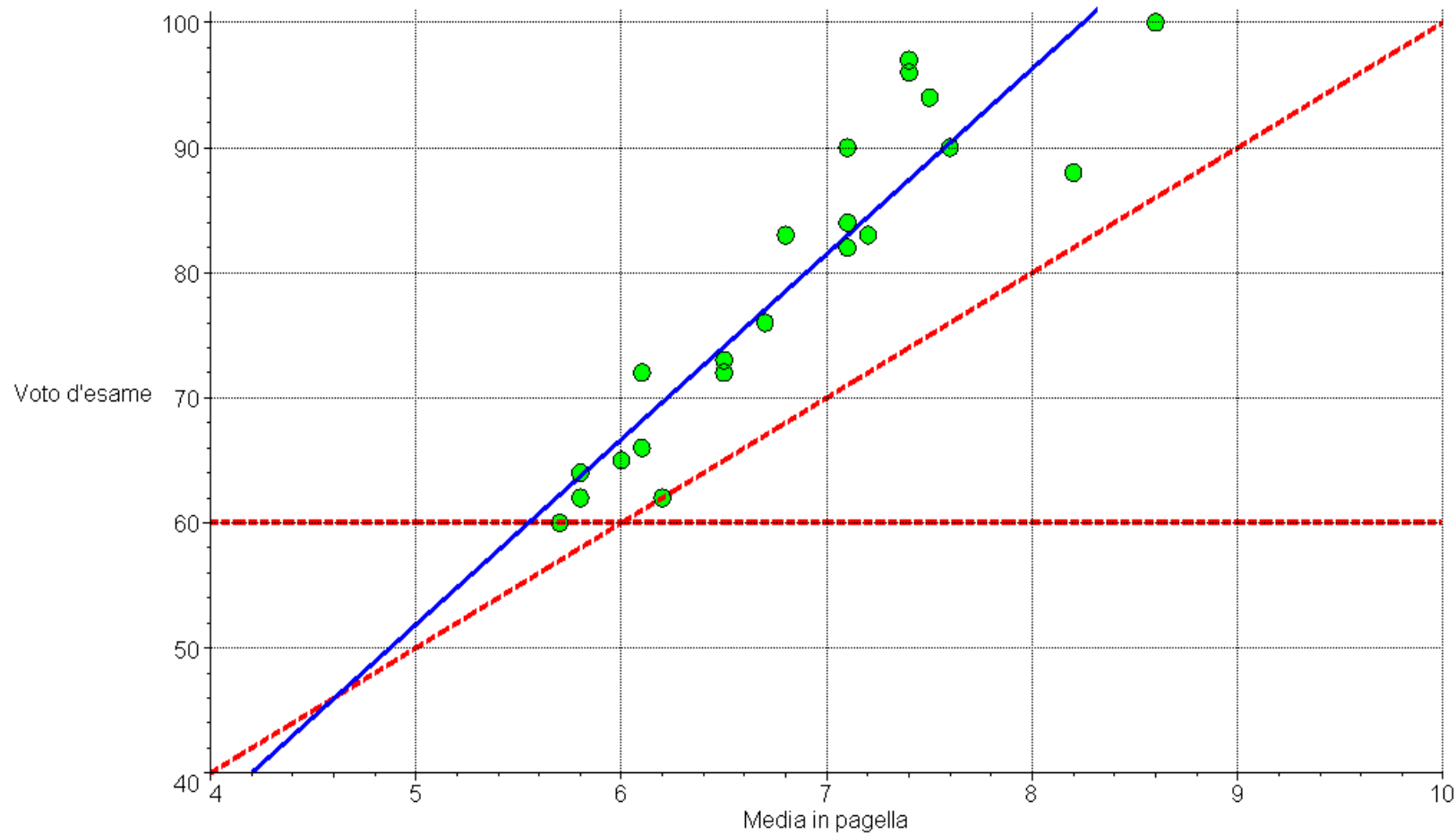
In ogni caso spiace osservare che in questo modo potrebbe consolidarsi l'idea di un esame in cui molto (troppo) dipende dalla commissione "che arriva" e poco dalla preparazione del singolo studente.

Teniamo conto che le commissioni sono formate da noi stessi (gli insegnanti in generale), e quindi i risultati esposti denunciano una nostra incapacità.

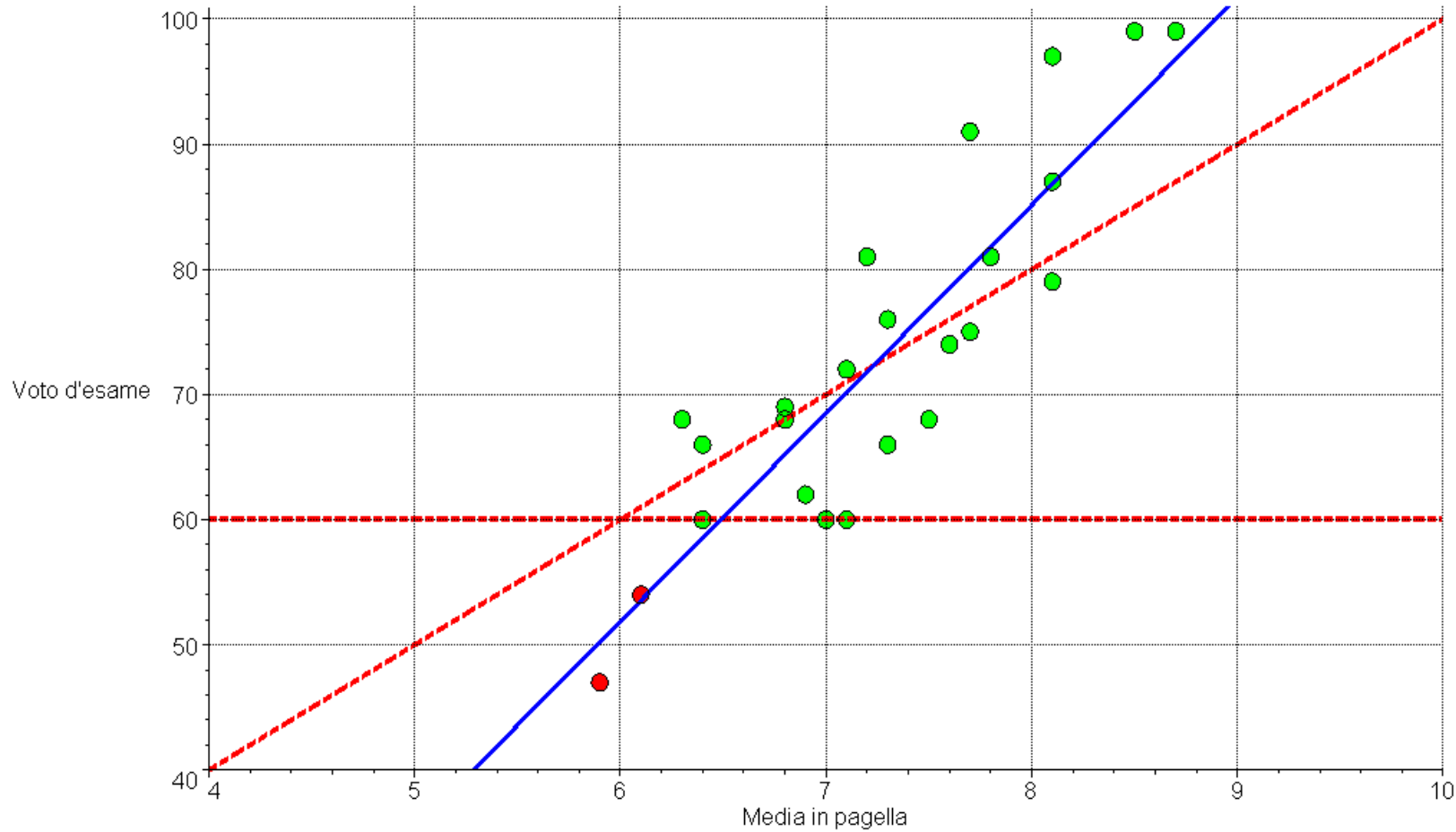
Abolire questo esame continua a sembrarmi l'ipotesi più serena, più aderente alla realtà, più economica, e soprattutto più equa per la valutazione.



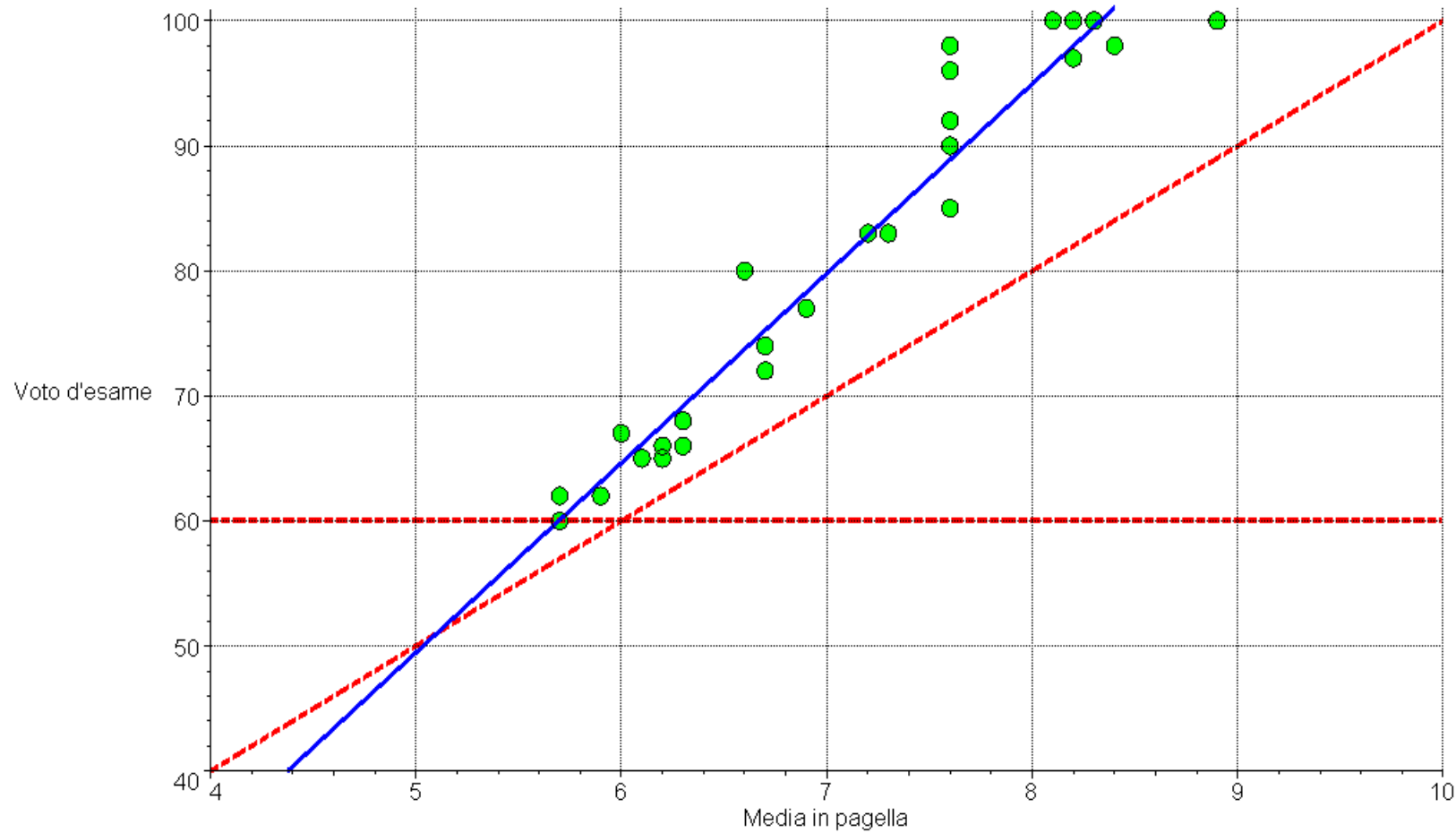
Media della classe	6.6	Scarto medio (regressione)	± 4.6
Media del voto d'esame	69.1	Scarto medio ($10x$)	$+ 3.0$
Retta di regressione	Voto = $62.6 + 10.8 (\text{Media} - 6)$		



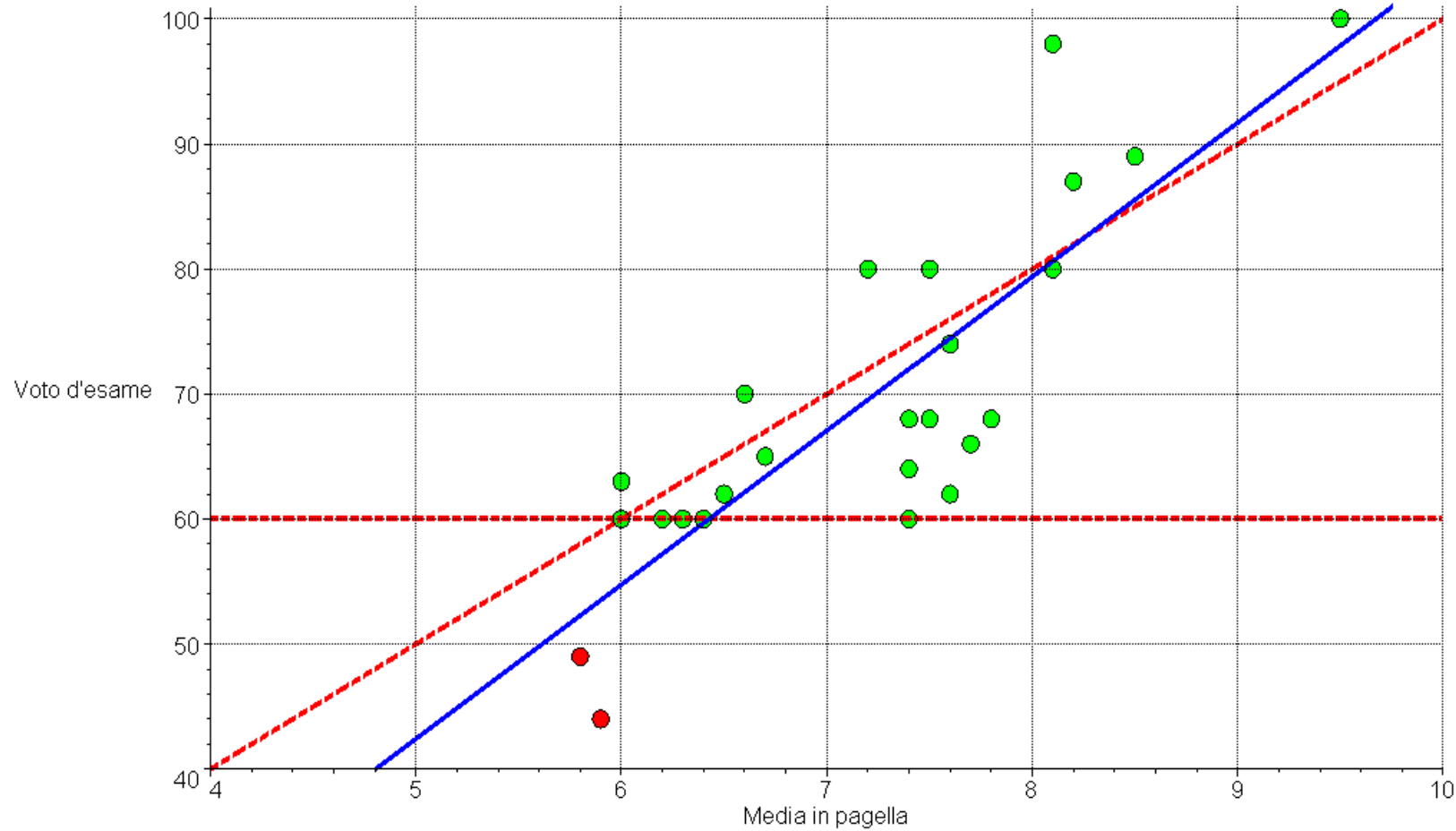
Media della classe	6.8	Scarto medio (regressione)	± 4.8
Media del voto d'esame	79.1	Scarto medio (10x)	+ 10.7
Retta di regressione	Voto = 66.7 + 14.8 (Media - 6)		



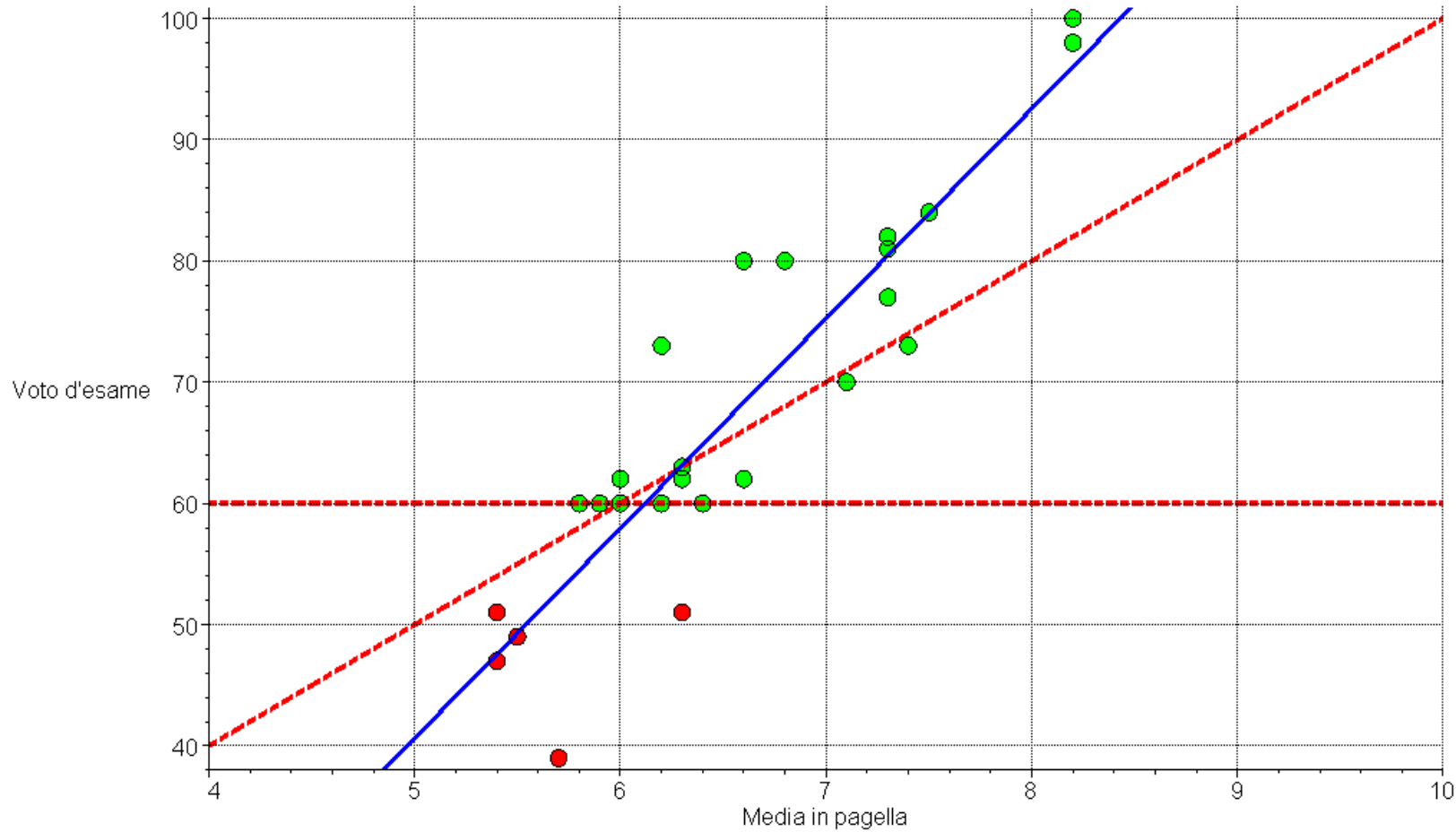
Media della classe	7.3	Scarto medio (regressione)	± 6.6
Media del voto d'esame	72.8	Scarto medio ($10x$)	$+ 0.2$
Retta di regressione	Voto = $51.9 + 16.6 (\text{Media} - 6)$		



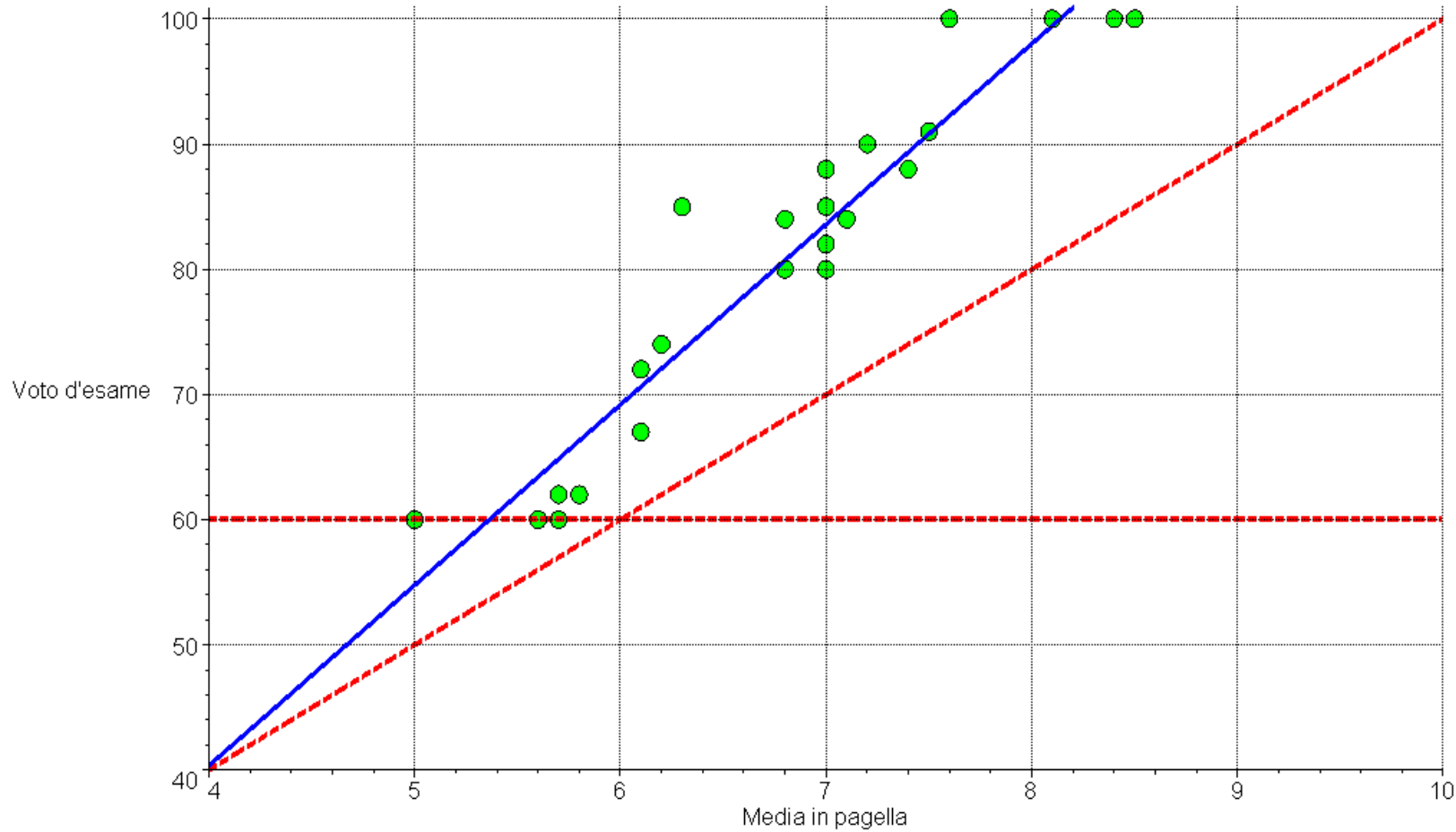
Media della classe	7.0	Scarto medio (regressione)	± 3.6
Media del voto d'esame	80.0	Scarto medio ($10x$)	+ 10.0
Retta di regressione	Voto = $64.6 + 15.2 (Media - 6)$		



Media della classe	7.2	Scarto medio (regressione)	± 7.4
Media del voto d'esame	69.5	Scarto medio (10x)	- 2.5
Retta di regressione	Voto = 54.7 + 12.3 (Media - 6)		



Media della classe	6.6	Scarto medio (regressione)	± 6.2
Media del voto d'esame	67.9	Scarto medio (10x)	+ 2.2
Retta di regressione	Voto = 57.9 + 17.3 (Media - 6)		



Media della classe	6.7	Scarto medio (regressione)	± 4.2
Media del voto d'esame	79.8	Scarto medio (10x)	+ 12.4
Retta di regressione	Voto = 69.2 + 14.4 (Media - 6)		