



VALUTAZIONE

Misuriamo il merito con l'h

LA VALUTAZIONE DELLA RICERCA: NECESSITÀ E OPPORTUNITÀ

La necessità, nel quadro di una politica di contenimento della spesa pubblica, di razionalizzare la spesa per l'Università italiana può essere un'opportunità per migliorarne la qualità ed il livello. Così, 474 milioni di euro, corrispondenti al 66% dei 720 milioni di euro, finalizzati a «promuovere e sostenere l'incremento qualitativo delle attività delle università statali e di migliorare l'efficacia e l'efficienza nell'utilizzo delle risorse», (legge 9 gennaio 2009 numero 1), sono stati ripartiti tra le Università sulla base di una valutazione della ricerca.

Come riportato nell'allegato1 del DM 21/12/2010 numero 655, i 474 M di euro di cui sopra, sono stati distribuiti, per il 70%, sulla base della partecipazione a finanziamenti pubblici nazionali (PRIN e FIRB), europei e internazionali e, per il restante 30%, sulla base del coefficiente di ripartizione destinato alle Aree dalla valutazione triennale della ricerca (VTR) 2001-2003, che peraltro è indipendente dalla qualità della ricerca di ciascuna Università.

E' chiaro che misurare la qualità e la produttività della ricerca scientifica sulla base dei finanziamenti pubblici ottenuti sarebbe come valutare l'efficienza dello smaltimento dei rifiuti sulla base dei fondi a quel fine impegnati. Napoli avrebbe in questo modo la palma d'oro! Lo stesso discorso potrebbe farsi per la Sanità e così via. Perciò, la distribuzione alle Università dei fondi per la premialità è un perfetto esempio del fatto che una corretta politica del merito è strettamente dipendente dall'applicazione di appropriati criteri di valutazione: se questi sono fallaci, il principio meritocratico può tramutarsi nel suo opposto. Nel caso della valutazione della ricerca, questa deve partire dall'analisi dei suoi prodotti, cioè lavori scientifici e, per quella applicata, anche brevetti. L'uso dei finanziamenti come parametro di valutazione ha un senso solo in relazione ai prodotti, al fine, eventualmente, di calcolare il rapporto costo/beneficio. Purtroppo, seppure antico, il problema della valutazione della ricerca scientifica non è stato risolto in maniera semplice, univoca e universalmente accettata; non solo, ma i metodi utilizzati sono diversi a seconda della scala nella quale la ricerca viene valutata. Per esempio, per valutare una serie di progetti di ricerca nell'ambito di un programma di finanziamento di specifici temi di ricerca, come è il caso dei programmi di ricerca europei o dei progetti strategici nazionali, il sistema corrente è quello della cosiddetta *study session* cioè del confronto diretto tra i membri di una commissione di esperti, possibilmente internazionale. E' evidente che questo metodo non è applicabile ad una scala più ampia, per esempio alla valutazione della produzione scientifica di un gran numero di lavori scientifici all'interno dello stesso settore disciplinare. E' questo il caso del CIVR, istituito nel 2004 per valutare la ricerca nazionale nel triennio 2001-2003 ed è anche quello utilizzato per selezionare i progetti di

ricerca del PRIN. In questo caso la valutazione di ciascun prodotto della ricerca è affidata ad almeno due valutatori anonimi, secondo uno schema ed una metodologia decisa da un panel di esperti per ciascuna delle 14 aree disciplinari o da un comitato di garanti. Gli esperti sono contattati e trasmettono i loro pareri al panel e interagiscono tra loro in maniera anonima e via internet. I problemi di tale metodo sono l'estrema parcellizzazione della valutazione che rende difficile se non impossibile l'applicazione di un uniforme metro di valutazione e consente di fatto, sotto la copertura dell'anonimato, l'esercizio della più estrema soggettività di giudizio. Su questo aspetto rinvio all'editoriale di Ernesto Carafoli su SciRE ([Valutazione della ricerca? No grazie!](#)).

I problemi connessi all'utilizzazione di questa procedura ha portato all'introduzione, nella valutazione della ricerca su scala nazionale ed internazionale, all'introduzione di metodi basati su parametri bibliometrici che esprimono l'impatto dei lavori sulla comunità scientifica, quantificato in termini di citazioni. Per esempio, nel DM numero 8 2010 del MIUR, con cui si tracciano le linee guida della valutazione quinquennale della ricerca (VQR) per il periodo 2003-2008, si introduce, accanto alla valutazione da parte di revisori anonimi, l'analisi delle citazioni.

Il numero di citazioni di un lavoro da parte della comunità scientifica non è una misura assoluta della qualità scientifica della pubblicazione ma sicuramente è un'espressione del suo impatto sull'elaborazione delle ipotesi e delle idee che formano la base dell'attività di ricerca.

Ma fino a che punto questo parametro è utilizzabile per valutare la ricerca di un ricercatore e di un'istituzione di ricerca e per paragonare istituzioni diverse e operanti in aree di ricerca diverse? Per esempio, che senso ha paragonare il numero bruto di citazioni di lavori appartenenti ad aree così diverse come quelle delle scienze chimiche, le cui pubblicazioni spesso riguardano argomenti di ricerca estremamente specialistici coltivati da un numero ristretto di ricercatori, e delle scienze mediche, che si rivolgono a comunità scientifiche ben più numerose, oppure, paragonare le citazioni di lavori di biologia cellulare o di immunologia, che spesso circolano su periodici internazionali generalisti come *Nature* e *Science*, e quindi accessibili ad una larghissima audience, con quelle di lavori di matematica, che circolano all'interno di una ristretta cerchia di autori, magari su periodici nazionali e per questo nemmeno computati da certe banche dati?

E' evidente che l'uso delle citazioni e degli indici da queste derivati, senza un'adeguata normalizzazione per la media delle citazioni dello specifico settore di ricerca, potrebbe produrre, almeno teoricamente, risultati fuorvianti. Ma, entrando nel merito, quali sono gli indici basati sulle citazioni correntemente utilizzati?

INDICI BIBLIOMETRICI E BANCHE DATI

Il primo e più antico e tuttora molto diffuso indice bibliometrico è l'impact factor (IF), sviluppato originariamente ad uso dei bibliotecari, come orientamento nella scelta dei periodici più diffusi ai quali abbonarsi. L'IF di un periodico scientifico è la media delle citazioni ottenuta dai lavori pubblicati in quel periodico nei due anni precedenti. Applicato ad un singolo lavoro scientifico, l'IF potrebbe teoricamente fornire una previsione sulle citazioni future di quel lavoro, a patto però che le citazioni che costituiscono l'IF siano distribuite su un numero sufficientemente ampio di lavori. Questo però non è il caso, dato che l'IF è determinato da una piccola percentuale di lavori, corrispondenti a circa il 5% del totale, che ottengono un elevatissimo numero di citazioni rispetto alla media. Per questo

motivo non esiste correlazione tra IF e citazioni effettive di ciascun lavoro e, come ciascuno di noi ha sperimentato per i propri articoli, spesso succede che lavori pubblicati su riviste ad alto IF ricevano meno citazioni di quelli pubblicati su riviste a basso IF. Inoltre, l'IF appiattisce le differenze tra ricercatori di buon livello appartenenti alla stessa settore disciplinare e che, come tali, pubblicano sugli stessi periodici. D'altra parte, l'IF ha valore *hic et nunc*, cioè per quel periodico e per l'anno nel quale è stato pubblicato il lavoro e quindi ha poco senso utilizzare l'IF dell'anno in corso per un lavoro pubblicato anni addietro, come è invece pratica comune nei concorsi universitari. Ma ciò che taglia la testa al toro e decreta l'inevitabile obsolescenza dell'IF come indice di valutazione della ricerca è di natura squisitamente pratica: la disponibilità di banche dati che forniscono gratuitamente le citazioni dei lavori e di software capaci di calcolare istantaneamente, utilizzando quelle citazioni, una serie di parametri utili per valutare la ricerca .

Attualmente sono disponibili tre diverse banche dati, due a pagamento, ISI Web of Science (Wos) e Scopus, e una gratuita, Google Scholar, i cui dati possono essere analizzati con due diversi software, [Publish or Perish](#) e un [add-on di Mozilla Firefox](#).

Le tre banche dati non sono equivalenti; per esempio, Scopus computa le citazioni solo a partire dal 1996 mentre ISIWos parte dal 1900 e Google scholar fin dall'800 (Darwin, The origin of Species, Appleton, 1859: 8.989 citazioni). Inoltre, la copertura dei diversi settori non è la stessa: per esempio, ISIWos, al contrario di Scopus e Google Scholar, utilizza come fonte un numero relativamente limitato di riviste di prestigio di lingua inglese sulla base dell'enunciato di Bradford (1934), secondo cui la ricerca di qualità è concentrata in un ristretto numero di riviste anglo-sassoni (sic!). Mentre ISIWos e Scopus utilizzano solo citazioni su riviste i cui lavori sono accettati solo dopo revisione da parte di esperti (*peer reviewed*), Google Scholar utilizza anche citazioni su libri e articoli di riviste *non-peer reviewed*. In pratica, le tre banche dati forniscono risultati, in termini di citazioni totali, qualitativamente e quantitativamente diversi. Tuttavia nella ricerca per autore, selezionando accuratamente i lavori ed eliminando le omonimie, ISIWos e Google Scholar danno risultati abbastanza in accordo tra loro se si utilizza come indice quello di Hirsch. Il vantaggio di questo indice è quindi anche quello di ridurre le differenze nelle citazioni totali tra le varie banche dati.

Esistono anche banche dati delle citazioni per specifici settori, come l'Astrophysics Data System (ADS) e il Physical Review Online Archive (PROLA). E' da notare che per quanto riguarda l'astrofisica, ADS utilizza un database più ampio di quello delle tre banche dati generaliste e questo spiega il fatto che i valori di citazioni totali e di indice di Hirsh ottenuti è sensibilmente (30-50%) superiore ai valori forniti dalle altre banche dati.

L'INDICE DI HIRSCH

L'indice di Hirsch (*h*) di un autore è il numero di lavori di quell'autore che hanno ottenuto un numero di citazioni non inferiore a quel numero. L'*h* è un vero e proprio uovo di Colombo dato che esprime in maniera mirabilmente concisa la consistenza e affidabilità dell'impatto della produzione scientifica di un autore. Un elevato *h* indica che l'autore ha una produttività che si è mantenuta ad un livello elevato per un periodo di tempo prolungato. Perciò l'*h* ha il vantaggio di attenuare le differenze tra diversi settori disciplinari caratterizzati da diverse emivite dei relativi lavori fornendo una parziale normalizzazione tra settori diversi. Ciò non esclude che l'*h*, quando venga applicato alla valutazione degli istituti di ricerca pubblici e privati, possa essere normalizzato alla media degli *h* di quel settore. Un esempio di come l'*h* possa essere utilizzato per valutare la ricerca delle

istituzioni pubbliche e private nazionali è stato fornito da Via Academy (VA), un'iniziativa del tutto autonoma di Mauro degli Esposti, professore associato presso l'Università di Manchester, e dei suoi collaboratori. VA ha prima creato una base di 1524 **Top Italian Scientists** (TIS), che include gli italiani che lavorano all'estero, e comprende ricercatori con h uguale o superiore a 31.

Raggruppando i TIS per istituzione di appartenenza è stata così ottenuta una classifica delle Top 50 Italian Institutions, che può essere eventualmente normalizzata sulla base dei ricercatori afferenti a ciascuna istituzione. L' h è naturalmente correlato all'età, ma questa caratteristica può costituire un plus se utilizzato per valutare la ricerca delle istituzioni, esprimendo la capacità dei suoi ricercatori di fare scuola.

CONCLUSIONE

Non c'è dubbio che suona fortemente riduttiva la pretesa di esprimere con un'unico numero il valore della ricerca di una vita o la produttività di un'intera istituzione. Ma, al fine di valutare la ricerca, l'indice di Hirsch è comunque preferibile a parametri, come la partecipazione a progetti nazionali o europei, correntemente utilizzati dal Ministero dell'Università. Siamo convinti che le classifiche delle istituzioni basate su un parametro obiettivo, trasparente e certamente conciso come l' h possa contribuire, eventualmente normalizzato, all'attuazione di una politica del merito nel finanziamento della ricerca e dell'Università in Italia.



Gaetano Di Chiara

Neuroscienze, Università di Cagliari