Significatività dell'indice h:

Il caso della Fisica

(P.R. e R.Mannella – Dipartimento di Fisica "E. Fermi" – Università di Pisa)

Introduzione (1)

- È aperta la discussione sulla possibilità di utilizzare in modo più o meno sistematico gli indicatori bibliometrici come strumento di valutazione dell'attività scientifica individuale e collettiva
- Un'ampia e recente (2007) analisi delle possibilità offerte dall'analisi bibliometrica appare nel *report* all'HEFCE predisposto dal *Centre for Science and Technology Studies* di Leiden
- Critiche severe sono state tuttavia rivolte contro l'uso di statistiche basate sul numero delle citazioni nella valutazione della produzione scientifica:

  - Kostov (1998) per primo mette in guardia sui diversi effetti sistematici che possono favorire o sfavorire la possibilità che un articolo sia citato

  - Il *Joint Committee on Quantitative Assessment of Research* nel suo *report* all'I.M.U. evidenzia l'infondatezza concettuale dell'uso dell'Impact Factor, ma anche l'inadeguatezza di <u>ogni</u> indicatore quantitativo nel giudizio di qualità

  - Editoriale di *Nature* (1.1.2009): Experts still needed

Introduzione (2)

- Ciò nonostante:
- Abbiamo bisogno di indicatori quantitativi dell'attività di ricerca
- Anche se largamente imperfetti e incapaci di cogliere molti aspetti qualitativi, gli indicatori bibliometrici sono necessari per:

  - contenere i tempi della valutazione

  - contenere i costi della valutazione

  - valutare le istituzioni scientifiche in forma aggregata

  - comparare aree differenti che non ammettono confronto qualitativo

Universalità nella distribuzione delle citazioni

- In particolare si può ovviare alle differenti caratteristiche delle comunità di ricerca grazie all'osservazione, dovuta a Radicchi *et al.* (2008) che la distribuzione delle citazioni, se rapportata al numero medio di citazioni tipico di ciascuna specifica disciplina (o anche sottodisciplina), segue un andamento universale descritto da una curva di tipo lognormale.

Universalità nella distribuzione (2)

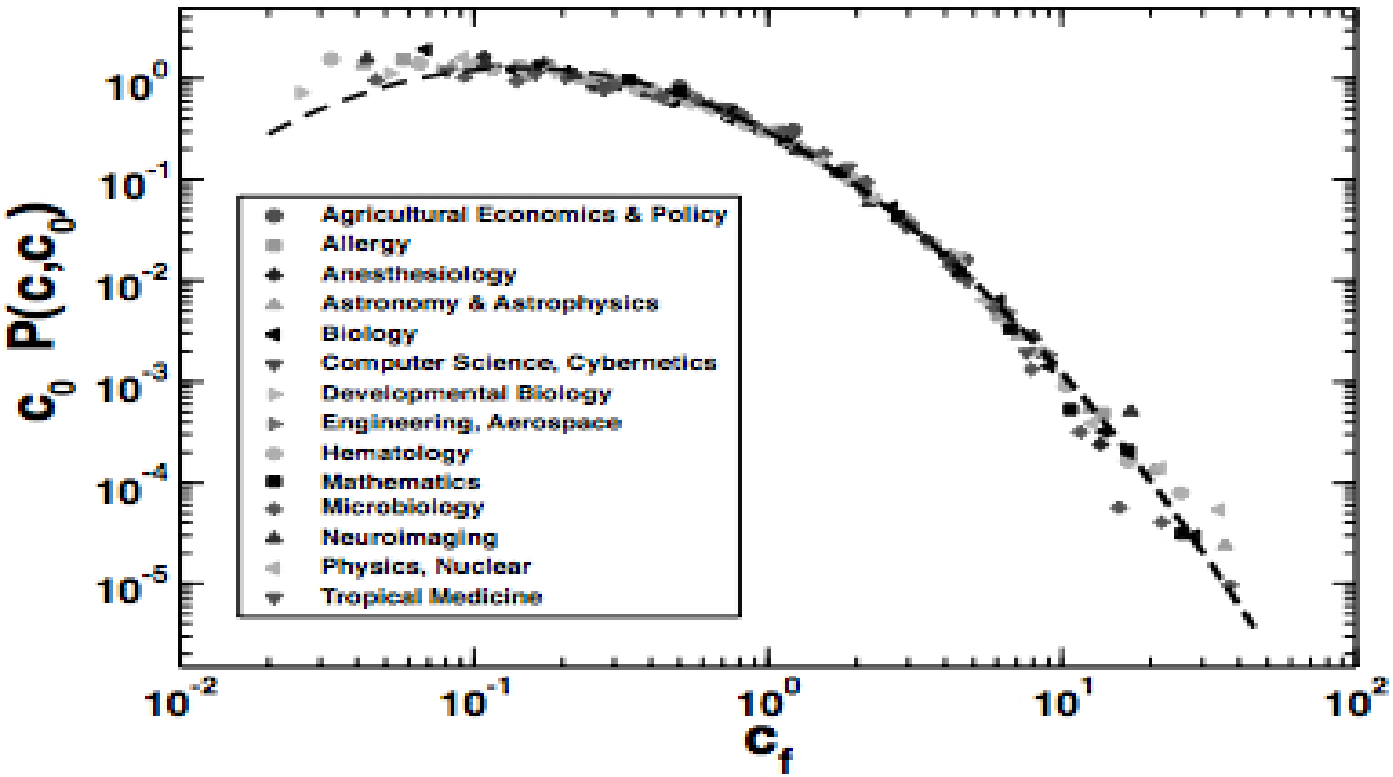


Fig. 2. Rescaled probability distribution $c_0\ P(c,c_0)$ of the relative indicator $c_f = c/c_0$, showing that the universal scaling holds for all scientific disciplines considered (see Table 1). The dashed line is a lognormal fit with $\sigma^2 = 1.3$.

L'indice h di Hirsch: presentazione (1)

- Il dibattito recente è spesso focalizzato in particolare sulla rilevanza e sul potere predittivo dell'indice h, introdotto dal fisico Hirsch nel 2005, che ha trovato rapida diffusione in diverse comunità scientifiche.
- L'indice H è definito dalla relazione per cui *uno scienziato ha un indice che vale h se ha almeno h pubblicazioni ciascuna delle quali ha ricevuto almeno h citazioni.*
- Esso combina quindi sinteticamente una misura di pura produttività con una di impatto scientifico, premiando in modo bilanciato entrambi gli aspetti dell'attività.
- Tuttavia il valore assoluto di H si dimostra dipendere in modo significativo dall'Area di ricerca, e anche dal settore particolare, a causa delle differenti dimensioni delle comunità scientifiche, delle modalità di organizzazione della ricerca, e anche dei differenti "stili" adottati nella compilazione delle bibliografie.

L'indice H di Hirsch: presentazione (2)

- Molte strategie sono state proposte per affrontare queste critiche.
- Tra le proposte di "miglioramento" vanno ricordate quelle di Egghe (2006), che definisce l'indice g per migliorare l'effetto cumulativo dei lavori molto citati, e le varianti indicate da Sidiropoulos *et al.* (2007)
- Numerose anche le critiche, tra cui quelle di Lehmann *et al.* (2006, 2008) e di Bornmann e Daniel (2007), alle quali Hirsch stesso ha tentato di offrire una risposta (2007).
- Lo studio di Jensen *et al.* (2009) ha mostrato (nel caso del CNRS), una significativa dipendenza dall'età dell'indice H diviso per il numero di anni di carriera, che sfavorirebbe i giovani, mentre la correlazione tra indice H e avanzamenti di carriera risulta non superiore al 50%.
- Malgrado ciò H resta l'indicatore con il maggior potere predittivo, seguito dal numero totale delle pubblicazioni, mentre il numero delle citazioni (sia totale che medio) ha scarsissimo potere predittivo.

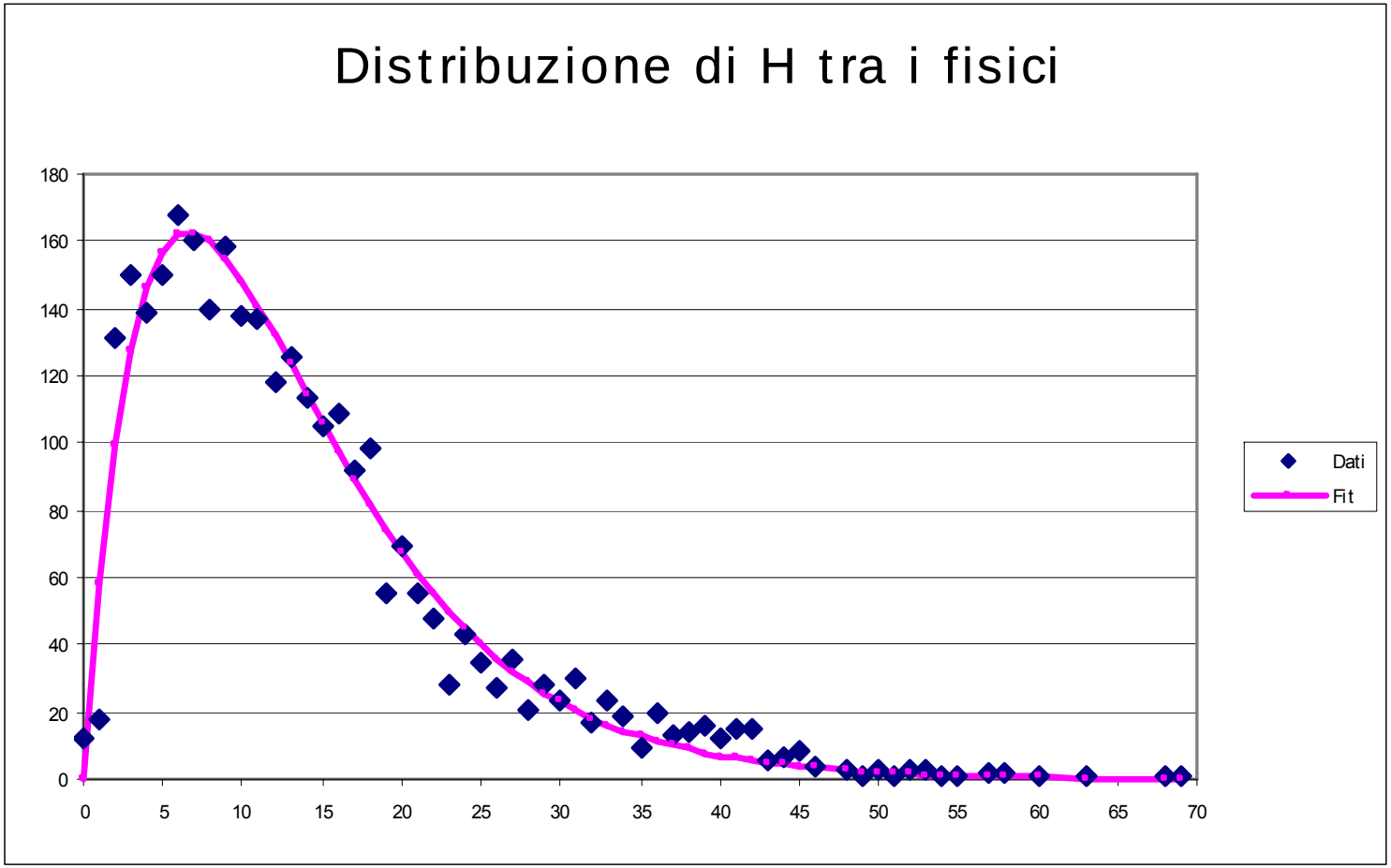
Il problema delle collaborazioni

- In alcuni settori di ricerca sarebbe anche necessario tener conto delle (talvolta grandi) differenze nel numero dei collaboratori che caratterizzano le differenti modalità di produzione scientifica, anche se la pura divisione per il numero N dei collaboratori appare impropria
- È stato proposto quindi di dividere il numero dei lavori per una potenza frazionaria di N, che andrebbe determinata empiricamente sulla base della dipendenza media del “peso” delle pubblicazioni (misurabile attraverso la dipendenza da N del numero medio delle citazioni), oppure sulla base della “facilità” della produzione collettiva (misurabile attraverso la dipendenza da N del numero medio annuo di pubblicazioni)

L'indice h dei fisici

- Per cercare di dare un'ulteriore base empirica alle considerazioni che stanno alla base dell'adozione di un qualche indicatore bibliometrico ai fini della valutazione quantitativa dell'attività di ricerca ci è parso opportuno investigare una comunità ampia ma ben delimitata: <u>i fisici universitari italiani attivi negli ultimi dieci anni</u>
- Si tratta di circa 3000 individui, per i quali sono disponibili i dati anagrafici e di carriera, e dei quali è stata analizzata la produzione scientifica sia mediante *Google Scholar* (meno preciso ma più ampio e di più facile utilizzo) che tramite la banca dati ISI Web of Science (analisi ancora preliminare)
- Per questioni di omonimia si sono dovuti eliminare dall'analisi circa un centinaio di soggetti, ma è ragionevole presumere che questa scelta non abbia introdotto alcun significativo effetto sistematico

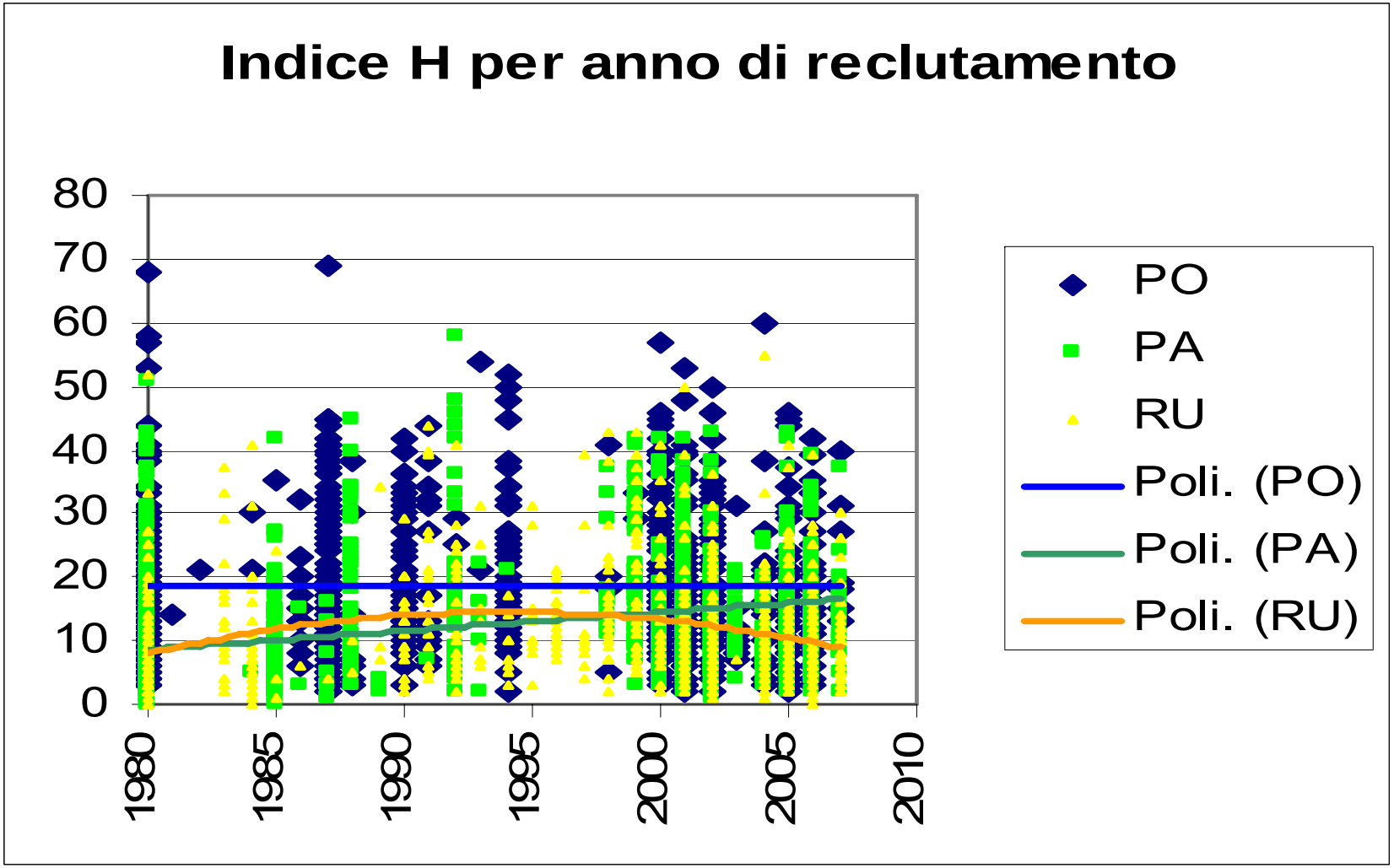
L'indice h dei fisici italiani (1)



L'indice h dei fisici italiani (2)

- I risultati sono ancora molto preliminari, ma alcune dipendenze e alcune correlazioni risultano abbastanza significative.
- La distribuzione di H è molto ben rappresentata da una distribuzione di tipo Gamma, fortemente asimmetrica rispetto al valore medio (H medio 14, H più probabile 6,5), e caratterizzata da una decrescita esponenziale della probabilità per alti valori di H (con coefficiente di decrescita 0,14).

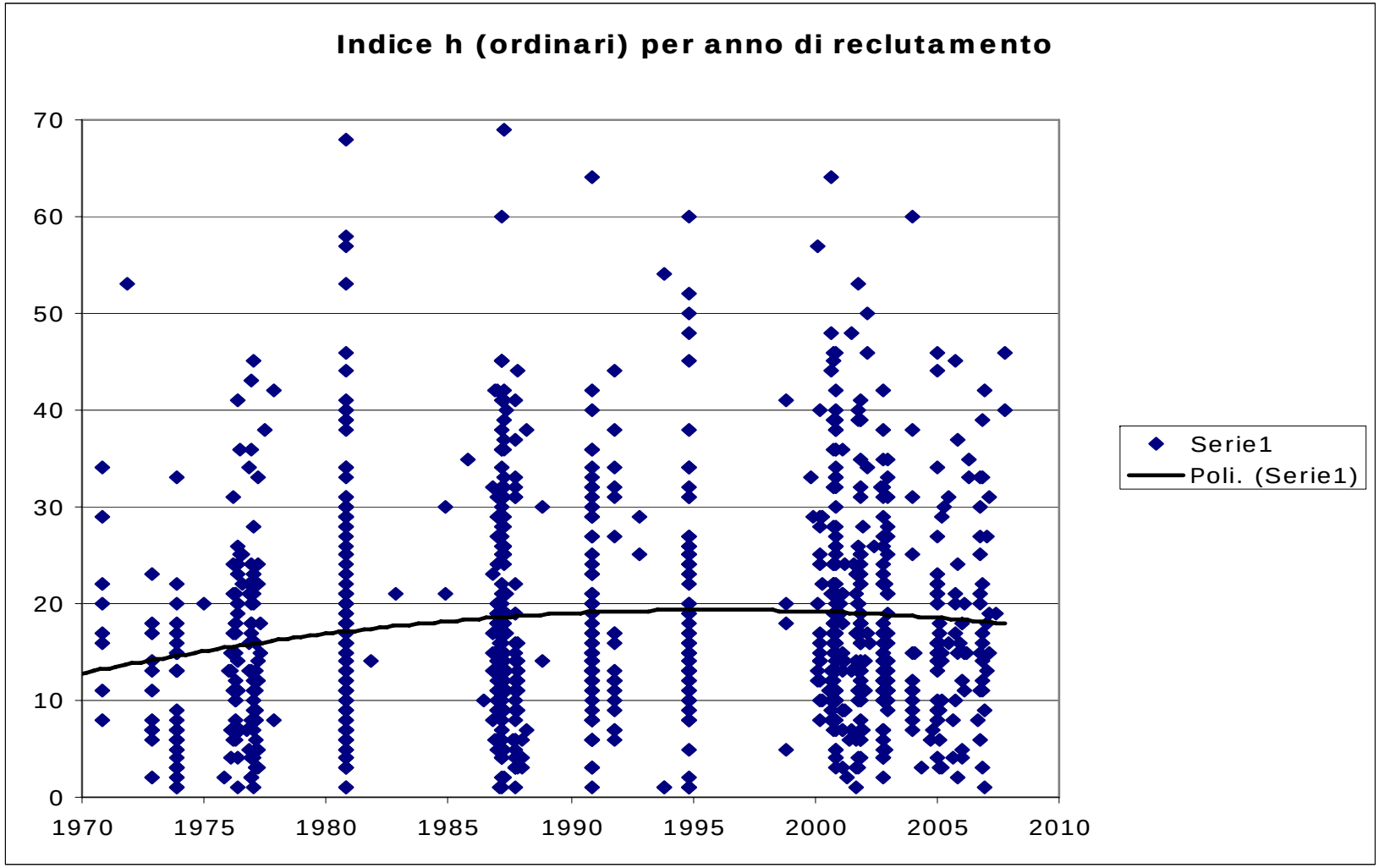
L'indice h dei fisici italiani (3)



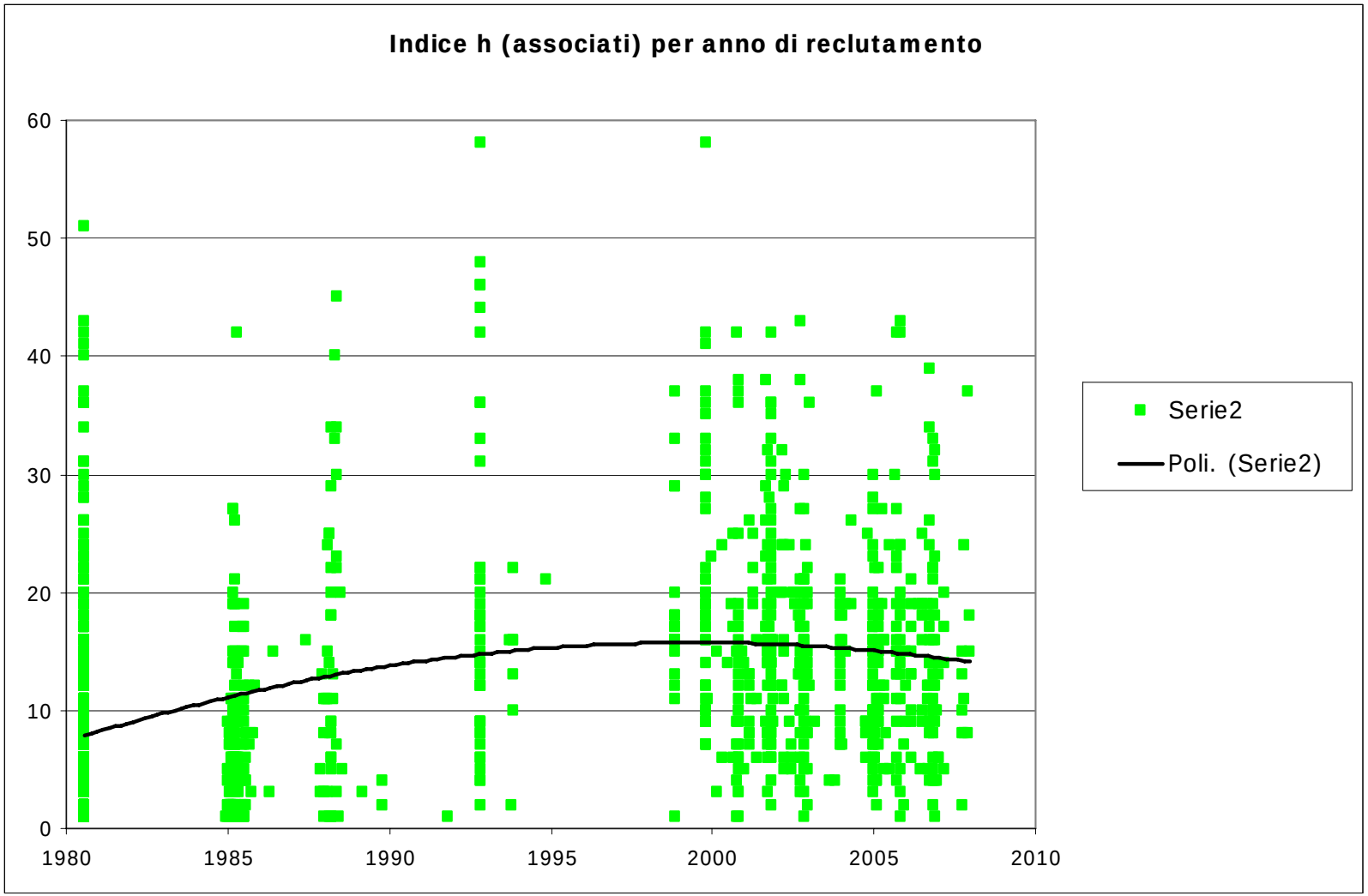
L'indice h dei fisici italiani (4)

- La dipendenza dall'anno di reclutamento mostra una stabilità assoluta per gli ordinari (H medio prossimo a 20), una tendenza alla crescita per gli associati (riconducibile soprattutto al basso livello medio dei reclutati negli anni '80 non promossi in seguito), e un'analoga tendenza per i ricercatori, tuttavia accompagnata in questo caso dalla già commentata riduzione dell'indice per i ricercatori reclutati più di recente.
- L'indice H medio al reclutamento dei ricercatori è prossimo a 10, mentre il valore medio al reclutamento degli associati è superiore a H=15, e quello degli ordinari è prossimo a H=20)

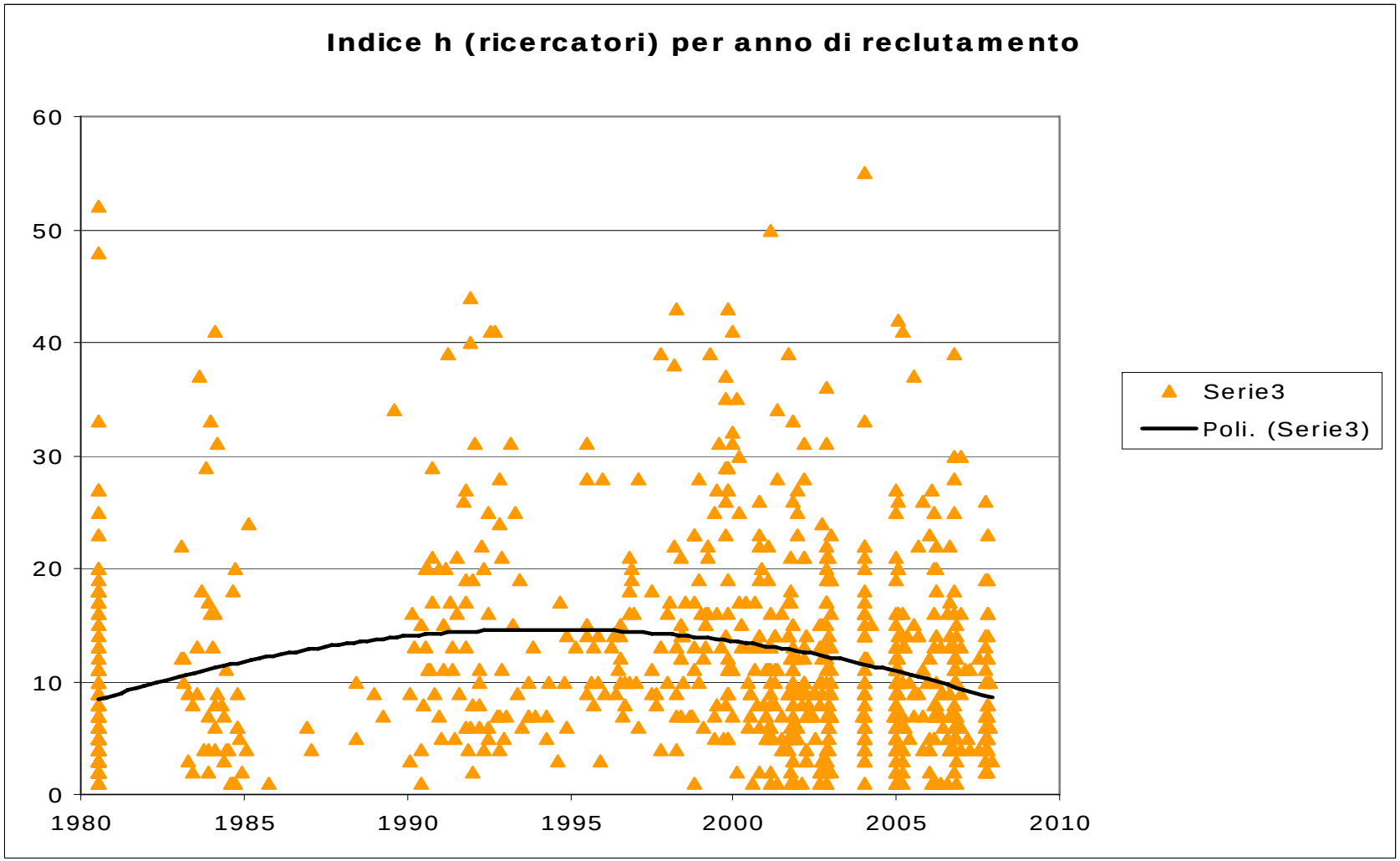
L'indice h degli ordinari



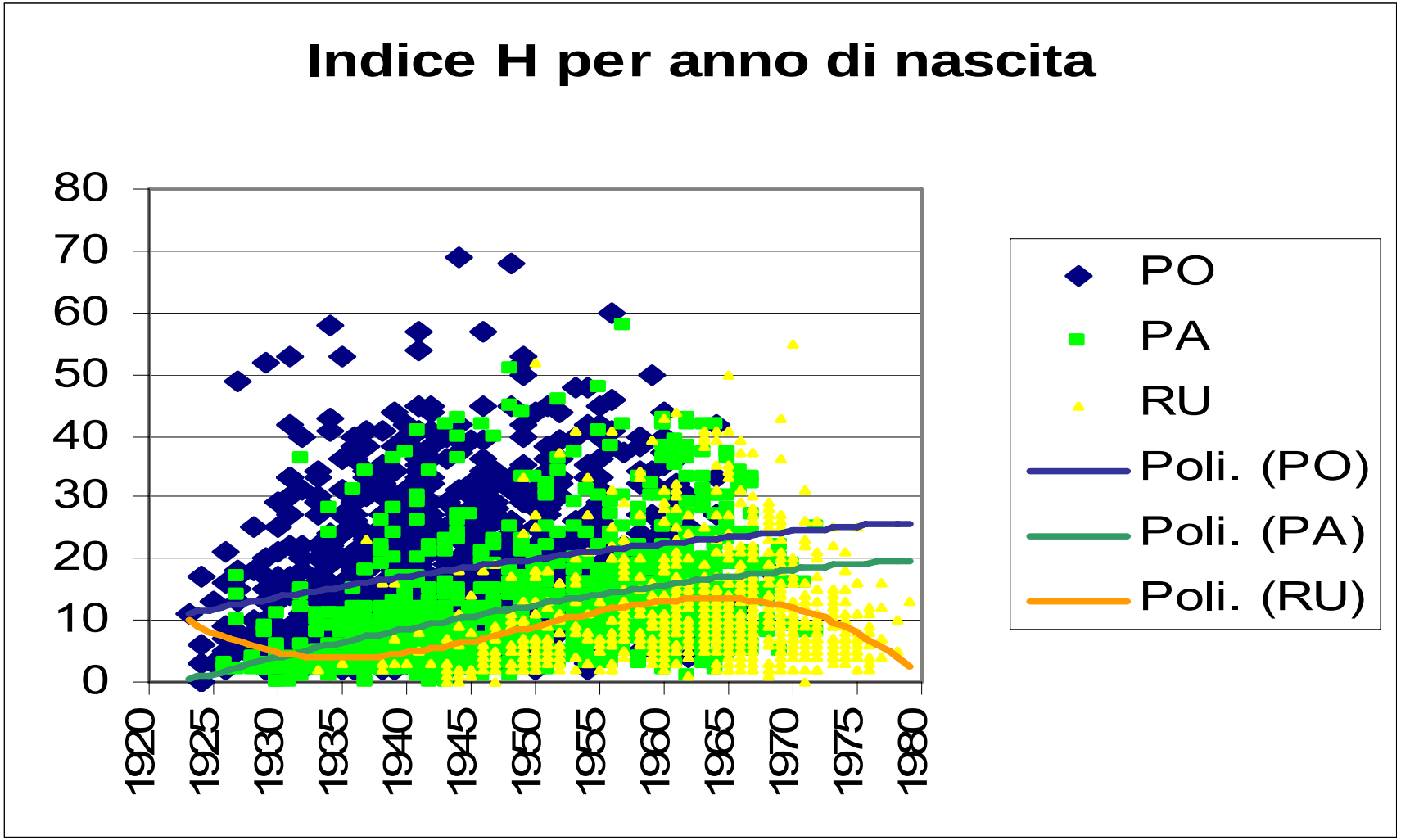
L'indice h degli associati



L'indice h dei ricercatori



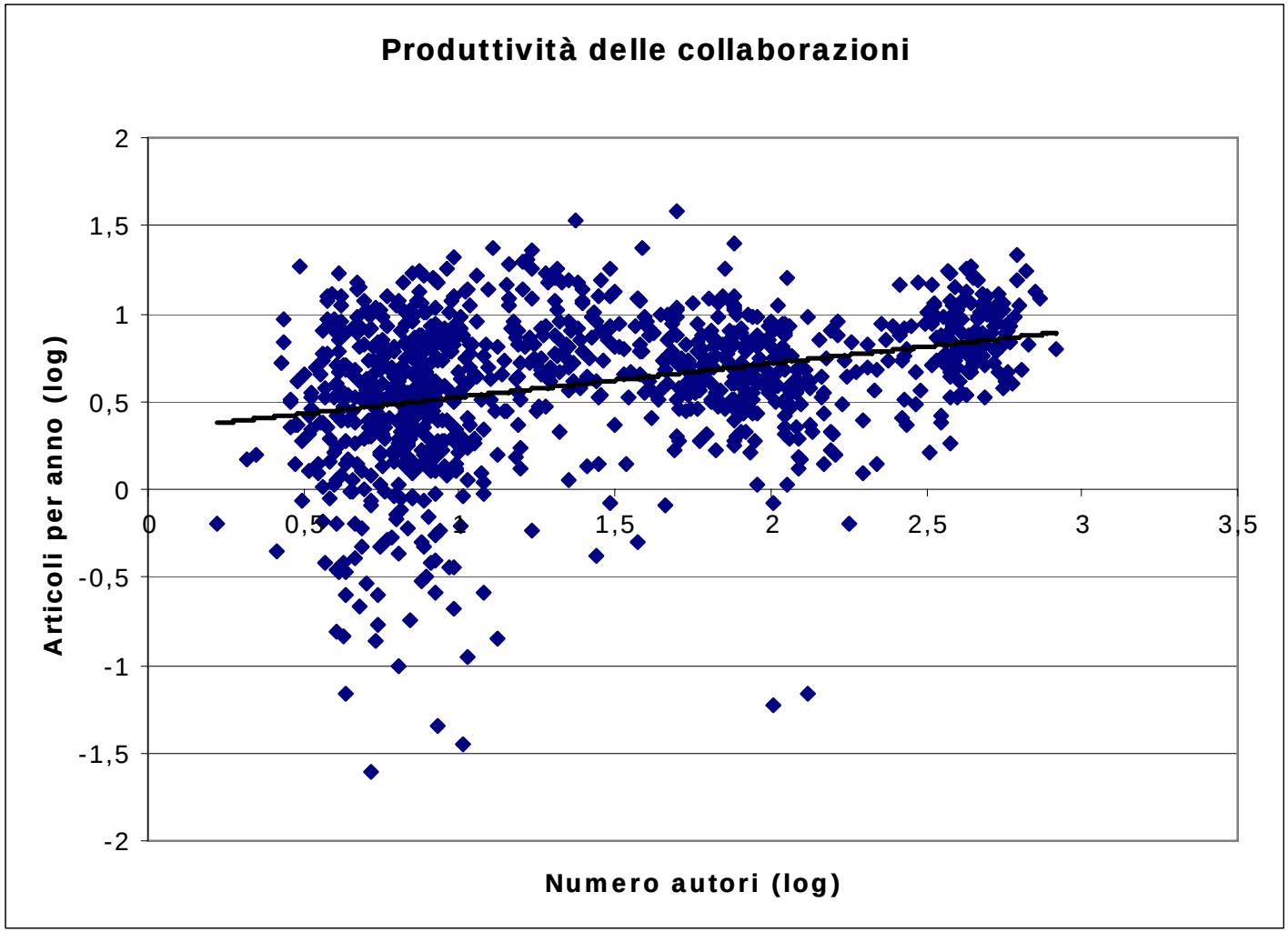
L'indice H dei fisici italiani (5)



L'indice H dei fisici italiani (6)

- La dipendenza dall'anno di nascita mostra una crescita costante per ordinari e associati (legata alla scarsa indicizzazione della produzione più remota e alla crescita nel tempo del numero medio di citazioni), mentre l'indice H dei ricercatori più giovani è decrescente, come d'altronde prevedibile a causa del limitato periodo di attività.

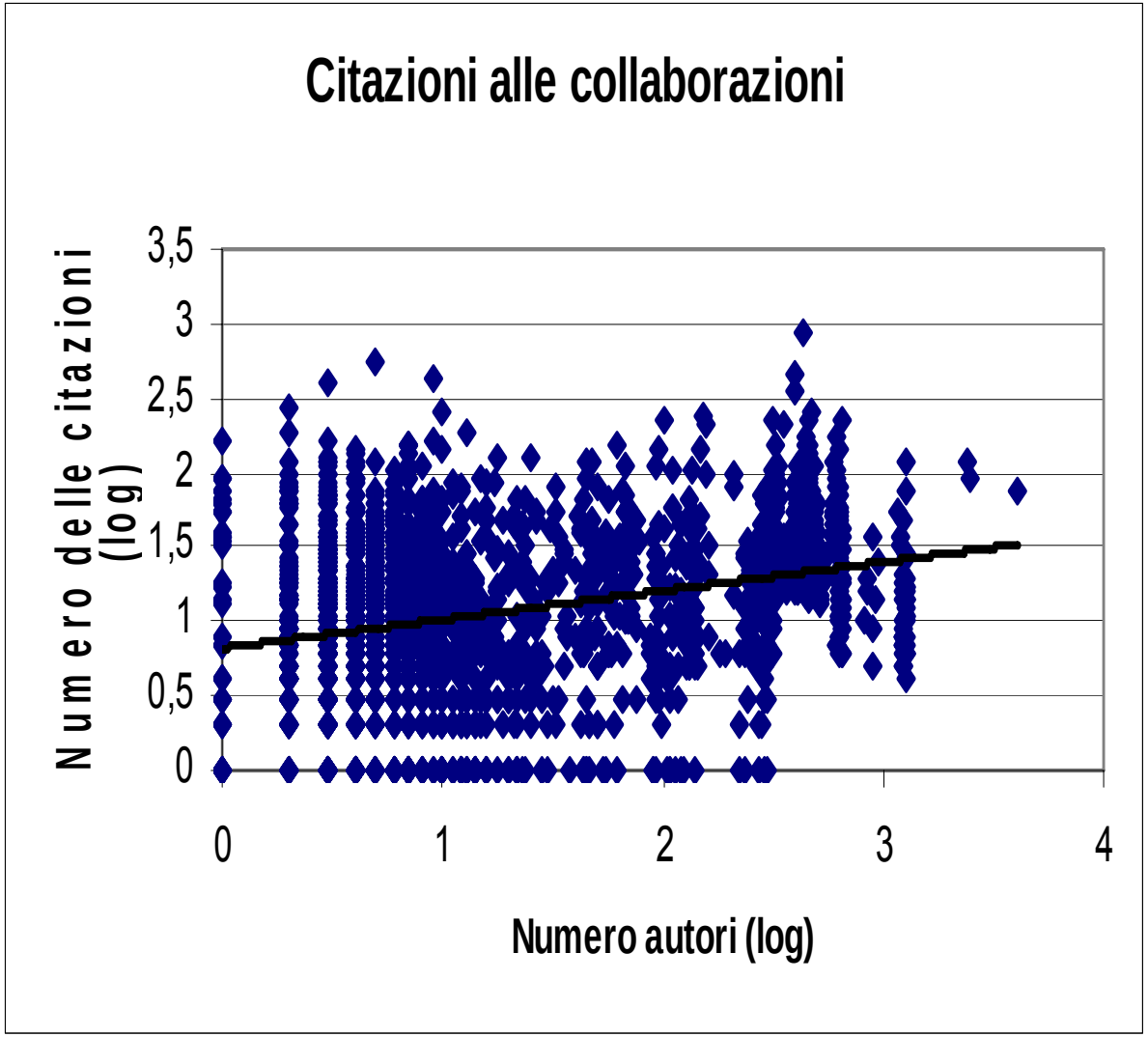
La dipendenza dal numero degli autori (1)



La dipendenza dal numero degli autori (2)

- Ogni indicatore bibliometrico dipende significativamente anche dalla peculiari modalità di produzione degli articoli scientifici che esistono in comunità di ricerca differenti
- In particolare tra i fisici esiste una forte variabilità nel numero dei collaboratori a una particolare attività di ricerca.
- Il grafico (in doppia scala logaritmica nella previsione di una legge di potenza) mostra l'esistenza di tre *cluster*, significativamente separati l'uno dall'altro e incentrati rispettivamente intorno ai 5-6, ai 60-70 e ai 400 collaboratori (medi per autore)
- L'esponente della legge di scala che lega il numero medio di articoli prodotti per anno al numero medio di collaboratori è 0,2
- Si tratta di un valore "basso" (che indica uno scarso beneficio di tipo "quantitativo" derivante dalla produzione in collaborazione) ma non trascurabile

La dipendenza dal numero degli autori (3)



La dipendenza dal numero degli autori (4)

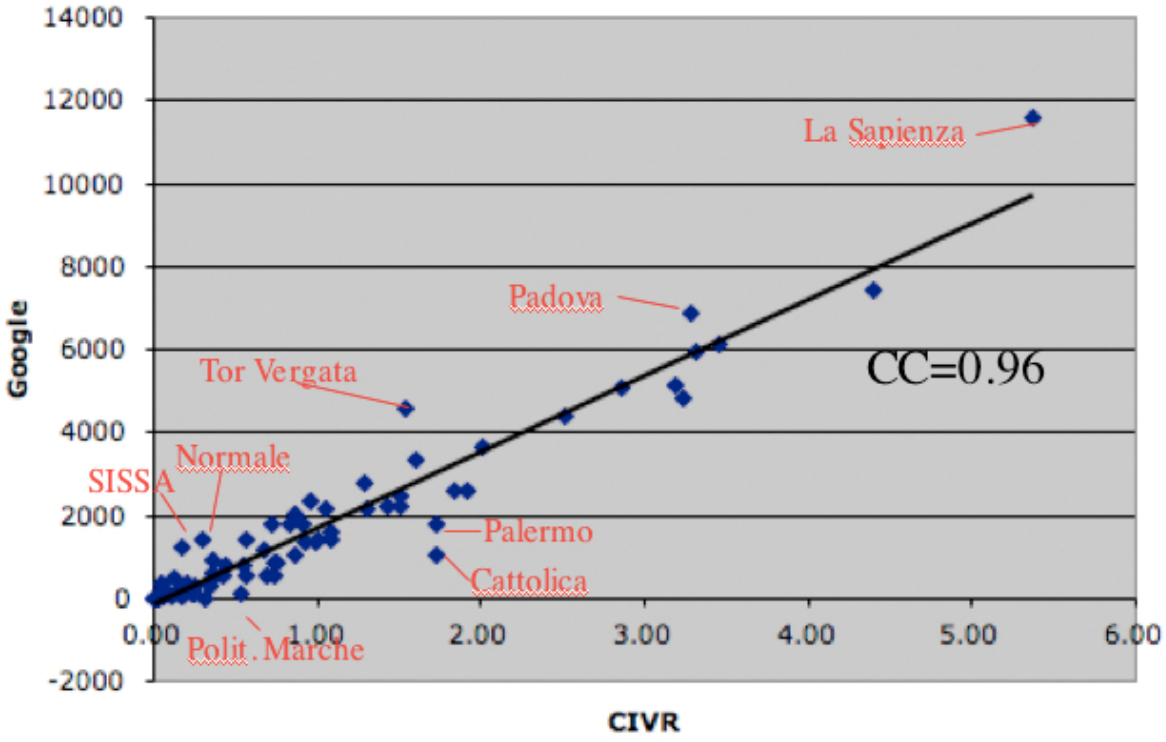
- Anche il numero delle citazioni per articolo può dipendere in modo significativo, oltre che dall'ambito, anche dalla modalità organizzative della ricerca e in particolare dal numero medio di collaboratori
- Anche in questo caso si riconoscono i *cluster* già menzionati, e anche in questo caso si nota una dipendenza, debole ma significativa, del numero delle citazioni dal numero dei collaboratori
- L'esponente della legge di scala (forse non casualmente) è di nuovo prossimo a 0,2
- In questo caso tuttavia è il complemento a 1 (ossia 0,8) che indica la riduzione dell'importanza del contributo individuale a un articolo dovuta alla collaborazione
- Poiché i due fattori sono entrambi rilevanti per la valutazione individuale, si potrebbe (nello spirito dell'indice h) considerarne la media geometrica e utilizzare il coefficiente di riduzione 0,5 (ossia in pratica dividere il numero dei lavori (o delle citazioni) per $\sqrt{n}$

La valutazione delle istituzioni scientifiche (1)

- Tutte le obiezioni che fondatamente vengono sollevate nei confronti degli indicatori bibliometrici acquistano un peso molto più ridotto quando si affronta il tema della valutazione aggregata e comparativa delle istituzioni di ricerca.
- Un'evidenza particolarmente significativa di tale affermazione è data dall'analisi effettuata da Cesareni (2007), che ha misurato l'indice H (collettivo) delle istituzioni di ricerca valutate dal CIVR per il periodo 2001-2003 e ha confrontato il risultato di tale misura con i punteggi attribuiti dal CIVR. La correlazione tra le due valutazioni è risultata essere 0,96.
- Ovviamente la rapidità e l'economicità delle valutazioni basate su indicatori bibliometrici sono incomparabilmente maggiori, e quindi la valutazione mediante indici è da raccomandarsi per un monitoraggio frequente e sistematico che non influisca sui giudizi individuali.

La valutazione delle istituzioni scientifiche (2)

CORRELAZIONE TRA LA VALUTAZIONE DEL CIVR E QUELLA OTTENUTA CON GOOGLE SCHOLAR



Questa presentazione è coperta, per le sue parti originali, da licenza Creative Commons:

attribuzione, non commerciale, condividi allo stesso modo

http://creativecommons.org/licenses/by-nc-sa/2.5/it/legalcode

rossi@df.unipi.it