

In questo numero:

3•Alessandra Petrucci

Editoriale

5•Alessandra Petrucci

Il riconoscimento della scientificità delle riviste: l'esperienza del CUN attraverso la Consultazione Pubblica

10•Aldo Pavan

Valutazione della Ricerca Scientifica e Comportamenti Accademici

13•Donato Posa, Sandra De Iaco, Sabrina Maggio, Monica Palma

Processi valutativi, qualità dei prodotti e interessi di ricerca

19•Maurizio Carpita

Valutazione della Ricerca e Classifiche delle Riviste Scientifiche nelle Scienze Statistiche: l'esperienza SIS

27•Francesco Bartolucci

Sulla misurazione del prestigio di una rivista scientifica

32•Alberto Baccini

La VQR di area13: una riflessione di sintesi

38•Daniela Cocchi

Gli statistici nell'area 13 italiana e nei settori ERC

Direttore Responsabile

Nicola Torelli

Direttore Scientifico

Giorgio Vittadini

Comitato Editoriale

Tommaso Agasisti (*tommaso.agasisti@polimi.it*)
Giancarlo Blangiardo (*giancarlo.blangiardo@unimib.it*)
Alberto Brugnoli (*Alberto.Brugnoli@eupolislombardia.it*)
Andrea Cammelli (*andrea.cammelli@unibo.it*)
Maurizio Carpita (*carpita@eco.unibs.it*)
Giuliana Coccia (*coccia@istat.it*)
Gianni Corrao (*giovanni.corrao@unimib.it*)
Alberto Daprà (*alberto.dapra@lispa.it*)
Tommaso Di Fonzo (*tommaso.difonzo@unipd.it*)
Alessandro Fassò (*alessandro.fasso@unibg.it*)
Paolo Giudici (*paolo.giudici@unipv.it*)
Gianpaolo Gualaccini (*ggualaccini@cons.cnel.it*)
Stefano Iacus (*stefano.iacus@unimi.it*)
Achille Lemmi (*achille.lemmi@unisi.it*)
Mario Mezzanzanica (*mario.mezzanzanica@unimib.it*)
Maria Gabriella Ottaviani (*mariagabriella.ottaviani@uniroma1.it*)
Monica Pratesi (*m.pratesi@ec.unipi.it*)
Antonio Quaglio (*antonio.quaglio@ilsole24ore.com*)
Carlo Zocchetti (*carlo_zocchetti@regione.lombardia.it*)

Società Italiana di Statistica

Salita de' Crescenzi, 26 – 00186 Roma
www.sis-statistica.it - sis@sis.statistica.it, tel. 06.6869845

Registrazione Tribunale

Iscritto nel Registro Stampa del Tribunale di Roma il 15.3.2012 al n. 56/2012
ISSN: 2282-233X

Editoriale

Alessandra Petrucci

Dipartimento di Statistica, Informatica, Applicazioni
Università degli Studi di Firenze

alessandra.petrucci@unifi.it

Questo numero speciale di *Statistica & Società* raccoglie le sintesi degli interventi dei relatori invitati – anche se non tutti presenti – alla tavola rotonda “Le esperienze e le evidenze della valutazione della ricerca nell’area 13 CUN” tenuta a Cagliari nell’ambito della XLVII Riunione Scientifica della Società Italiana di Statistica.

Il tema della tavola rotonda sulla valutazione della ricerca nella nostra area riveste grande interesse ed è senza dubbio di forte attualità. Si sono, infatti, svolti e si stanno tuttora svolgendo due grandi momenti per la vita scientifica delle organizzazioni universitarie e dei singoli componenti: l’esercizio di valutazione della qualità della ricerca (VQR) e le prime due tornate dell’abilitazione scientifica nazionale (ASN). Tuttavia, al tempo in cui si è realizzata la tavola rotonda erano stabili solo i risultati della VQR e le relazioni si riferiscono in modo diretto agli aspetti che riguardano questo esercizio.

La volontà di organizzare questo momento di discussione che comprendesse tutte le “voci” dell’Area 13 (Scienze economiche e statistiche) è sostenuta dal desiderio di portare a conoscenza della comunità scientifica degli statistici (di ogni raggruppamento disciplinare) un esempio del metodo di lavoro adottato negli otto anni di mandato come rappresentante al Consiglio Universitario Nazionale (CUN) nell’affrontare criticità e ricercare soluzioni sempre attraverso il dialogo, la riflessione comune e soprattutto la conoscenza delle peculiarità dei diversi “macrosettori” che compongono l’area che, per sua definizione, è un’area assai variegata, complessa, portatrice di differenti istanze non sempre in armonia.

Il contributo di **Alessandra Petrucci** illustra in modo sintetico alcuni dei risultati che sono emersi dalla Consultazione Pubblica organizzata dal CUN nel 2013 e tesa ad acquisire le indicazioni e le osservazioni di quanti, in ragione delle loro competenze ed esperienze scientifiche e professionali, volessero contribuire all’assunzione delle migliori determinazioni in merito ai requisiti che presiedono al riconoscimento di scientificità delle pubblicazioni e degli altri prodotti della ricerca.

Aldo Pavan, componente del direttivo dell'Accademia Italiana di Economia Aziendale (AIDEA), mette in luce come l'attività di valutazione rileva per i comportamenti che induce più che per le classifiche che realizza e spinge verso percorsi di crescita della ricerca che siano originali e non di mera omologazione.

Il contributo di **Donato Posa**, che ha coordinato la tavola rotonda, ripercorre i processi valutativi che hanno riguardato la ricerca italiana dell'Area 13, i criteri utilizzati e le ripercussioni sui contributi derivanti da specifici interessi di ricerca e individua come elemento di forte criticità la classificazione delle riviste.

Maurizio Carpita, già componente del consiglio direttivo della SIS, affronta in modo più specifico la valutazione della collocazione editoriale dei prodotti della ricerca. In particolare, viene descritta la procedura utilizzata dalla SIS per la predisposizione delle classifiche delle riviste compiuta nel 2012.

Francesco Bartolucci, componente del Gruppo di Esperti della Valutazione (GEV) dell'Area 13, analizza in dettaglio gli aspetti peculiari che concorrono alla classificazione di una rivista scientifica ipotizzando infine un modello basato su difficoltà nel pubblicare in una certa rivista come misura collegata al livello reputazionale.

Il contributo di **Alberto Baccini**, Presidente dell'Associazione Italiana per la Storia dell'Economia Politica (STOREP), offre invece una sintesi critica dell'esercizio di Valutazione della Qualità della Ricerca (VQR 2004-2010) sottolineando le molte criticità emerse, in particolar modo nell'Area 13, di carattere metodologico e tecnico e mostrando la distorsione dei risultati.

Daniela Cocchi, già membro del Consiglio Nazionale dei Garanti per la Ricerca (CNGR) e già presidente della SIS, illustra le peculiarità della valutazione delle discipline statistiche, saldamente inserite nell'Area 13 del CUN, proponendo una strategia per ampliare la gamma delle riviste di interesse per gli statistici e sottolineando la necessità di definire collegamenti con gli argomenti di ricerca classificati secondo l'European Research Council (ERC).

In conclusione, questo numero di *Statistica & Società* mostra un panorama ampio e differenziato di “visioni” della valutazione della ricerca nell'Area 13 e costituisce un esempio al momento unico di raccolta di esperienze e riflessioni su questo argomento che mi auguro possa costituire un'apertura di nuovi confronti anche appassionati ma soprattutto produttivi e con l'obiettivo di “guardare avanti”.

Il riconoscimento della scientificità delle riviste: l'esperienza del CUN attraverso la Consultazione Pubblica

Alessandra Petrucci

Consigliere CUN – Area 13 Scienze Economiche e Statistiche
Dipartimento di Statistica, Informatica, Applicazioni
Università degli Studi di Firenze

alessandra.petrucci@unifi.it

Sommario

Il 23 luglio 2013 si è chiusa la Consultazione Pubblica telematica per l'istituzione dell'Anagrafe Nazionale Nominativa dei Professori e dei Ricercatori e delle Pubblicazioni Scientifiche (ANPrePS). Con questa iniziativa il Consiglio Universitario Nazionale (CUN) ha inteso acquisire indicazioni e osservazioni in merito ai requisiti che presiedono al riconoscimento di scientificità delle pubblicazioni e degli altri prodotti della ricerca che, per ogni professore ordinario o associato e per ogni ricercatore, devono essere contenute nell'ANPrePS.

1. Introduzione

Secondo la definizione che ne fornisce la Commissione europea (Commissione Europea, 2002, p. 14), le consultazioni pubbliche possono essere considerate come il processo attraverso il quale le amministrazioni intendono coinvolgere le parti interessate all'elaborazione delle politiche, prima che la decisione finale venga presa. La consultazione così intesa viene quindi identificata non già come una pratica concertativa, né solo come una forma di pubblicizzazione di scelte politiche già confezionate, bensì quale vero e proprio *metodo di indagine*, che ha lo scopo di rendere esplicite le premesse decisionali, consentendo la raccolta di dati e opinioni e la circolazione delle informazioni, per mettere in luce – eventualmente in una fase più avanzata della ricerca – i vantaggi e gli svantaggi percepiti rispetto alle diverse ipotesi di intervento. Con la Consultazione Pubblica, il Consiglio Universitario Nazionale ha voluto rendere partecipato e trasparente il percorso decisionale volto alla formulazione della proposta che, ai sensi dell'art. 3---ter, comma 2, della legge 9 gennaio 2009 n. 1, è chiamato a presentare al Ministro dell'Istruzione, dell'Università e della Ricerca, perché, sentita l'Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (ANVUR), si stabiliscano, con apposito decreto ministeriale, i «criteri identificanti il carattere scientifico delle pubblica-

zioni». L'obiettivo era quello di raccogliere contributi funzionali all'individuazione dei «requisiti minimi» idonei a identificare il carattere scientifico delle pubblicazioni e degli altri prodotti della ricerca, ai fini del loro riconoscimento e inserimento nell'Anagrafe nazionale nominativa dei Professori e dei Ricercatori e delle pubblicazioni scientifiche (ANPrePS). Stanti le finalità della Consultazione Pubblica, non rientrava, fra gli obiettivi, l'acquisizione di indicazioni e osservazioni in merito alla valutazione qualitativa dell'attività di ricerca scientifica.

In questo lavoro si riporta una sintesi dell'Autore degli esiti della Consultazione Pubblica che riguardano solo la sezione rivolta specificatamente alle riviste scientifiche.

Infine, si sottolinea che il questionario non ha inteso, in alcun modo, rappresentare orientamenti, scelte e soluzioni già approvate e adottate dal Consiglio Universitario Nazionale (CUN).

2. Caratteristiche della Consultazione Pubblica

La Consultazione è stata condotta tramite un'indagine *online* le cui domande erano state predisposte al solo fine di acquisire osservazioni in merito alle principali questioni concernenti l'identificazione dei criteri di scientificità delle pubblicazioni e la

loro applicazione, nonché l'individuazione degli ambiti ai quali riferirli e delle condizioni ai quali subordinarli. L'indagine ha avuto come destinatari il personale strutturato e non strutturato delle Università, degli Istituti Speciali e delle Scuole Superiori, degli Istituti di Alta Formazione Dottorale, e degli Enti di ricerca pubblici e privati nazionali, regionali e locali; i dottorandi e dottori di ricerca, gli editori attivi nel settore dell'editoria accademica e professionale, i componenti dei Nuclei di Valutazione degli Atenei, degli Osservatori di Ateneo o di analoghe strutture comunque denominate. L'accesso al questionario era libero e non prevedeva nessuna procedura previa di registrazione. Era tuttavia necessario fornire in chiusura informazioni sulla propria attività e dati personali, in assenza dei quali le risposte fornite non sarebbero state prese in considerazione. Il questionario *online* era articolato in quattro sezioni per un totale di 63 domande: la prima sezione concernente le pubblicazioni scientifiche, gli elementi che ne caratterizzano la scientificità, le tipologie di pubblicazioni e la loro diffusione; la seconda relativa alle riviste scientifiche, alle modalità e alle procedure di revisione dei contributi ospitati, alla diffusione e all'applicazione dei criteri di scientificità; la terza, dedicata ai prodotti della ricerca diversi dalle pubblicazioni scientifiche; la quarta necessaria a raccogliere i dati

di natura anagrafica e quelli idonei a qualificare i partecipanti. La maggior parte delle domande richiedeva di indicare una sola fra diverse opzioni. Erano presenti anche alcune risposte multiple. Non era obbligatorio rispondere a tutte le domande. In corrispondenza di tutte le sezioni e sottosezioni, vi era la possibilità di intervenire, in appositi spazi, con contributi liberi. Per maggiori informazioni si rimanda al Rapporto reperibile nel sito del CUN (www.cun.it)

Nei 90 giorni di apertura della Consultazione, avviata il 23 Aprile 2013 e chiusa il 23 Luglio 2013, circa 17.000 persone si sono collegate almeno una volta al sito della Consultazione.

Complessivamente sono stati registrati 16.195 questionari, di cui 10.443 iniziati ma non completati e 5.752 completati. I 5.752 questionari correttamente completati sono stati sottoposti a un processo di validazione predefinito al fine di accertare l'affidabilità delle risposte. Sono stati individuati 5.678 questionari considerati attendibili e completi delle informazioni richieste e solo su questi ultimi è stata condotta l'analisi.

I partecipanti alla Consultazione sono risultati nell'89% personale strutturato dell'Università, appartenente a tutte le aree CUN anche se in percentuale diversa (Fig. 1). Il restante 11% invece è composto da figure che non appartengono al mon-

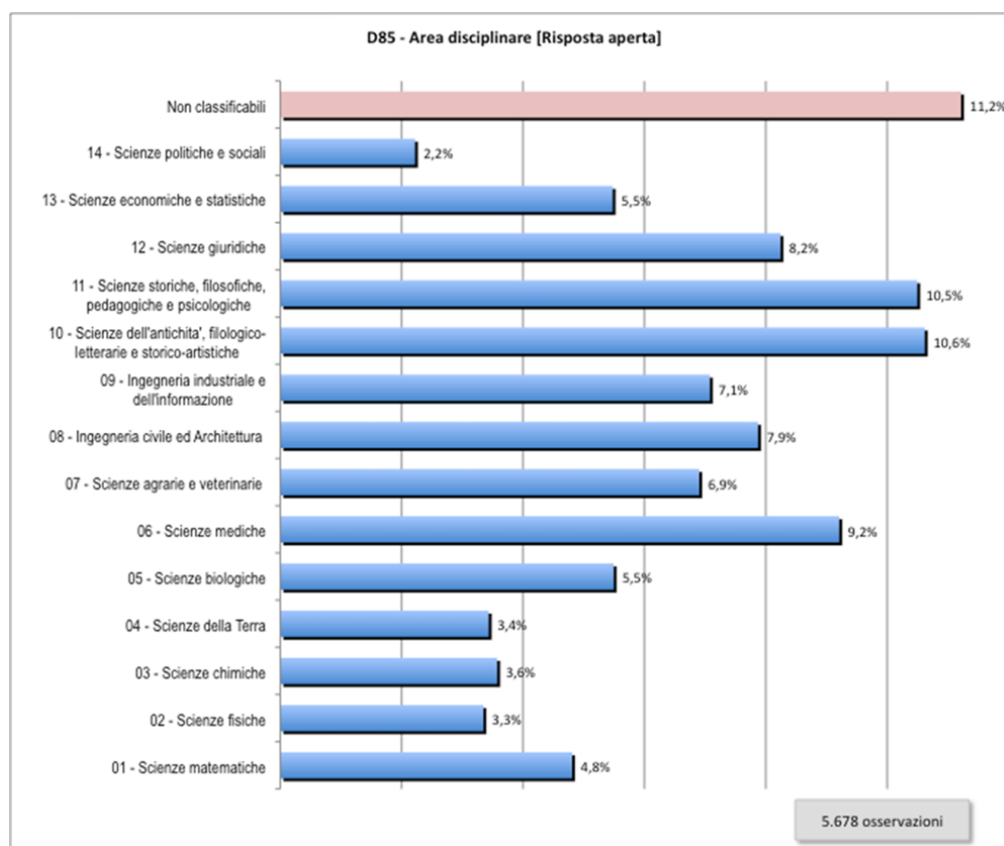


Figura 1: Area disciplinare dei rispondenti accademici

do accademico (es. editori specializzati, ecc.).

Al personale docente e ricercatore accademico è stato chiesto di indicare oltre al settore scientifico disciplinare (SSD) di afferenza anche la propria posizione all'interno dell'Università. Il 27% dei rispondenti è professore ordinario, il 26% è professore associato mentre la quota dei ricercatori è leggermente superiore (31%).

Per cercare di caratterizzare meglio il rispondente è stato chiesto quali sono le tipologie di pubblicazione scientifica nelle quali si è maggiormente espressa l'attività scientifica dei rispondenti. La modalità maggiormente scelta è stata il contributo in rivista seguita da contributo in volume e la monografia (era possibile indicare più tipologie). È stato quindi chiesto ai partecipanti se siano stati impegnati in attività di *peer review* negli ultimi tre anni e il 77% ha risposto positivamente.

3. La sezione riviste scientifiche

Come già ricordato il questionario era suddiviso in

più sezioni. In particolare, vengono riportati alcuni degli esiti riguardanti la seconda sezione che propone domande volte a: identificare i criteri perché una rivista possa considerarsi scientifica; verificare quali condizioni siano atte a garantirne una congrua diffusione; valutare le modalità di applicazione dei criteri di scientificità.

Al fine di identificare le procedure e le modalità di revisione considerate adeguate a certificare la scientificità di una rivista viene chiesto ai partecipanti di pronunciarsi in merito all'importanza di quelle proposte (Fig. 2).

Sono considerate particolarmente qualificanti quelle procedure di revisione per l'accettazione dei contributi che prevedono l'anonimato dei revisori (79,0%), anche nella forma del "doppio cieco" (82,0%). Altrettanto qualificante è giudicata la previsione di revisioni tramite raccolta di pareri pubblici, volontari o sollecitati dal comitato scientifico-editoriale (74,1%). È stato quindi chiesto se sia necessario individuare altri criteri per definire la scientificità di una rivista. Solo il 12% ritiene utile individuarne anche altri.

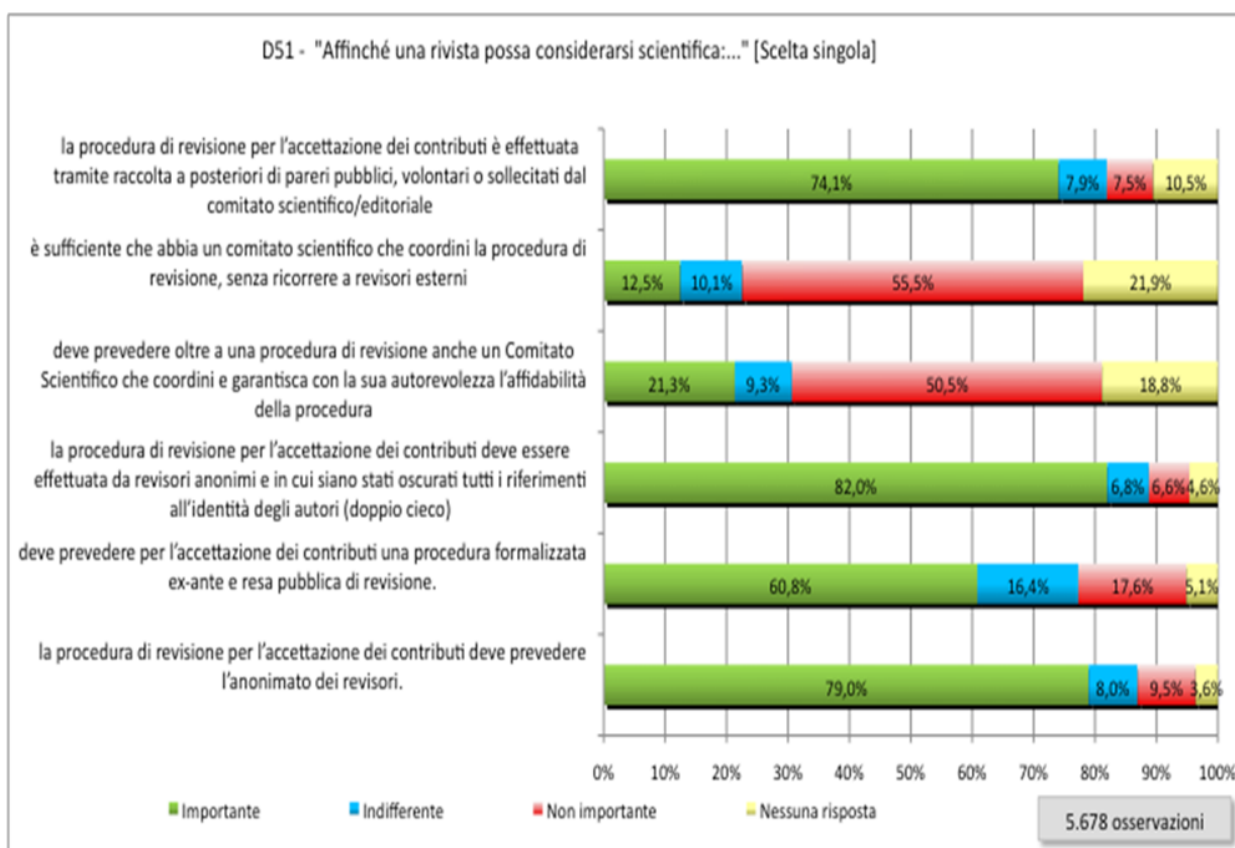


Figura 2: Procedure di revisione per l'accettazione dei contributi in rivista

Per quanto riguarda la diffusione, viene chiesto ai partecipanti di esprimersi, secondo una scala d'importanza, in merito ad alcune condizioni che pos-

sano ritenersi idonee a garantire una congrua diffusione di una rivista scientifica (Fig. 3).

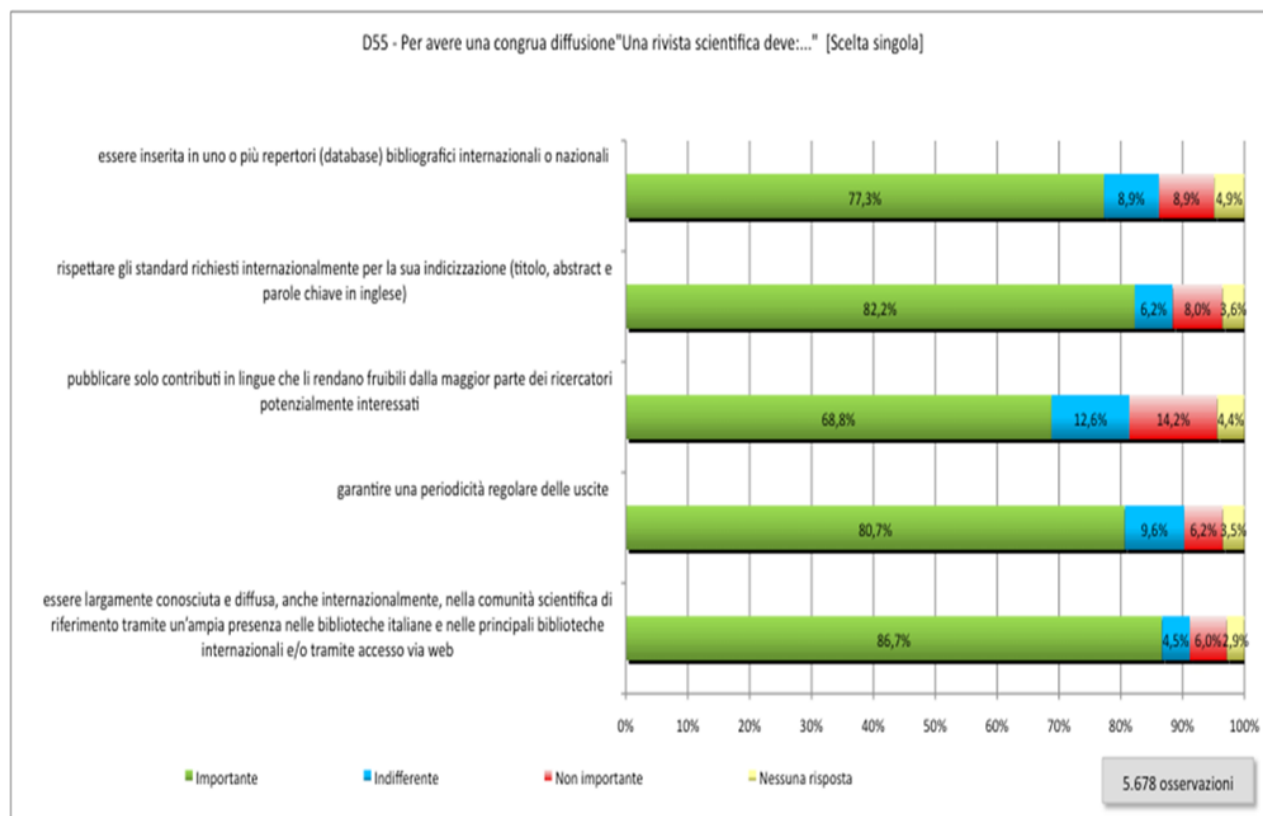


Figura 3: Condizioni per garantire una congrua diffusione di una rivista scientifica

Tutte le modalità indicate sono state ritenute importanti dalla maggior parte dei partecipanti: l'86,7% ritiene che sia importante la presenza in biblioteche italiane e nelle principali biblioteche internazionali o che ne sia comunque garantito l'accesso via web; l'82% ritiene che sia importante il rispetto degli standard richiesti internazionalmente per la sua indicizzazione; l'80% ritiene importante che sia assicurata una periodicità regolare delle uscite; il 77,3% ritiene importante l'inserimento in uno o più data base internazionali o nazionali; il 68,8% ritiene importante che siano pubblicati solo contributi in lingue che li rendano fruibili dalla maggior parte dei ricercatori potenzialmente interessati. Per verificare quali siano gli ambiti e le possibili finalità applicative dei criteri di scientificità delle riviste, oggetto dei precedenti quesiti, si chiede perciò ai partecipanti se, a loro avviso, una pubblicazione possa considerarsi scientifica per il solo fatto di apparire in una rivista considerata scientifica.

Le risposte sono sostanzialmente negative: per il 54% dei rispondenti la presenza di un articolo in una rivista considerata scientifica non è condizione sufficiente a garantirne la scientificità (Fig. 4).

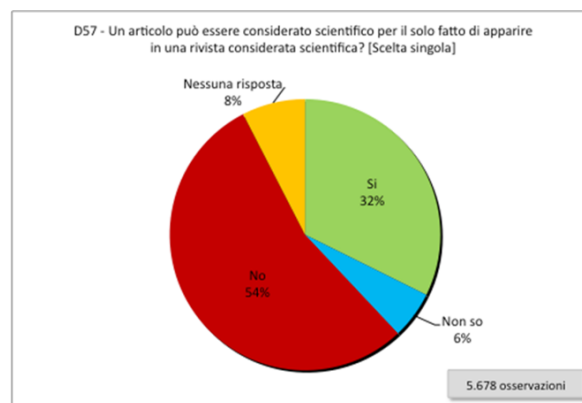


Figura 4: Distribuzione percentuale dei rispondenti alla domanda circa la scientificità di un articolo pubblicato in una rivista considerata scientifica

Si tratta di un'indicazione dalla quale si evince come gli indicatori esterni di scientificità non possano essere assunti come criteri per la valutazione qualitativa del prodotto, ossia per scopi differenti dal mero riconoscimento dei requisiti minimi di scientificità necessari ad autorizzare la collocazione delle pubblicazioni in banche dati, quali l'AN-PrePS.

4. Conclusione

Gli esiti della Consultazione Pubblica condotta dal CUN sui criteri idonei a identificare il carattere scientifico delle pubblicazioni e degli altri prodotti della ricerca, ai fini del loro inserimento e della loro collocazione nell'ANPrePS, consentono di proporre alcune considerazioni conclusive in merito a quale sia l'atteggiamento e quali gli orientamenti diffusi in proposito nella comunità scientifica. L'elevata partecipazione, da considerarsi tale anche in ragione della complessità dell'indagine, articolata in 63 domande cui si aggiungeva la possibilità, di cui molti si sono avvalsi, di approfondire e precisare le proprie posizioni con contributi liberi, testimonia la grande attenzione prestata dalla comunità scientifica al tema. I partecipanti, appartenenti per l'89% al mondo accademico, hanno mostrato consapevolezza e conoscenza delle questioni di volta in volta portate al loro esame, manifestando di condividere nella totalità i principi che, anche in ambito internazionale, presiedono al riconoscimento della scientificità.

Per quanto riguarda il riconoscimento dei criteri di scientificità delle riviste, emerge una chiara indicazione a ritenere che nessun requisito esterno, ap-

partenente alla sede di pubblicazione, possa essere assunto quale criterio per la valutazione qualitativa dei contributi in essa accolti, potendo essere utilizzato solo per le finalità dell'inserimento nell'ANPrePS.

Riferimenti bibliografici

Commissione europea, 2002, *Verso una cultura di maggiore consultazione e dialogo. Principi generali e requisiti minimi per la consultazione delle parti interessate ad opera della Commissione*, COM(2002)704 def., Bruxelles.

CUN, 2013, *Consultazione Pubblica per l'istituzione dell'Anagrafe Nazionale Nominativa dei Professorie dei Ricercatori e delle Pubblicazioni Scientifiche (ANPrePS)*, Roma (https://www.cun.it/uploads/4533/rapporto_esiti_consultazione_pubblica_cun.pdf?v=)

Presidenza del Consiglio. Dipartimento della funzione pubblica. Ufficio per l'innovazione delle pubbliche amministrazioni, 2001, *La consultazione nell'analisi dell'impatto della regolazione*, a cura di Antonio La Spina e Sabrina Cavatorto. Soveria Mannelli, Rubbettino.

Valutazione della Ricerca Scientifica e Comportamenti Accademici

Aldo Pavan

Consiglio Direttivo AIDEA – Accademia Italiana di Economia Aziendale
Dipartimento di Scienze Economiche e Aziendali
Università degli Studi di Cagliari

apavan@unica.it

Sommario

L'attività di valutazione rileva per i comportamenti che induce più che per le classifiche che realizza. L'eccessiva enfasi sulla ricerca riduce l'interesse degli universitari verso la didattica, il servizio al territorio e all'accademia e impoverisce l'utilità sociale delle università. Il rilievo attribuito alle riviste di fascia A appiattisce la ricerca nazionale sugli interessi dei relativi *editor*. Occorre favorire percorsi di crescita della ricerca che siano originali e non di mera omologazione.

1. Premessa

Le attività umane sono spesso oggetto di valutazione, a livello sia individuale sia collettivo. Nelle organizzazioni le prestazioni di uffici, reparti ecc. sono oggetto di misurazione tanto nei risultati raggiunti quanto nelle risorse utilizzate. Tali misurazioni consentono la formazione di classifiche e l'attribuzione di riconoscimenti che, infine, riguardano anche i singoli individui. La logica di tali azioni non è tuttavia finalizzata alle classifiche e ai premi, ma cerca di influire sui comportamenti così che, dato un sistema premiante ben conosciuto a preventivo, ci si attende che i comportamenti si adeguino a esso e gli obiettivi dell'organizzazione siano perseguiti con maggior efficacia ed efficienza. Tale argomento è, in linea di principio, valido per le università e nell'ambito delle scienze sociali; le une e le altre presentano peculiarità non trascurabili. L'attività accademica è tra le più complesse che la società esprima, non solo e tanto per la profondità intellettuale richiesta, ma anche e forse soprattutto per la molteplicità di compiti che la compongono. Un osservatore può vedere il professore universitario prima di tutto quale produttore di nuova conoscenza o quale operatore dell'alta formazione. Eppure quello stesso professore è normalmente chiamato ancora a dare servizio al proprio ateneo nelle diverse funzioni manageriali o professionali che gli statuti prevedono; non pochi sono impegnati in organismi nazionali e internazionali, nelle società scientifiche, ecc. Di norma molto apprezzata e incentivata è l'attività che professori e ricercatori svolgono in collaborazione con le istituzioni pubbliche e private dei territori di riferimento

e oltre. Normalmente il sentire comune, dentro e fuori delle università, attribuisce valenza maggiore all'attività di ricerca così che le progressioni di carriera, ovunque nel mondo, sono funzione della qualità e quantità di prodotti scientifici riconducibili al singolo accademico. La ricerca scientifica è osservabile quale attività preminente dell'universitario, ma è inscindibilmente legata alle altre di didattica e servizio, tanto all'accademia quanto al territorio. Fare ricerca scientifica significa produrre nuova conoscenza nell'ambito costituito dalla propria disciplina scientifica. Il prodotto materiale di tale attività è prosaicamente chiamato prodotto della ricerca e si concretizza in un articolo su rivista, un'opera monografica, un brevetto ecc. La qualità di tali prodotti è, per definizione, valutata "tra pari", vale a dire da altri soggetti appartenenti alla data comunità scientifica. Chi può infatti esprimersi con competenza su contributi che spostino in avanti le conoscenze in campo matematico se non i matematici? Tale ovvia osservazione ha necessità di svilupparsi in prassi operative virtuose che assicurino valutazioni effettuate con onestà intellettuale e secondo procedure che assicurino disponibilità verso ogni approccio, tanto tradizionale quanto innovativo o addirittura rivoluzionario. Le comunità scientifiche sono pertanto in perenne ricerca di soddisfacenti procedure di valutazione. In assenza di criteri ottimi, si ritiene comunemente necessario che le regole di valutazione della produttività scientifica siano note a priori e condivise. La ricerca nell'ambito delle scienze sociali assume caratteri peculiari rispetto alle scienze della natura. Il ricercatore studia comportamenti dei quali è partecipe. Non sappiamo se e quanto la realtà sociale

sia data per sé e possa essere scoperta o sia frutto dell'attività intellettuale e pertanto modificata o addirittura costruita dalla investigazione. I confini tra essere e dover essere sono tutt'altro che chiari e definiti. Il ricercatore sociale si trova di fatto maggiormente legato alle realtà nelle quali vive; ostacoli di tipo culturale, linguistico, logistico ecc. rendono spesso arduo lo studio di realtà sociali lontane.

Le tradizioni nazionali convivono da sempre con la tendenza alla internazionalizzazione e dunque all'unitarietà delle discipline. Quest'ultimo fenomeno è andato rafforzandosi negli ultimi decenni in coerenza con la forte accelerazione dei processi di globalizzazione della società e dell'economia; sono di conseguenza evidenti le tendenze all'omologazione verso i modelli di comportamento propri dei paesi dominanti.

2. Attività di AIDEA

L'Accademia Italiana di Economia Aziendale – AIDEA – è la società scientifica di riferimento di oltre 1.800 aziendalisti italiani, costituenti nel complesso poco meno del 40% di tutti i docenti di area 13. L'Accademia condivide senza riserve l'esigenza che la valutazione della qualità della ricerca sia rafforzata anche nel senso di una continua attività di revisione critica e di adeguamento all'evoluzione delle discipline. In tale ottica i processi di internazionalizzazione assumono un ruolo irrinunciabile in quanto volti all'ampliamento del confronto e all'allineamento verso l'alto delle comunità scientifiche nazionali.

Tale posizione è stata affermata in svariati documenti pubblici, disponibili nella pagina web. In tale direzione AIDEA ha posto in essere molteplici comportamenti concreti. Tra questi, nel 2007 ha avviato un progetto di classificazione delle migliori riviste scientifiche internazionali (Journal Rating AIDEA); il primo documento è stato poi aggiornato nel 2009 e nel 2012. Nel 2009 ha avviato un processo di progressivo accreditamento delle Riviste Nazionali: sono 27 le riviste che hanno superato le verifiche condotte sul piano dei requisiti scientifici, gestionali, organizzativi e di diffusione fissati a priori. Dal 2013 ha avviato un simile progetto di accreditamento di collane di monografie. Sempre dal 2013 ha avviato il progetto "Tassonomia degli studi aziendali" al fine di definire una mappa dei saperi e delle discipline coerente con la frontiera della conoscenza scientifica moderna.

3. Le esperienze recenti di valutazione della ricerca

L'esercizio VQR 2004-10 ha costituito un momento di svolta dell'intero sistema della ricerca nazionale. Esso è stato condotto secondo criteri e modalità che sono stati oggetto di severe critiche. È stato da più parti osservato che "le regole del gioco" sono state rese note troppo tardi, così che sono stati misurati e valutati comportamenti posti in essere quando erano vigenti criteri tradizionali, spesso molto differenti. I criteri della VQR sono diversi rispetto a quelli dell'Abilitazione Scientifica Nazionale – ASN – così che i risultati di valutazione sono stati disomogenei e spesso in conflitto. Le società scientifiche hanno in generale lamentato lo scarso dialogo con l'Agenzia di Valutazione – ANVUR. I criteri bibliometrici sono stati utilizzati anche nei settori definiti non bibliometrici, secondo modalità che sono state oggetto di rilevanti critiche metodologiche. Anche nell'area aziendale si lamenta che non vi è stato alcun coinvolgimento delle società scientifiche, né di AIDEA né delle altre specifiche di settore: il GEV 13 ha di fatto operato in condizioni di isolamento. Non può pertanto sorprendere se la comunità degli aziendalisti ha percepito i risultati di tale esercizio come qualcosa di estraneo e punitivo. In presenza di una tradizione centrata prevalentemente sulle monografie in lingua italiana e sui metodi qualitativi, le prime sono state penalizzate sin dalla stessa impostazione delle schede di valutazione; i secondi sono stati sfavoriti dal prevalente riferimento a riviste a dominio nord americano, da sempre orientate ai metodi quantitativi. L'insoddisfazione per le modalità di svolgimento della VQR è stata poi accentuata dal fatto che i suoi risultati stanno avendo, di fatto, ricadute di grande rilievo sui diversi aspetti della vita universitaria, quali la distribuzione delle risorse inter-ateneo e l'accreditamento dei dottorati.

4. Prospettive

La valutazione della ricerca è svolta, si è detto, tra pari; essa non può pertanto prescindere dal coinvolgimento delle specifiche comunità scientifiche. Tale coinvolgimento attiene tanto ai criteri da adottare nei processi di valutazione, quanto alla rappresentatività dei soggetti chiamati a svolgere i ruoli di rilievo in tali processi – es. composizione GEV, scelta dei revisori. I criteri bibliometrici semplificano i processi di valutazione, ma pongono il fuo-

co esclusivamente sui “contenitori” piuttosto che sui contenuti. Il loro utilizzo in settori che non ne fanno abituale uso può facilmente portare a risultati paradossali. È evidente il rischio che le agende di ricerca siano improntate opportunisticamente a criteri di mera “resa bibliometrica”.

La necessità di internazionalizzazione è condivisa e fortemente voluta dalla comunità degli aziendalisti, si ritiene tuttavia che essa 1) non possa essere esasperata sulle riviste di maggior prestigio internazionale, 2) debba convivere con la dimensione nazionale.

Le riviste di fascia A, di natura strettamente aziendale, sono in numero limitato, tale da non consentire uno “spazio” fisico sufficiente; sono poi a prevalente dominanza nord americana: in esse i metodi quantitativi assumono un’enfasi che in non pochi casi sembra prevaricare sulla rilevanza degli oggetti e sul merito dei contenuti. Un orientamento a esse di una comunità di oltre 1800 professori e ricercatori 1) non è materialmente realizzabile, 2) porterebbe a un totale appiattimento delle agende di ricerca con quanto è in tali ambienti considerato di interesse e 3) alla perdita di contatto con le imprese e le altre organizzazioni economiche nazionali e dei territori. Si ritiene pertanto che si debba procedere con gradualità nel processo di crescita scientifica e internazionalizzazione, attraverso 1) una valutazione

meno elitaria delle pubblicazioni internazionali che sposti il *divide* attuale tra fascia A e “tutto il resto” per includere tra i prodotti di pregio le pubblicazioni che assicurino comunque rigore nel processo di accettazione; 2) la valorizzazione dei prodotti scientifici in lingua italiana; 3) l’internazionalizzazione delle riviste e collane nazionali. Le discipline aziendali hanno una forte tradizione scientifica basata sulle monografie in lingua italiana. Non è auspicabile che esse siano totalmente marginalizzate a favore degli articoli su rivista. Occorre per esse procedere sulla strada della valorizzazione attraverso percorsi rigorosi e affidabili di valutazione e accettazione. I criteri di valutazione VQR e ASN devono essere omogeneizzati così da dare ai giovani ricercatori messaggi univoci su cosa le loro comunità scientifiche intendano per “buona ricerca”. Pur riconosciuta la preminenza della ricerca scientifica tra le attività accademiche, è necessario che essa sia inquadrata in un contesto di giudizio di singoli e gruppi che tenga conto delle altre componenti dell’attività accademica: didattica, servizio all’accademia, servizio al territorio. Ciò al fine di favorire un percorso di crescita internazionale che sia realizzato mantenendo l’utilità sociale delle università e l’originalità delle scuole italiane e contrastando i processi di loro mera omologazione ai modelli dominanti di origine anglo-americana.

Processi valutativi, qualità dei prodotti e interessi di ricerca

Donato Posa, Sandra De Iaco, Sabrina Maggio, Monica Palma

Dipartimento di Scienze dell'Economia - Sez. Matematico-Statistica
Università del Salento

donato.posa@unisalento.it, sandra.deiaco@unisalento.it, sabrina.maggio@unisalento.it,
monica.palma@unisalento.it

Sommario

In questo lavoro, si intende ripercorrere i processi valutativi che hanno riguardato la ricerca italiana dell'Area 13, i criteri utilizzati e le ripercussioni sui contributi derivanti da specifici interessi di ricerca. Si conclude con una proposta di nuovi criteri di valutazione in grado di supportare processi valutativi trasparenti, basati su principi noti *ex-ante*, volti a valorizzare gli svariati interessi di ricerca, quelli di ampio respiro e quelli di nicchia.

1. Introduzione

Le attività di valutazione dei prodotti della ricerca che hanno coinvolto e sconvolto il mondo accademico italiano rientrano, a vario titolo, nei processi valutativi degli atenei e dei singoli ricercatori. I criteri di massima sono spesso condivisibili in linea generale: infatti, come sottolinea la stessa ANVUR, il “giudizio di qualità” si deve basare sui “criteri di rilevanza, originalità/innovazione, internazionalizzazione”, ovvero sul “posizionamento del prodotto della ricerca nel panorama internazionale”. Fino a questo punto, non si è aggiunto nulla di nuovo, dal momento che molti gruppi di ricerca di tutta Italia lavorano da decenni e decenni nel rispetto di questi principi. Del resto il CUN in un suo documento del maggio del 2011, sottolineava che l'Area 13, dal punto di vista scientifico e della ricerca, è naturalmente caratterizzata dai seguenti aspetti:

- produzione scientifica prevalentemente su base pubblicistica (articoli su riviste, monografie, libri, capitoli di libri, saggi in volumi collettanei);
- pubblicazioni in gran parte in lingua italiana e inglese (talvolta in francese);
- crescente presenza, per diversi SSD, di riviste che rispondono a criteri bibliometrici internazionalmente riconosciuti, anche se esiste ancora una presenza rilevante di riviste, anche internazionali, che non rispondono a tali criteri;
- uso di metodologie di valutazione consolidate dalla tradizione;

- utilizzo, nell'ambito delle metodologie di valutazione, anche di criteri basati sulla rilevanza e sulla qualità della sede editoriale di pubblicazione.

Il CUN ha anche evidenziato che la grande articolazione delle discipline interne all'Area 13, consente solo parzialmente di adottare criteri e riferimenti univoci nella valutazione dei risultati della ricerca scientifica e, in particolare, pesano i diversi oggetti di studio, i diversi metodi prevalenti, il diverso tasso di internazionalizzazione degli studi, nonché le diverse modalità pubblicistiche, come sopra precisato.

In questo lavoro, dopo una breve digressione sui dettami del bando VQR 2004-2010, relativamente alle tipologie di prodotti della ricerca da sottoporre a valutazione, e sulle modalità di valutazione degli stessi prodotti, viene presentata una discussione dettagliata sui risultati dell'esercizio di valutazione della qualità della ricerca, nonché sulle criticità emerse nell'intero processo valutativo. Infine, viene proposto il ricorso a nuovi criteri di valutazione che, riconoscendo l'ampia eterogeneità dell'Area 13, possano debitamente valorizzare gli interessi di ricerca di ogni studioso afferente a tale Area.

2. Prodotti della ricerca e criteri di valutazione

Con la pubblicazione del Bando VQR 2004-2010, avvenuta nel novembre 2011, l'ANVUR ha inteso

avviare l'esercizio della Valutazione della Qualità della Ricerca in Italia. In particolare, per ciascuna delle 14 Aree disciplinari definite dal Comitato Universitario Nazionale (CUN), l'ANVUR ha nominato un Gruppo di Esperti della Valutazione (GEV), composto da studiosi, anche stranieri, di riconosciuta esperienza e qualità scientifiche, che, di concerto con la stessa ANVUR, ha definito i criteri per la valutazione della qualità dei prodotti della ricerca afferenti a ricercatori (a tempo determinato ed indeterminato), assistenti, professori di prima e seconda fascia incardinati in tutti gli Atenei Italiani, nonché ricercatori e tecnologi degli Enti di Ricerca, e successivamente ha decretato le valutazioni per ogni Ateneo e per le rispettive Strutture Dipartimentali. Il numero dei componenti del GEV è stato definito dall'ANVUR in base al numero atteso di prodotti da valutare nelle diverse Aree disciplinari. Nel caso dell'Area 13 "Scienze Economiche e Statistiche", sono stati nominati ben 36 esperti valutatori, raggruppati nei Sub-GEV "Economia", "Economia Aziendale e Finanza" e "Statistica e Matematica Applicata".

È importante evidenziare che, tra gli esperti valutatori, 12 (33%) non risultano nell'elenco dei docenti e ricercatori del MIUR, ma sono studiosi appartenenti ad università straniere o istituti di ricerca non accademici. D'altra parte, dei 24 esperti "reclutati" dal mondo accademico nazionale, il 50% si occupa di discipline economiche (ovvero afferisce ai SSD SECS-P/01, SECS-P/02, SECS-P/03 e SECS-P/05), il 33% afferisce ai settori che si occupano di economia, organizzazione finanza aziendale (SSD SECS-P/07, SECS-P/08, SECS-P/09, SECS-P/10), mentre solo uno studioso appartiene al settore SECS-S/01 "Statistica".

Con riferimento ai prodotti della ricerca (articoli su riviste; libri, capitoli di libri ed atti di congressi dotati di ISBN; edizioni critiche, traduzioni e commenti scientifici; brevetti; composizioni, disegni, manufatti, prototipi, banche dati e software), i membri del GEV sono stati chiamati ad esprimere un giudizio di qualità che tenesse in considerazione la rilevanza, l'originalità/innovazione, l'internazionalizzazione. Il giudizio sintetico conclusivo, che è stato espresso per ben 11941 prodotti (il 63% di questi era rappresentato da articoli su rivista), è stato infine attribuito ad ogni prodotto della ricerca in base ad uno dei seguenti livelli di merito: Eccellente, Buono, Accettabile, Limitato, Non valutabile. Tuttavia, i processi valutativi della qualità della ricerca avviati con il bando VQR, si intrecciano e si combinano con quelli della ASN, dei processi di autovalutazione dipartimentale (SUA), di processi

di composizione degli organi collegiali. In effetti, per quanto sia stato ben chiarito che i risultati VQR riguardano l'Ateneo e le strutture dipartimentali, e non i singoli ricercatori, in realtà, in un modo o nell'altro se ne è tenuto conto nella valutazione dei singoli.

3. Risultati dell'esercizio di valutazione

L'esercizio di valutazione che ha interessato il mondo degli accademici, in veste sia di studioso valutatore, che di studioso valutato, ha avuto inizio nel momento in cui è stato necessario definire, in maniera puntuale, i criteri per la valutazione dei prodotti della ricerca, individuando *in primis* gli strumenti della valutazione. Come è noto, il bando della VQR prevedeva due strumenti di valutazione, ovvero lo strumento bibliometrico e la *peer review*. Quest'ultimo strumento è stato adottato per la valutazione di ogni prodotto della ricerca diverso da articolo su rivista, posto che lo stesso prodotto non sia stato preventivamente etichettato dal cosiddetto Gruppo di Consenso nominato all'interno del GEV, come prodotto "limitato" o, peggio, "non valutabile".

D'altra parte, per la valutazione degli articoli su rivista è stato utilizzato lo strumento bibliometrico e per un campione composto da almeno il 10% degli articoli, anche con il metodo della *peer review*. A tal fine, il GEV ha individuato, da un lato, un campione casuale di prodotti della ricerca – campione casuale stratificato (Cicchitelli, Herzel, Montanari, 1997; Cochran, 1977) per Sub-GEV – da valutare in *peer review* e, dall'altro, un elenco di riviste valido per ciascun Sub-GEV, nonché dei criteri, da utilizzare per la valutazione degli articoli pubblicati su rivista. Di seguito, viene fornita una breve descrizione della classifica realizzata dal Sub-GEV "Statistica e Matematica Applicata" nell'ambito della VQR, e viene proposta una comparazione con la classifica sviluppata dagli studiosi aderenti alla Società Italiana di Statistica (SIS), in maniera indipendente dalla prima e con criteri diversi e, talvolta, complementari a quelli adottati dal GEV. Si osservi che le classifiche appena menzionate hanno fatto riferimento a numerosità completamente differenti tra loro, ovvero

- Elenco VQR
n. riviste: 1907, di cui 449 (23,5%) di classe A;
n. riviste per l'Area Statistica: 445, di cui 114 (25,6%) di classe A;

- Elenco SIS per il SECS-S/01
n. riviste: 700, di cui 37 (5,3%) di classe Aa+ e Aa.

È interessante evidenziare, infine, che sulla base degli stessi criteri generali previsti dall'ANVUR, le riviste di classe A, censite per l'Abilitazione Scientifica Nazionale (ASN) 2012-2013, sono state numericamente superiori a quelle di classe A appartenenti all'elenco VQR. Si tratta infatti di ben 6600 riviste, con 2300 (35%) riviste di classe A, di cui 714 di Statistica. Resta poi da chiarire perché l'elenco delle riviste di classe A ai fini del l'ASN sia stato usato solo per il calcolo delle mediane e non per la valutazione dei candidati, che invece sono stati valutati sempre sulla base dell'elenco VQR.

3.1 Classifica VQR

L'elenco iniziale delle riviste predisposto per ciascun Sub-GEV dell'Area 13 (possibilmente anche per differenti aree di ricerca presenti in uno stesso Sub-GEV), è stato realizzato a partire dalla lista delle riviste presenti su ISI-Web of Science (WoS), che risultano essere rilevanti per le aree di ricerca del Sub-GEV. Tuttavia, tale elenco iniziale è stato ampliato per includere altre riviste che, pur non risultando su WoS, presentano un elevato impatto su Google Scholar, e sono individuabili dalla lista UGov del Cineca, ovvero risultano essere riviste su cui gli studiosi italiani hanno pubblicato nel periodo 2004-2010.

La classifica delle riviste, presenti nei suddetti elenchi, formulati per ciascun Sub-GEV, è stata stilata tenendo in considerazione la distribuzione di alcuni indicatori bibliometrici delle riviste stesse, quali l'*Impact Factor* (IF), l'IF a 5 anni (5IF) e l'*Article Influence Score* (AIS). In particolare, ricorrendo ai percentili della distribuzione di tali indicatori bibliometrici, sono state definite 4 classi di merito che, secondo un ordinamento decrescente, risultano essere A, B, C, D.

È immediato riconoscere alcuni punti di debolezza in questo impianto teorico di individuazione dei prodotti della ricerca meritevoli di valutazione, di definizione degli elenchi delle riviste e di classificazione delle stesse. In particolare, è interessante evidenziare quanto segue.

1. Un prodotto della ricerca è un concetto talmente ampio, tanto da essere riconosciuto giustamente in numerose forme di produzione dell'attività scientifica di uno studioso (articoli su rivista, libri e capitoli di libri, invenzioni,

brevetti e così via). Assegnare ad un ristretto Gruppo di Consenso (composto soltanto da due membri del GEV) il compito di stabilire se un prodotto della ricerca diverso dall'articolo su rivista, ad esempio un volume, sia "limitato" e come tale non meriti la valutazione in *peer review*, determina certamente un declassamento dello stesso prodotto scientifico, il quale, al contrario, potrebbe contenere approfondimenti teorico-pratico di valore scientifico almeno pari a quelli contenuti in un articolo pubblicato su rivista.

2. Non è da trascurare il possibile disallineamento tra i contenuti di U-Gov, utilizzati per stilare l'elenco delle riviste ed i prodotti successivamente consegnati per la valutazione. Alcuni prodotti pubblicati su riviste non censite, e quindi non presenti nell'elenco delle riviste "elette", sono stati valutati "limitati", sebbene le riviste in questione a) fossero riconosciute dall'ISI e con un IF ben superiore alle riviste censite presenti nel primo quartile (della graduatoria in ordine decrescente), b) fossero presenti negli elenchi di altre aree disciplinari e si collocassero nel primo o nel secondo quartile di tali elenchi.
3. L'emanazione di un bando che *ex-post* fornisca genericamente l'indicazione di prodotto da considerarsi della ricerca e dei criteri di valutazione e che demandi ad un gruppo di esperti il compito di stilare un elenco di riviste "elette" con l'attribuzione della sotto-area di riferimento, non è condivisibile.

3.2 Classifica SIS

Nel periodo 2005-2012, il Direttivo della SIS, le Commissioni di lavoro appositamente create, la base societaria, hanno prestato estrema attenzione al tema della valutazione della ricerca, giungendo dapprima a definire due liste (Frosini, 2008): una relativa alle riviste scientifiche di Statistica (SSD SECS-S/01-02-03 e 05) e l'altra relativa alle riviste di Demografia (SSD SECS-S/04). Tali riviste sono state classificate, a seconda della rinomanza, della diffusione e della qualità scientifica (eccellente, mediamente buona/elevata, mediamente accettabile/discreta) loro riconosciute, in 3 categorie (rispettivamente AA, A, B). Al fine di consolidare tali classifiche, è stata effettuata successivamente, tra tutti i soci SIS, un'indagine volta a rilevare le riviste di potenziale interesse scientifico. L'indagine ha portato alla definizione di 4 differenti graduatorie, distinte per i SSD SECS-S/01-02, SECS-S/03, SECS-S/04 e SECS-S/05, in cui le riviste sono sta-

te classificate in 3 classi di merito. Questa indagine fornisce elementi significativi che testimoniano la spiccata eterogeneità e varietà degli ambiti di pubblicazione degli studiosi afferenti alle Scienze Statistiche. Nel 2012 è stata definita una proposta di sintesi in cui vengono formulate due classifiche: quella effettuata con criterio SIS, ovvero ottenuta integrando nella classifica dell'indagine sui soci SIS, la classifica della Commissione Frosini, e quella effettuata con criterio ISI-WoS ovvero ottenuta suddividendo in 3 classi (A, B, C) le riviste della classifica SIS in base all'IF della Subject Category di riferimento (A se la rivista presenta IF superiore alla mediana, B se presenta IF inferiore alla mediana e C se non presenta IF).

4. Criticità della valutazione

Sulla base delle considerazioni appena riportate, relativamente alle differenti classifiche prodotte in sede di esercizio della VQR e in ambito associativo, è immediato riscontrare la difficoltà di definire liste e criteri di valutazione validi in senso generale in un contesto scientifico, come è quello riguardante l'Area 13, marcatamente eterogeneo e multidisciplinare. D'altra parte la produzione scientifica degli studiosi afferenti alle discipline statistiche può talvolta non trovare la giusta collocazione e la corretta classe di merito in un mero elenco di riviste scientifiche, che, per quanto completo ed esaustivo possa essere, determina comunque una limitazione ed uno snaturamento della Statistica, quale disciplina trasversale che trova applicazione in numerosi ambiti di ricerca. Quindi, l'evidente criticità risiede proprio nella procedura di catalogazione e classificazione delle riviste. Quali riviste considerare? Quali indicatori della qualità usare? Sulla base delle risposte a queste domande, che possono essere tutte valide e ragionevoli, possono scaturire differenti elenchi e classifiche. Inoltre, se si considera che, secondo quanto stabilito dall'ANVUR, l'analisi basata sui criteri di qualità deve convergere ad un giudizio sintetico sul prodotto della ricerca articolato su quattro livelli (Eccellente; Buono; Accettabile; Limitato) e che tali livelli fondamentalmente coincidono con la ripartizione dell'elenco delle riviste in quartili, è chiaro che tutto dipende da quante riviste ricadono o non ricadono nell'elenco stesso.

Infine, è importante evidenziare che il processo di valutazione proposto/imposto dall'ANVUR ha determinato negli studiosi un meccanismo automatico e cautelativo che spinge gli stessi studiosi a

pubblicare i propri prodotti della ricerca basando la scelta non più sulla rivista più pertinente con i suoi interessi di ricerca, bensì sulla rivista che è più probabile che venga riconosciuta nel settore e censita. Si aggiunga a latere che il principio della libertà della scienza è sancito dalla Costituzione italiana (art.33 comma 1) e, come tale, deve essere sempre garantito e mai limitato.

5. Verso nuovi criteri di valutazione della qualità della ricerca

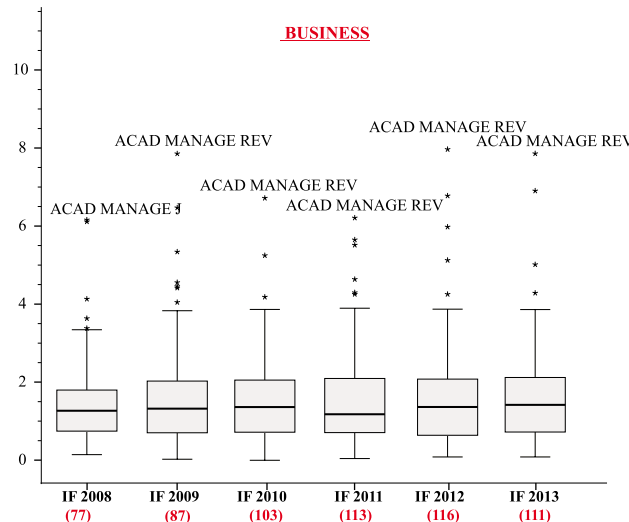
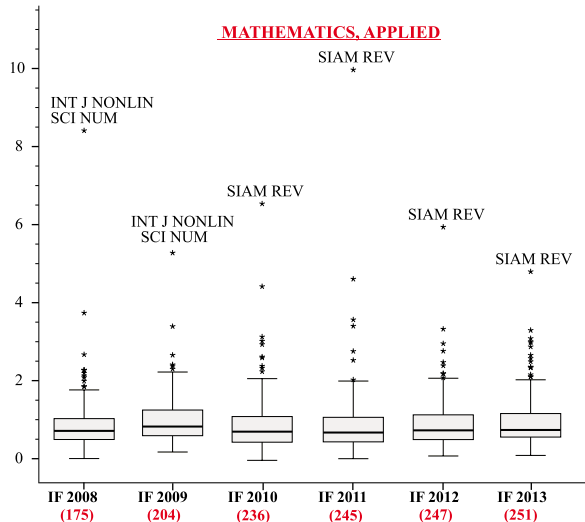
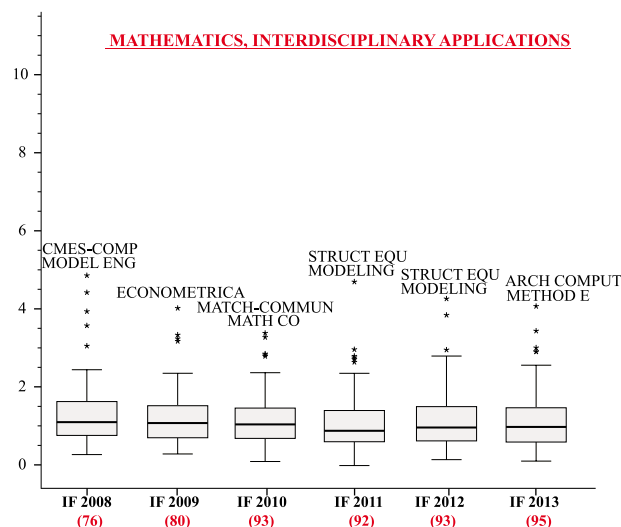
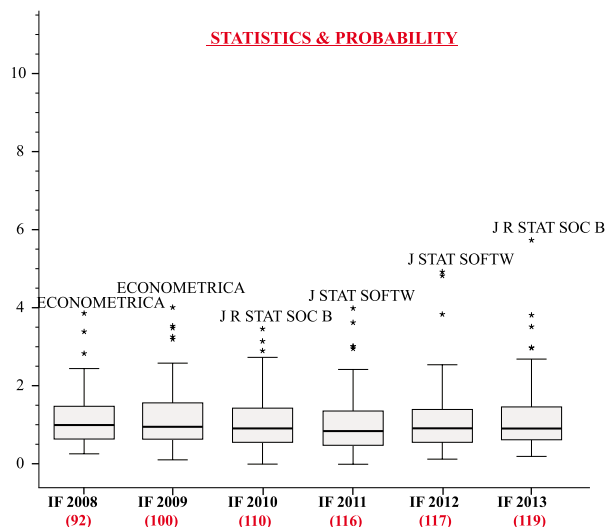
La definizione di nuovi e più generali criteri di valutazione della qualità della ricerca è sicuramente un'impresa ardua ed ambiziosa. Tuttavia, la consapevolezza che qualunque lista più o meno numerosa di riviste scientifiche mancherà pur sempre della caratteristica di esaustività, porta alla chiara affermazione del principio che occorre predisporre metodologie di valutazione di un articolo, che siano applicabili autonomamente e senza il ricorso alla consultazione di elenchi preventivamente definiti. Il ricorso a criteri di valutazione generali, che garantiscano la libertà della ricerca e rispettino gli interessi di ricerca di ciascuno studioso, non produrranno ancora situazioni in cui prodotti pubblicati su una rivista riconosciuta dall'ISI vengano paradossalmente collocati nell'ambito di prodotti limitati, a causa dell'irragionevole mancanza di quella rivista nella liste delle "elette". Il principio che dovrebbe essere largamente condiviso è che il riconoscimento dell'ISI è, di per sé garanzia di rilevanza internazionale e di ampia diffusione. La procedura di valutazione proposta si sviluppa attraverso le seguenti fasi:

1. individuare una Subject Category di riferimento per il macro-settore (ad esempio, facendo riferimento a ISI-WoS, per Statistica potrebbe essere "Statistics and Probability"),
2. considerare la distribuzione dell'IF delle riviste appartenenti alla suddetta Subject Category,
3. definire classi o *range* di qualità (ad esempio, ricorrendo ai quartili della distribuzione dell'IF della Subject Category), che siano validi per la valutazione di qualunque articolo pubblicato da uno studioso appartenente a quel macro-settore, su qualunque rivista.

Questa procedura può essere ragionevolmente applicata anche per periodi di durata pluriennale,

considerata la stabilità nel tempo della distribuzione dell'IF per una determinata Subject Category. Nella Figura 1 sono rappresentati i box plots (Posa, De Iaco, Palma, 2008) delle distribuzioni degli IF delle riviste appartenenti, rispettivamente, alle Subject Categories "STATISTICS & PROBABILITY", "MATHEMATICS, APPLIED", "OPERATIONS RESEARCH & MANAGEMENT SCIENCE", "MATHEMATICS, INTERDISCI-

PLINARY APPLICATIONS", "BUSINESS" e "ECONOMICS" per gli anni 2008-2013: è immediato riscontrare la stabilità dell'IF mediano nel sessennio considerato. In particolare, oltre alla mediana, anche il primo e il terzo quartile risultano stabili nel tempo. Tuttavia, le distribuzioni osservate sono caratterizzate da alcuni valori anomali sulla coda destra.



(segue)

(continua)

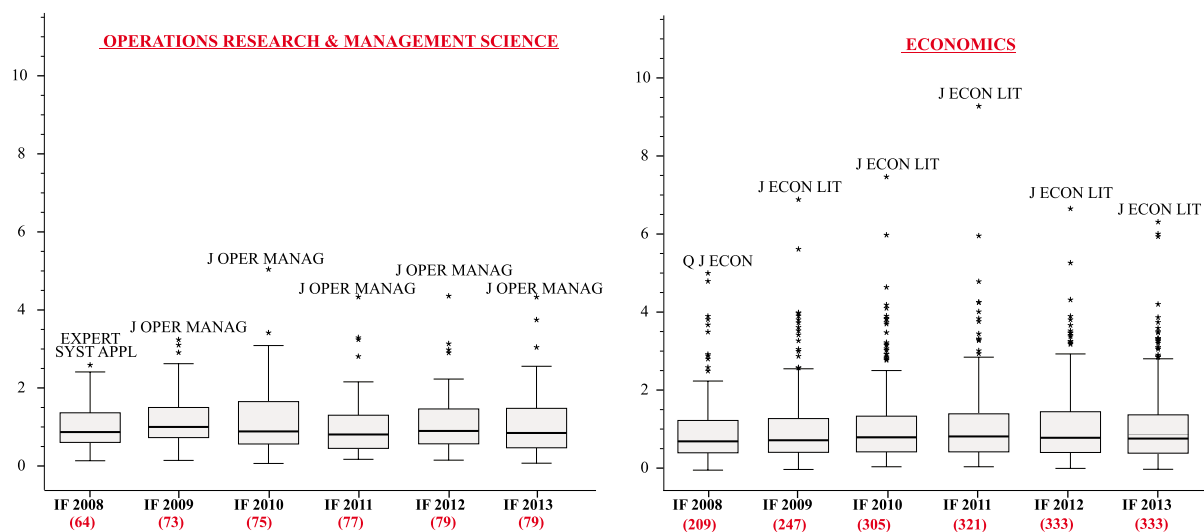


Figura 1: box plots delle distribuzioni degli IF delle riviste appartenenti alle Subject Categories considerate (in parentesi il numero di riviste per anno - elaborazioni su dati ISI Web of Knowledge).

Riferimenti bibliografici

Cicchitelli, G., Herzel, A., Montanari, G.E., 1997, *Il campionamento statistico*, Il Mulino, Bologna.

Cochran, W.G., 1977, *Sampling techniques*, J. Wiley & Sons, New York.

Frosini, B.V., 2008, *Valutazione della ricerca e valutazione delle riviste scientifiche in ambito statistico*, Statistica & Società, VI Speciale, 39-47.

IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY, IBM Corp. IBM Corp., 2011.

Posa, D., De Iaco, S., Palma, M., 2008, *Fondamenti di Statistica descrittiva*, Seconda Edizione, Giapichelli, Torino.

Sitografia

- <http://admin-apps.webofknowledge.com>
- http://www.anvur.org/rapporto/files/Area13/VQR2004-2010_Area13_RapportoFinale.pdf
- http://anvur-miur.cineca.it/sites/anvurmiur/files/imputation_and_ranking_of_journals_04_9_2012.pdf
- <http://new.sis-statistica.org/sis/valutazione-della-ricerca-nelle-scienze-statistiche/>
- <http://www.roars.it/online/>

Valutazione della Ricerca e Classifiche delle Riviste Scientifiche nelle Scienze Statistiche: l'esperienza SIS

Maurizio Carpita

Dipartimento di Economia e Management, Università degli Studi di Brescia
Consigliere del Direttivo della Società Italiana di Statistica (2010-2014)

maurizio.carpita@unibs.it

Sommario

Dopo aver ricordato il modello multidimensionale proposto dalla Società Italiana di Statistica (SIS) per la valutazione della ricerca individuale nell'ambito delle Scienze Statistiche, si descrivono in sintesi gli strumenti bibliometrici utilizzati per valutare la collocazione editoriale dei prodotti della ricerca. È quindi descritta la procedura adottata dalla SIS per la predisposizione delle classifiche di riviste scientifiche che hanno pubblicato articoli degli statistici italiani tra il 2004 e il 2010, compiuta nel 2012 in occasione dell'ultimo esercizio di valutazione nazionale della ricerca (VQR) promosso dall'ANVUR.

1. L'esperienza della Società Italiana di Statistica nella valutazione della ricerca

Ogni ricercatore sa bene quanto è complessa la sua attività, quanto numerosi e difficilmente quantificabili siano gli elementi che la compongono. Questa complessità è ben sintetizzata nell'introduzione di un interessante articolo pubblicato, ormai qualche anno fa, sulla prestigiosa rivista *Nature*:

"No scientist's career can be summarized by a number. He or she spends countless hours troubleshooting experiments, guiding students and postdocs, writing or reviewing grants and papers, teaching, preparing for and organizing meetings, participating in collaborations, advising colleagues, serving on editorial boards and more – none of which is easily quantified."
(Abbott et al., 2010, p. 860).

La definizione di criteri condivisi di valutazione della ricerca anche in aree relativamente limitate non è semplice: nel caso delle scienze statistiche sono presenti approcci di valutazione piuttosto diversi, che distinguono non solo i "metodologi" dagli "applicati" ma anche – vista trasversalità della disciplina e pur rimanendo nel solo ambito dell'Area 13 – ambiti scientifici eterogenei. Questa situazione si è ancora di più accentuata nel tempo, grazie allo sviluppo di collaborazioni tra discipline diverse, ma con obiettivi comuni. Malgrado tali difficoltà, si è resa sempre più urgente l'esigenza di

"misurare" l'attività di ricerca, i cui risultati sono utilizzati in particolare al fine di ripartire i finanziamenti nazionali e internazionali necessari per la sua realizzazione (Si veda ad esempio la piattaforma americana STAR METRICS: www.starmetrics.nih.gov; Largent e Lane, 2012). In rete si sono sviluppate iniziative utili per promuovere la ricerca e sempre più ricercatori riescono a misurare ma soprattutto a dare visibilità alla loro attività grazie a iniziative come Google Scholar (scholar.google.it), Research Gate (www.researchgate.net), ResearcherID (www.researcherid.com) di Thomson Reuters, ORCID (www.orcid.org) di Taylor & Francis, Elsevier, Wiley e Kudos (www.growkudos.com). Nell'ultimo decennio il Direttivo e diversi soci della Società Italiana di Statistica (SIS) si sono impegnati e confrontati sul tema della valutazione della ricerca in ambito statistico: tutta la documentazione prodotta finora alla quale fa riferimento questo contributo è disponibile sul sito www.sis-statistica.it, nella sezione *Valutazione della ricerca nelle scienze statistiche*.

Il modello generale di valutazione della ricerca proposto dalla SIS (Nicolini et al. 2010), che è in linea con l'impostazione del ricordato articolo di *Nature* ed è condiviso anche da altri ricercatori dell'Area 13 (Ferrini e Tucci, 2011), riconosce la natura multidimensionale dell'attività ricerca: come mostra la Figura 1, una dimensione considera le attività organizzative (distinte da un lato in attività di partecipazione a gruppi di ricerca e convegni e dall'altro in attività editoriali e di aggiornamento) e una dimensione riguarda le attività di pubblicazione.

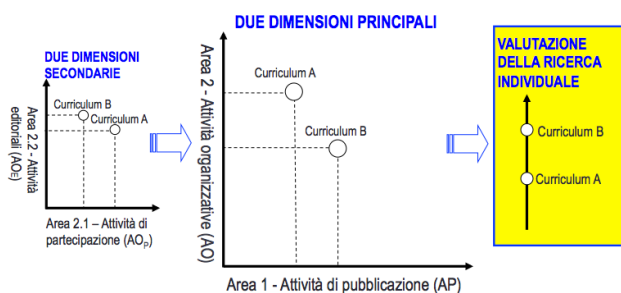


Figura 1: Il modello SIS di valutazione della ricerca individuale in Statistica (Nicolini et al. 2010)

Il sistema di valutazione proposto prevede il calcolo di 20 indicatori, riferiti a ogni dimensione (9 per l'attività di pubblicazione e 11 per le attività organizzative) e sotto-dimensione (5 per le attività di partecipazione a gruppi di ricerca e convegni e 6 per le attività editoriali e di aggiornamento). Per rendere questo modello operativo, in particolare per sintetizzare i 20 indicatori, nel 2010 è stata condotta un'indagine tra i soci SIS allo scopo di ricavare i pesi di importanza dei diversi aspetti considerati. In questo modo è stato possibile predisporre un prototipo di sistema di valutazione informatizzato, che permette di ricavare una scheda sintetica per descrivere l'attività di ricerca svolta in un certo periodo da un singolo ricercatore (Fig. 2).

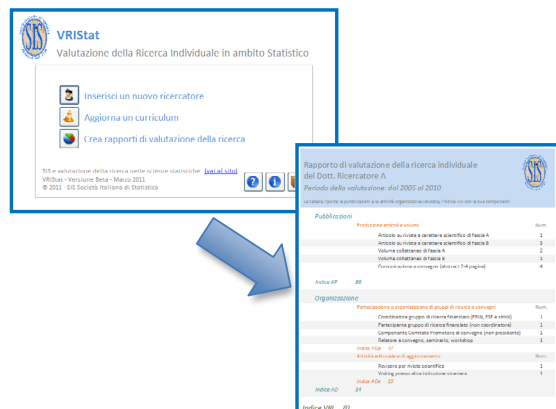


Figura 2: Schermata iniziale ed esempio di curriculum del programma VRISat-SIS (Carpita 2011)

All'interno di questa cornice, uno degli elementi qualificanti e attualmente più considerati per la valutazione della ricerca individuale è la classificazione dei "prodotti" della ricerca, in particolare articoli in riviste e volumi collettanei, monografie e atti di convegni a diffusione nazionale e internazionale (d'Ovidio 2011). Questa esigenza è oggi molto sentita in ambito accademico, a seguito della diffusione dei risultati dell'esercizio di valutazione delle strutture universitarie denominato VQR 2004-

2010, promosso dall'ANVUR e concluso a luglio 2013. In quell'occasione, Gruppi di Esperti della Valutazione (GEV) delle varie aree hanno classificato migliaia di prodotti della ricerca, attribuendoli a 4 classi di merito, utilizzate successivamente per la valutazione delle strutture universitarie di interesse. Non solo nelle aree cosiddette bibliometriche ma anche nell'area 13, una parte consistente dei prodotti scientifici considerati non è stata valutata tramite la classica peer-review (revisione dei ricercatori appartenenti alla stessa disciplina), ma considerando la sua collocazione editoriale: in altri termini, il livello scientifico dell'articolo è stato valutato (indirettamente) utilizzando una misura dell'impatto (diffusione) della rivista scientifica che lo ha pubblicato. Se questa operazione ha offerto da un lato indubbi vantaggi in termini di riduzione dei tempi e dei costi della valutazione, dall'altro ha modificato (non necessariamente in termini positivi) i comportamenti dei ricercatori e ha alimentato un acceso dibattito (www.lavoce.info/valutazione-continua/ e www.roars.it/online/category/vqr/).

Nel prossimo paragrafo si ricordano opportunità e rischi che derivano dall'utilizzo acritico della bibliometria per classificare riviste scientifiche ai fini della valutazione della ricerca. Nel paragrafo 3 si presenta la proposta dalla SIS che ha portato a metà del 2012, proprio in occasione dell'esercizio VQR 2004-2010, alla diffusione delle liste con la classificazione di riviste scientifiche di interesse per le scienze statistiche.

2. La bibliometria delle riviste scientifiche

È utile per i ricercatori conoscere la diffusione e la reputazione delle riviste scientifiche che potrebbero pubblicare i risultati dei propri studi. Queste informazioni servono anche ai finanziatori e a tutti gli interessati all'impatto dei risultati della ricerca. Chi partecipa ai progetti europei sa bene che è necessario dedicare molto tempo alla compilazione di dettagliati *time-sheet*, documenti nei quali si dichiara con dettaglio orario il tempo impiegato nell'attività di ricerca. Molto più facilmente verificabile ed efficiente è porre invece l'attenzione sull'output della ricerca, rendicontando la "produzione scientifica", ovvero il numero di *working paper*, *proceeding* e articoli in riviste e libri scientifici attesi ed effettivi alla fine delle diverse fasi del progetto ("deliverables of work packages").

Classifiche di riviste (ma anche di *working paper* e *proceeding*) e di editori di libri scientifici possono quindi aiutare a individuare il livello di diffusione (impatto) potenziale dei prodotti scientifici dei progetti di ricerca.

La bibliometria è la disciplina che si occupa di predisporre gli indicatori utili per classificare i prodotti editoriali della ricerca scientifica. Il *Journal Impact Factor* (JIF) di Thomson Reuters (calcolato per le sole riviste contenute nel database citazionale Web of Science - WoS) e il *Source Normalized Impact per Paper* (SNIP) di Elsevier (calcolato per le sole riviste indicizzate nel database citazionale Scopus) sono certamente tra gli indicatori bibliometrici più noti e considerati dai ricercatori di molte aree scientifiche per la scelta della rivista alla quale sottomettere i loro articoli.

È bene tenere presente che, per poter essere inserite nelle banche dati WoS e Scopus, le riviste devono superare un processo di selezione piuttosto rigido, che verifica il soddisfacimento di alcuni requisiti minimi di scientificità (stabilità della frequenza di pubblicazione, presenza di un comitato scientifico e di un livello di diffusione internazionale, adozione di un adeguato processo di selezione *peer-review*). Gli indicatori bibliometrici sono quindi calcolati considerando il numero di citazioni ricevute in media dagli articoli pubblicati in un certo periodo dalle riviste nella banca dati di appartenenza. JIF e SNIP non sono però gli unici indicatori utilizzabili per misurare l'impatto delle riviste scientifiche (Bollen et al 2009 ne considerano ben 39) e la proliferazione di metriche diverse crea disorientamento tra i ricercatori (Van Noorden, 2010). Inoltre, JIF, SNIP e gli altri indicatori bibliometrici presentano vantaggi ma anche numerosi limiti (Adams, 2014; Baccini, 2010; Frosini, 2007 e 2011; Setti, 2013). Prima di tutto è necessario evidenziare che gli indicatori bibliometrici misurano la diffusione media degli articoli pubblicati da una rivista, non la qualità scientifica intrinseca dei suoi singoli articoli: generalmente molti di questi non ricevono nemmeno una citazione e solo pochissimi ricevono molte citazioni. Come è ben noto agli statistici, in situazioni di elevatissima asimmetria positiva della distribuzione delle citazioni, l'impiego della media come misura di sintesi è sconsigliato.

Un ulteriore elemento di criticità riguarda la difficoltà di confrontare i valori di un indicatore bibliometrico delle riviste dedicate a diverse aree scientifiche, le cosiddette *subject categories* in WoS: per quanto riguarda la categoria di riferimento per le riviste delle scienze statistiche (*Statistics and Probability* in WoS), alcuni studi hanno evidenzia-

to che i relativi indicatori bibliometrici sono generalmente più bassi di quelli delle riviste di altre categorie (Van Nierop, 2009) e che la considerazione di una finestra temporale più ampia di quella usualmente adottata per calcolare tali indicatori permette di valutarne meglio l'effettivo impatto (Van Nierop, 2010). La già ricordata trasversalità delle ricerche basate su metodi e modelli statistici limita quindi la possibilità di comparare riviste che pubblicano articoli di statistici che svolgono ricerche in ambiti diversi: non è possibile confrontare direttamente il JIF di una rivista di statistica con il JIF di una rivista di economia, ingegneria o medicina. La multidisciplinarietà di molte riviste rende tali confronti ancora più problematici (Leydesdorff e Rafols, 2009).

Un altro noto problema degli indicatori bibliometrici riguarda la loro "manipolazione": in questi ultimi anni sono emersi numerosi casi di citazioni improprie (Van Noorden, 2013) e nessuna rivista, nemmeno la più blasonata a livello mondiale, può essere certa di pubblicare articoli di qualità scientifica indiscutibile (blogs.nature.com/news/2013/06/new-record-66-journals-banned-for-boosting-impact-factor-with-self-citations.html).

I casi di articoli apparsi su riviste di grande prestigio internazionale e poi ritirati per evidenti frodi sono diversi: di recente hanno fatto scalpore due articoli pubblicati su *Nature* e dedicati a STAP, un metodo per trasformare cellule umane in staminali (*Nature* 2014c). Si ricordano anche il caso delle due riviste bosniache presenti in Web of Science (Sipka, 2012) e quello più recente del *Journal of Vibration and Control*, che ha deciso di ritirare ben 60 articoli dopo aver scoperto che un ricercatore ne aveva influenzato il processo di *peer-review* (Fountain, 2014). Questi fatti, che rappresentano una quota marginale delle riviste e degli articoli pubblicati, dimostrano comunque la presenza di un attento controllo (purtroppo *ex-post*) della qualità degli articoli pubblicati.

Un altro elemento di criticità associato all'impiego degli indicatori bibliometrici riguarda la loro stabilità nel tempo (Ogden e Bartley, 2008; Chen et al., 2014). Da alcuni anni il Centre for Science and Technology Studies (CWTS) dell'Università di Leiden calcola (utilizzando la tecnica bootstrap) i *journal indicators stability intervals* (www.journalindicators.com) per valutare l'attendibilità degli indicatori bibliometrici ricavabili da Scopus. La Figura 3.a mostra le prime 10 riviste, delle 163 considerate di area statistica, ordinate rispetto all'indice SNIP del 2013 con i relativi intervalli di stabilità.

Title	P	SNIP	Stability interval
1 <i>Statistics Surveys</i>	14	6.99	
2 <i>Journal of the Royal Statistical Society, Series B: Statistical Method...</i>	84	4.87	
3 <i>Annals of Mathematics</i>	251	4.45	
4 <i>Journal of Statistical Software</i>	224	4.37	
5 <i>Annals of Statistics</i>	298	3.12	
6 <i>Journal of Machine Learning Research</i>	739	2.66	
7 <i>Journal of Mathematical Imaging and Vision</i>	166	2.54	
8 <i>Journal of Business and Economic Statistics</i>	135	2.50	
9 <i>Technometrics</i>	99	2.20	
10 <i>International Statistical Review</i>	60	2.13	

Statistics Surveys

Source type: Journal
 Publisher: Institute of Mathematical Statistics
 Areas: Statistics, Probability and Uncertainty, Statistics and Probability

Print ISSN: 1935-7516
 Electronic ISSN:

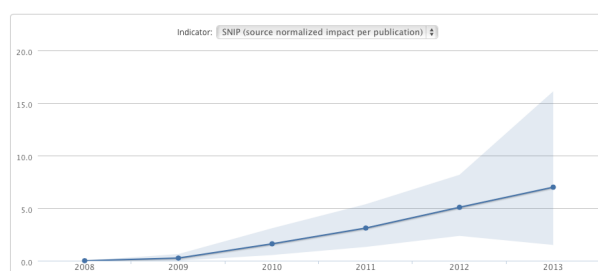


Figura 3: a) Prime 10 riviste della sotto-area Statistics and Probability secondo l'indicatore SNIP 2013 e relativi stability intervals; b) SNIP 2008-13 della rivista Statistics Surveys (www.journalindicators.com/indicators/)

È evidente che la rivista di maggiore impatto, *Statistics Surveys* (SNIP 6,99), presenta anche un intervallo molto ampio (da 1,5 a 16,1), mentre il *Journal of the Royal Statistical Society - Series B* (SNIP 4,87) mostra un intervallo molto più stretto (da 3 a 7,2), che comunque si sovrappone a quelli delle altre otto riviste considerate. Nella Figura 3.b è rappresentato l'andamento dell'indice SNIP di *Statistics Surveys* dal 2008 al 2013: si vede chiaramente l'evidente incremento dell'ampiezza del suo intervallo di stabilità. Il comportamento anomalo dell'indicatore SNIP di *Statistics Surveys* si può facilmente spiegare considerando che si tratta di una rivista *open-access* fondata nel 2007 da *American Statistical Association*, *Bernoulli Society*, *Institute of Mathematical Statistics* e *Statistical Society of Canada*, che pubblica articoli di rassegna dei metodi statistici: è quindi evidente che alcuni articoli siano molto citati. Questo esempio evidenzia chiaramente i rischi di utilizzare ordinamenti basati sull'uso puntuale degli indicatori bibliometrici.

Un ultimo elemento che negli ultimi anni ha modificato sostanzialmente le modalità di diffusione dei risultati della ricerca ma che nei prossimi anni diventerà probabilmente ancora più importante è la diffusione delle riviste presenti su internet ad accesso libero (*open-access*), modalità sempre

più incentivata anche a livello europeo (European Commission, 2013). Considerando l'esempio più famoso, la rivista *open-access* (a pagamento per gli autori) *PLoS ONE*, in soli sette anni e mezzo ha pubblicato più di 100 mila articoli (31.507 solo nel 2013) (blogs.plos.org/everyone/2014/06/23/plos-one-publishes-100000th-article/), accettando in media circa il 69% delle sottomissioni. Alcuni ritengono che in un futuro ormai prossimo il metodo della *peer-review* sarà sostituito da metodi di revisione pubblica in rete, con gli strumenti di *Altmetrics* (Alternative metrics; altmetrics.org/manifesto/), che offrono la possibilità di valutare l'impatto di una pubblicazione scientifica utilizzando le informazioni presenti in rete:

"The web enables new ways to measure scholarly productivity and impact, making available tools and data that can reveal patterns of intellectual activity and impact that were previously invisible: mentions, acknowledgments, endorsements, downloads, recommendations, blog posts, tweets." (Cronin e Sugimoto, 2014)

Si stima che sono oggi accessibili tramite web almeno 114 milioni di articoli scientifici scritti in inglese, quasi 100 milioni dei quali indicizzati da Google Scholar e meno di 50 milioni da WoS (Khabisa e Giles 2014). È quindi aumentata notevolmente in questi ultimi tempi l'attenzione verso queste nuove fonti informative per la bibliometria (De Bellis, 2009, capitolo 8) e sempre più ricercatori sostengono la maggiore rappresentatività degli indicatori d'impatto della ricerca ottenuti da Google Scholar rispetto a WoS e Scopus (Delgado e Repiso, 2013); alcuni si sono già spinti a utilizzare per il calcolo degli indicatori bibliometrici i dati provenienti dai social network e da Wikipedia (Priem e Hemminger, 2010; Evans e Krauthammer, 2011; Bik e Goldstein, 2013; Thelwall et al., 2013). È bene tenere presente che queste moderne fonti informative presentano ancora grandi problemi di attendibilità e qualità dei dati, tali da renderne sconsigliabile l'uso esclusivo (Aguillo, 2012; Delgado López-Còzar et al., 2014). La lista aggiornata periodicamente da Jeffrey Beall dedicata ai cosiddetti "predatory publishers" (scholarlyoa.com/2014/01/02/list-of-predatory-publishers-2014/) è aumentata nel corso degli anni e nel 2014 conta 477 editori *open-access* con comportamenti editoriali considerati non conformi agli standard scientifici.

Tutte queste difficoltà di calcolo, interpretazione e quindi utilizzo degli indicatori bibliometrici hanno portato molte istituzioni a sconsigliarne l'uso auto-

matico: dopo alcuni esperimenti pilota, gli esercizi di valutazione della ricerca in corso nel Regno Unito (REF: *Research Excellence Framework*) e in Australia (ERA: *Excellence in Research Assessment*) hanno deciso di utilizzare le citazioni bibliometriche solo come strumento utile alla cosiddetta *informed peer-review* (www.ref.ac.uk/background/bibliometrics/). Inoltre, contro l'uso acritico della bibliometria, sono state pubblicate nel 2012 la *San Francisco Declaration on Research Assessment* (am.ascb.org/dora/; researchanalytics.thomsonreuters.com/statement_re_sfdra/; www.elsevier.com/connect/san-francisco-declaration-on-research-assessment-dora-elseviers-view) e nel 2013 la *Statement on Appropriate Use of Bibliometric Indicators* (www.ieee.org/publications_standards/publications/rights/bibliometrics_statement.html). Pur con queste difficoltà, è indubbio che la bibliometria offre utili indicazioni per misurare l'impatto potenziale delle riviste scientifiche (il solo fatto di essere presenti in WoS o Scopus indica il soddisfacimento dei requisiti minimi di scientificità) e può quindi essere uno degli elementi (ma non l'unico e nemmeno il prevalente) per valutare la qualità di un articolo scientifico.

3. Le classifiche SIS delle riviste scientifiche per le Scienze Statistiche

Come ricordato alla fine del primo paragrafo, in occasione dell'esercizio VQR 2004-2010 anche la SIS – come altre società scientifiche – è stata invitata a dare il proprio parere riguardo alla classificazione delle riviste scientifiche. A tale proposito, è bene precisare che le proposte delle società scientifiche non erano vincolanti per i GEV, che hanno deciso in piena autonomia i criteri ritenuti più adeguati per la valutazione della ricerca nelle loro aree di competenza.

Per favorire l'attività di valutazione, l'ANVUR ha messo a disposizione della SIS gli elenchi delle riviste inserite nella banca dati CINECA dai ricercatori dei settori scientifici delle scienze statistiche di Area 13 (SECS-S01/05), con l'ulteriore informazione del numero di articoli da questi pubblicati dal 2004 al 2010. La Tab. 1 mostra il numero di statistici censiti dal MIUR, di articoli e di riviste presenti nella banca dati CINECA nei sette anni considerati: risulta evidente il gran numero di riviste che nel periodo considerato hanno pubblicato articoli di statistici italiani.

Numero di...	SECS-S01	SECS-S02	SECS-S03	SECS-S04	SECS-S05
Statistici MIUR	444	26	148	75	69
Articoli CINECA 04-10	2.785	362	836	636	430
Riviste CINECA 04-10	745	170	288	211	221

Tabella 1: Numero di statistici, articoli e riviste presenti nella banca dati CINECA 2004-2010 (area scienze statistiche)

A conferma delle differenti scelte editoriali dei ricercatori appartenenti alla stessa area disciplinare, la Tab. 2 evidenzia che solo una parte degli articoli da loro pubblicati è presente in WoS (dal 70% per SECS-S01 al 56% per SECS-S03) o appartiene alle *subject category* di riferimento previste dalla SIS (dal 26% per SECS-S01 al 40% per SECS-S03).

Riviste...	SECS-S/01	SECS-S/02*	SECS-S/03	SECS-S/04	SECS-S/05
in SC di riferimento	26%	7%	40%	30%	37%
in SC altre	44%	72%	16%	8%	26%
non in WoS	30%	21%	44%	62%	37%
Totale	100%	100%	100%	100%	100%

* Oltre a quelle previste per SECS-S/01

Tabella 2: Composizione % degli articoli pubblicati nelle riviste presenti nella banca dati CINECA 2004-2010 (area scienze statistiche)

La Figura 4 rappresenta invece, con riferimento al solo settore scientifico SECS-S01 (ma risultati analoghi si ottengono per gli altri settori), il fenomeno della *coda lunga* (Anderson, 2007) delle riviste che hanno pubblicato gli articoli degli statistici italiani SECS-S01 nel periodo 2004-2010: poche riviste con molti articoli e – di conseguenza – molte riviste con solo 1 o 2 articoli.

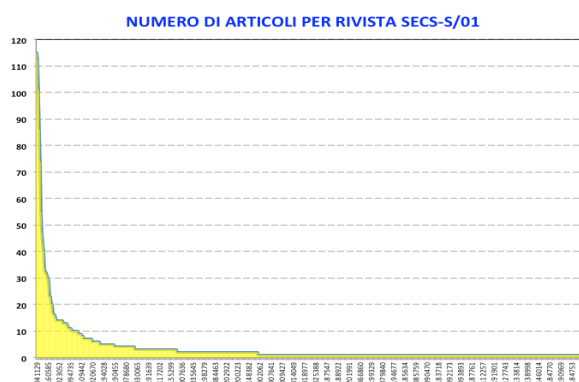


Figura 4: Distribuzione del numero di articoli per riviste SECS-S01 presenti nella banca dati CINECA 2004-2010 (area scienze statistiche)

Questi risultati confermano la trasversalità delle scienze statistiche, che trovano applicazioni in ambiti teorici e applicati anche molto diversi: limitarsi a considerare esclusivamente gli articoli pubblicati nelle riviste del solo ambito statistico rischia di ridurre sensibilmente il raggio d'azione della disci-

plina, con un indebolimento delle potenzialità di collaborazioni interdisciplinari dei suoi ricercatori. Considerati questi risultati preliminari, per classificare le riviste sono state combinate due diverse fonti informative, che hanno portato a due distinte classifiche: la classifica *bibliometrica* (ottenuta dall'impiego del JIF – normalizzato rispetto alla mediana della *subject category* di riferimento – e dello SNIP) e quella *reputazionale* (ottenuta considerando i risultati di alcune indagini sulla rilevanza delle riviste di statistica svolte tra i soci negli anni precedenti). La scelta di combinare le due diverse classifiche per lo stesso gruppo di riviste è stata adottata con l'obiettivo di attenuare almeno in parte le possibili distorsioni di entrambi i metodi. Poiché l'esercizio VQR svolto dall'ANVUR prevedeva di classificare le riviste rispetto a 4 classi di merito (A: eccellente, B: buona, C: accettabile, D: limitata), per costruire le classifiche bibliometrica e reputazionale è stato usato lo stesso criterio. Le due liste ottenute sono esemplificate per il settore disciplinare SECS-S01 nelle ultime due colonne a destra della Tab. 3. Combinando queste due liste è stata quindi ricavata la classifica finale VQR, riportata nella terza colonna della Tab. 3: questa classifica soddisfa anche il requisito (a dire il vero quanto meno curioso) che le riviste nelle 4 classi siano distribuite con le percentuali 20%-20%-10%-50% previste dal GEV dell'Area 13.

RIVISTA	ISSN	CLASSE VQR	CLASSE SIS	WoS SubCatRef	IN SCOPUS	CLASSE REPUTAZIONALE	CLASSE BIBLIOMETRICA
ANN STAT	0090-5364	A	Aa+	StatProb	SI	A	A
BIOSTATISTICS	1465-4644	A	Aa+	StatProb	SI	A	A
ECONOMETRICA	0012-9682	A	Aa+	StatProb	SI	A	A
J AM STAT ASSOC	0162-1459	A	Aa+	StatProb	SI	A	A
J R STAT SOC A STAT	0964-1098	A	Aa+	StatProb	SI	A	A
J R STAT SOC B	1369-7412	A	Aa+	StatProb	SI	A	A
REV ECON STAT	0034-6535	A	Aa+	SocSciMM	SI	A	A
STAT COMPUT	0960-3174	A	Aa+	StatProb	SI	A	A
STAT MED	0277-6715	A	Aa+	StatProb	SI	A	A
STAT SCI	0883-4237	A	Aa+	StatProb	SI	A	A
ANN APPL PROBAB	1050-5164	A	Aa	StatProb	SI	A	A
ANN PROBAB	0091-1798	A	Aa	StatProb	SI	A	A
BIOMETRICS	0006-341X	A	Aa	StatProb	SI	A	A
BIOMETRIKA	0006-3444	A	Aa	StatProb	SI	A	A
CHEMOMETR INTELL LAB	0169-7439	A	Aa	StatProb	SI	B	A
COMB PROBAB COMPUT	0963-5483	A	Aa	StatProb	SI	B	A
COMPUT STAT DATA AN	0167-9873	A	Aa	StatProb	SI	A	A
DATA MINING KNOWL & DISCOVERY	1384-5810	A	Aa	ComSciAI	SI	B	A
ECONOMET REV	0747-4938	A	Aa	StatProb	SI	A	A
ECONOMET THEOR	0266-4666	A	Aa	StatProb	SI	A	A
IEEE TRANSACTIONS ON RELIABILITY	0018-9529	A	Aa	EngEE	SI	B	A
J APPL ECONOM	0883-7252	A	Aa	SocSciMM	SI	A	A
J BUS ECON STAT	0735-0015	A	Aa	StatProb	SI	A	A
J COMPUT GRAPH STAT	1061-8600	A	Aa	StatProb	SI	A	A
J ECONOMETRICS	0304-4076	A	Aa	SocSciMM	SI	A	A
J EDUC BEHAV STAT	1076-9986	A	Aa	SocSciMM	SI	B	A
J QUAL TECHNOL	0022-4065	A	Aa	StatProb	SI	B	A

Tabella 3: Classifica delle riviste del settore SECS-S01 presenti nella banca dati CINECA 2004-2010 (area scienze statistiche)

Con l'indicatore SNIP 2010 si è infine ritenuto utile ottenere una classifica delle riviste in 9 classi di merito (senza vincoli sulle percentuali di ogni classe), riportata nella colonna 4 della Tab. 3 (per i dettagli della procedura si veda www.sis-statistica.it, sezione *Valutazione della ricerca nelle scienze statistiche*).

Come ricordato all'inizio di questo paragrafo, i GEV hanno considerato quanto proposto dalle

società scientifiche, ma hanno poi adottato autonomamente il criterio di valutazione ritenuto più idoneo: in particolare, il GEV dell'Area 13 non ha considerato (almeno esplicitamente) il criterio reputazionale oppure l'indicatore SNIP per la classificazione di riviste senza JIF, ritenendo invece di imputare degli JIF stimati con un modello di regressione che utilizza le citazioni da Google Scholar (con tutte le criticità ricordate nel secondo paragrafo).

Non è quindi sorprendente che le classifiche ottenute dalla SIS e dal GEV 13 in occasione dell'esercizio VQR 2004-2010 siano piuttosto diverse. La Tabella 4 mostra la prima parte dell'elenco ottenuto ordinando le riviste rispetto al numero di articoli pubblicati dagli statistici dell'Area 13 durante il periodo considerato (seconda colonna: Num): la prima rivista in classe A per la lista del GEV dell'Area 13 – *Scienze Economiche e Statistiche* (ultima colonna: VQR) è *Statistics in Medicine*, che con 30 articoli si colloca in 16° posizione.

Riviste ordinate per numero articoli pubblicati	ARTICOLI		IF	SNIP	SIS	SIS-VQR	VQR
	Num	Cum. %					
COMPUTATIONAL STATISTICS & DATA ANALYSIS	115	4	X	X	Aa	A	B
STATISTICAL METHODS & APPLICATIONS	112	8	X	X	Ab	A	C
STATISTICA & APPLICAZIONI	101	12			Bb	B	D
STATISTICA	99	15			Bb	B	D
STATISTICA APPLICATA	86	18			Bb	B	D
JOURNAL OF STATISTICAL PLANNING AND INFERENCE	74	21	X	X	Ab	A	B
QUADERNI DI STATISTICA	55	23			Bb	B	D
METRON	47	25			Bb	A	C
ENVIRONMETRICS	43	26	X	X	Bb	A	B
RIVISTA ITALIANA DI ECONOMIA, DEMOGRAFIA E STATISTICA	40	28			Db	D	D
COMMUNICATIONS IN STATISTICS. THEORY AND METHODS	40	29	X	X	Ab	A	D
JOURNAL OF APPLIED STATISTICS	34	30	X	X	Ab	A	D
INDUZIONI	32	32			Db	D	D
APPLIED STOCHASTIC MODELS IN BUSINESS AND INDUSTRY	32	33	X	X	Bb	A	C
STATISTICS & PROBABILITY LETTERS	31	34	X	X	Ab	A	D
STATISTICS IN MEDICINE	30	35	X	X	Aa+	A	A

Tabella 4: Classifica delle riviste del settore SECS-S01 presenti nella banca dati CINECA 2004-2010 (area scienze statistiche)

Secondo la classifica ottenuta applicando il doppio criterio bibliometrico-reputazionale alla VQR (penultima colonna: SIS-VQR), in questo elenco parziale sono presenti in classe A ben 10 riviste. Si evidenzia che la storica rivista *Metron* si trova per la SIS-VQR in classe A (data la sua rilevanza reputazionale) mentre per la VQR si trova in classe C (data l'assenza in WoS). Anche la ben nota *Statistics & Probability Letters* è in classe A per la SIS-VQR e addirittura in classe D per la VQR. In soli tre casi (*Rivista Italiana di Economia, Demografia e Statistica*, *Induzioni* e la già ricordata *Statistics in Medicine*) le classificazioni VQR e SIS-VQR offrono risultati concordanti.

Si noti infine che questo elenco, composto da sole 16 riviste, rappresenta ben il 35% del totale degli articoli pubblicati dagli statistici italiani tra il 2004 e il 2010: si può quindi ben capire quale può essere stato l'effetto della classificazione VQR sulla va-

lutazione della produzione scientifica nelle scienze statistiche.

4. Considerazioni conclusive

Dopo aver ricordato il modello multidimensionale di valutazione della ricerca proposto dalla SIS, in questo contributo sono state evidenziate le numerose difficoltà di utilizzo della bibliometria (soprattutto quando è impiegata per classificare articoli scientifici) e l'esperienza fatta in occasione dell'esercizio VQR 2004-2010 promosso dall'ANVUR. Come si è provato ad argomentare, le classifiche proposte dalla SIS sono risultate piuttosto diverse da quelle adottate dal GEV dell'Area 13: evidentemente questa situazione non ha favorito la possibilità di dare indicazioni chiare per orientare le scelte dei ricercatori. La classifica del GEV dell'Area 13, oltre ad aver avuto un impatto *ex-post* sulla valutazione della produzione di letteratura scientifica in ambito statistico del periodo 2004-2010, ha molto probabilmente già influenzato anche la produzione scientifica degli statistici italiani degli ultimi quattro anni e soprattutto la influenzerà in futuro (diversi atenei utilizzano i criteri della VQR, ad esempio per la ripartizione dei fondi di ricerca). Comunque, anche la classifica predisposta dalla SIS non è esente da limiti, dovuti in parte alla necessità di attenersi ai criteri VQR e in parte all'oggettiva difficoltà di misurare l'effettivo impatto delle riviste scientifiche per una disciplina così trasversale.

Al fine di rivedere e aggiornare le classifiche del 2012, il Direttivo SIS ha nominato una nuova Commissione con i seguenti obiettivi:

- aggiornare gli indicatori bibliometrici
- affrontare il problema delle *subject category*
- valutare la rilevanza di riviste non statistiche
- aggiornare l'approccio reputazionale

La nuova Commissione si dovrà anche chiedere se la direzione indicata ai giovani ricercatori delle scienze statistiche sia effettivamente quella corretta e utile perché abbiano un futuro nel mondo della ricerca internazionale.

Riferimenti bibliografici

Abbott A., Cyranoski D., Jones N., Maher B., Schiermeier Q. and Van Noorden R., 2010, *Metrics: Do metrics matter?*, Nature, 465, June, 860-862.

Adams J., 2014, *The citation game*, Nature, 510, 26 June, 470-471.

Albert B., 2013, *Impact factor distortions*, Science, 340, 17 May, 787.

Aguillo I.F., 2012, *Is Google Scholar useful for bibliometrics? A webometric analysis*, Scientometrics 91, 2, 343-351.

Anderson C., 2007, *La coda lunga. Da un mercato di massa a una massa di mercati*, Codice Edizioni, Torino.

Baccini A., 2010, *Valutare la ricerca scientifica*, Il Mulino, Bologna.

Bik H.M., Goldstein M.C., 2013, *An introduction to social media for scientists*, PLoS Biology, 11, 4, 1-8.

Bollen J., Van de Sompel H., Hagberg A., Chute R., 2009, *A principal component analysis of 39 scientific impact measures*, PLoS ONE, 4, 6, 1-11.

Campbell P., Grayson M., 2014, *Assessing Science*, Nature Outlook, 24 July, Issue 7510.

Carpita M., 2011, *VRISat: Measuring the italian statisticians' research activities*, Quaderni del Dipartimento di Scienze Statistiche "Paolo Fortunati", Serie Ricerche, Alma Mater Studiorum Università di Bologna, 1-6.

Chen K.M., Jen T.H., Wu M., 2014, *Estimating the accuracies of journal impact factor through bootstrap*, Journal of Informatics, 8, 181-196.

Cronin B., Sugimoto C.R. (a cura di), 2014, *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, MIT Press. (mitpress.mit.edu/books/beyond-bibliometrics)

D'Ovidio F.D., 2011, *La misura della produttività scientifica individuale e delle strutture di ricerca*, in Fabbris L. (a cura di), *Criteri e indicatori per misurare l'efficacia delle attività universitarie*, 12, 299-329, Cleup, Padova.

De Bellis N., 2009, *Bibliometrics and citation analysis – From the science citation index to cybermetrics*, Scarecrow Press Inc, UK.

Delgado E., Repiso R., 2013, *The impact of scientific journals of communication: comparing Google Scholar metrics*, Web of Science and Scopus, Comunicar, 41, XXI, 45-52.

Delgado López-Còzar E., Robinson-Garcia N., Torres-Solinas D., 2014, *The Google scholar experiment: How to index false papers and manipulate bibliometric indicators*, Journal of the Association for Information Science and Technology, 65, 3, 446-454.

European Commission, 2013, *Guidelines on Open*

Access to Scientific Publications and Research Data in Horizon 2020, ver. 1.1, December.

(ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf).

Evans P., Krauthammer M., 2011, *Exploring the Use of Social Media to Measure Journal Article Impact*, AMIA Annu Symp Proc, 374-381 (www.ncbi.nlm.nih.gov/pmc/articles/PMC3243242/#!po=12.5000).

Frosini B.V., 2007, *Valutazione della ricerca e valutazione delle riviste scientifiche in ambito statistico*, Statistica & Società, 6, 39-47.

Frosini B.V., 2011, *La valutazione della ricerca scientifica: scopi e criteri di valutazione*, in Frosini B.V. (a cura di), *La valutazione della ricerca e la valutazione della didattica*, Vita e Pensiero, Milano, 13-83.

Khabsa M., Giles C.L., 2014, *The number of scholarly documents on the public web*, PLoS ONE 9(5), e93949. doi:10.1371/journal.pone.0093949.

Nicolini G., Carpita M., Civardi M., Manisera M., Manzi M., Peracchi F., 2010, *Analisi dei criteri per la valutazione della ricerca individuale in ambito statistico, Relazione finale della Commissione SIS Riforma dei criteri condivisi per la valutazione della ricerca e revisione delle equivalenze dei titoli accademici europei ed internazionali da applicare alle "chiamate dirette"*, Società Italiana di Statistica, Roma.

Ferrini S., Tucci M.P., 2011, *Evaluating research activity: impact factor vs. research factor*. *Electronic Journal of Applied Statistical Analysis: Decision Support Systems and Services Evaluation*, 2, 1, 25-53 (siba-ese.unisalento.it/index.php/ejasa_dss/article/viewFile/3947/10610).

Fountain H., 2014, *Science journal pulls 60 papers in peer-review fraud*, The New York Times, Science, July 10.

Largent M.A., Lane J.I., 2012, *STAR METRICS and the Science of Science Policy*, Review of Policy Research, 29, 3, 431-438.

Leydesdorff L., Rafols I., 2009, *A global map of science based on the ISI subject categories*, Journal of the American Society for Information Science and Technology, 60, 2, 348-362.

Nature, 2014a, *Assessing Science. Lessons from Australia and New Zealand*, Outlook, 511, 24 July, 7510.

Nature, 2014b, *Metrics market*, Editorial, 510, 26 June.

Nature, 2014c, *STAP retracted*, Editorial, 511, 3 July.

Ogden T.L., Bartley D.L., 2008, *The ups and downs of journal impact factors*, Annals of Occupational Hygiene, 52, 73-82.

Priem J., Hemminger B.M., 2010, *Scientometrics 2.0: Toward a new metrics of scholarly impact on the social Web*, First Monday 15, 7 (firstmonday.org/ojs/index.php/fm/article/view/2874/2570#author).

Setti G., 2013, *Bibliometric indicators: Why do we need more than one?* IEEE Access, 1, 232-246.

Sipka P., 2012, *Legitimacy of citations in predatory publishing: The case of proliferation of papers by Serbian authors in two Bosnian WoS-indexed journals*, CEES Occasional Series, 12-2.

Thelwall M., Haustein S., Larivière V., Sugimoto C.R., 2013, *Do Altmetrics Work? Twitter and Ten Other Social Web Services*. PLoS ONE 8(5), e64841.

Van Nierop E., 2009, *Why do statistics journals have low impact factors?*, Statistica Neerlandica, 63, 1, 52-62.

Van Nierop E., 2010, *The introduction of the 5-year impact factor: does it benefit statistics journals?*, Statistica Neerlandica, 64, 1, 71-76.

Van Noorden R., 2010, *A profusion of measures*, Nature, 465, 17 June, 864-866.

Van Noorden R., 2013, *Brazilian citation scheme ousted. Thomson Reuters suspends journals from its rankings for 'citation stacking'*, Nature, 500, 29 August, 510-511.

Sulla misurazione del prestigio di una rivista scientifica

Francesco Bartolucci

Dipartimento di Economia
Università di Perugia

bart@stat.unipg.it

Sommario

Con riferimento ad un certo gruppo di riviste scientifiche di Statistica, si commenta l'andamento di alcuni indicatori di qualità e diffusione, quali l'Impact Factor e l'h-index, instaurando un confronto con i risultati di alcuni studi svolti nell'ambito della Società Italiana di Statistica per misurarne il prestigio. Viene anche suggerito di definire il prestigio in termini di difficoltà nel pubblicare in una certa rivista e viene suggerito un metodo per misurare tale difficoltà.

1. Introduzione

Il tema della classificazione delle riviste scientifiche ha attualmente un ruolo centrale sia a livello nazionale che internazionale. A livello nazionale queste classifiche vengono utilizzate in ambito di Abilitazione Scientifica Nazionale e sono state utilizzate per il recente esercizio di Valutazione della Qualità della Ricerca riferito al periodo 2004-2010. A livello internazionale, inoltre, esistono vari enti che periodicamente mettono a disposizione classifiche di riviste scientifiche, tra cui si segnalano Thomson Reuters, tramite Web of Science (WoS), ed Elsevier, tramite Scopus.

Si segnala che anche Google, tramite la sezione Metriche di Google Scholar (GS), mette ora a disposizione una classifica di riviste per singoli settori. Questi dati sono consultabili anche tramite Publish or Perish. Per una discussione approfondita si rinvia a Bartolucci, Dardanoni, Peracchi (2013).

Le classifiche di riviste scientifiche di cui sopra fanno uso di indicatori della qualità di una rivista che sono basati essenzialmente sul numero delle citazioni degli articoli pubblicati in un certo periodo. Di conseguenza questi indicatori vanno a misurare l'impatto della rivista sulla comunità scientifica. Tra questi il più noto è il cosiddetto Impact Factor (IF) che, per una certa rivista scientifica, corrisponde al numero medio di citazioni che in un dato anno

sono riferite agli articoli pubblicati nel biennio precedente in quella rivista.

L'IF viene utilizzato in particolare nell'ambito della classifica fornita da WoS insieme ad altri indicatori aventi una maggiore complessità di calcolo. Un altro indicatore molto noto è l'h-index (Hirsch, 2005), utilizzato ad esempio da GS (Metriche) con riferimento a un orizzonte temporale di 5 anni. Una rivista ha un valore di questo indice pari ad h se in questa rivista sono stati pubblicati h articoli con almeno h citazioni ciascuno, mentre tutti gli altri articoli hanno non più di h citazioni ciascuno.

Con riferimento alle materie statistiche, nella presente nota viene esaminata, limitatamente alle migliori riviste, la classifica basata sugli indicatori di cui sopra e su altri indicatori aventi una struttura simile all'IF.

In particolare, vengono commentati alcuni risultati riguardanti le riviste che sono ritenute di maggiore pregio, facendo notare che per queste riviste alcuni indicatori basati sul numero di citazioni non sono pienamente in linea con il prestigio percepito. Ciò permette anche di identificare in modo più preciso a cosa corrisponda il *prestigio* di una rivista utilizzando dei dati provenienti da studi condotti dalla Società Italiana di Statistica (SIS). Infine, viene ipotizzato un modello di valutazione basato su indicatori di difficoltà nel pubblicare in una certa rivista come misura collegata al livello reputazionale.

2. Le migliori riviste di Statistica secondo IF

Nella Tabella 2 viene riportata la lista delle migliori 30 riviste secondo IF al 2013 (fornita da WoS) per il settore "STATISTICS & PROBABILITY" con il valore dell'h-index a 5 anni e dei seguenti indici insieme alla rispettiva posizione in graduatoria:

- IF a 5 anni (IF5): variante dell'IF che tiene conto delle citazioni degli articoli pubblicati nei 5 anni precedenti;
- numero di citazioni complessivamente contegiate da WoS;
- numero di articoli pubblicati dalla rivista nel 2013;
- Article Influence Score (AIS): indice basato sulle stesse informazioni utilizzate da IF5, pesando le citazioni in modo diverso a seconda della provenienza.

Questo gruppo corrisponde sostanzialmente al quartile delle riviste migliori secondo l'IF, visto che la lista in questione include nel complesso 119 riviste per il 2013. Inoltre, con riferimento a queste riviste, i valori dell'indice di correlazione (anche nella versione di Spearman) tra coppie di indicatori sono riportati in Tabella 1.

	Indice di correlazione				
	IF5	Citazioni	Articoli	AIS	h-index
IF	0,808	0,299	-0,125	0,603	0,365
IF5		0,397	-0,103	0,759	0,592
Citazioni			0,403	0,576	0,742
Articoli				-0,185	0,281
AIS					0,768
	Indice di correlazione di Spearman				
	IF5	Citazioni	Articoli	AIS	h-index
IF	0,705	0,358	-0,092	0,350	0,324
IF5		0,509	-0,141	0,732	0,493
Citazioni			0,258	0,438	0,701
Articoli				-0,289	0,510
AIS					0,533

Tabella 1: Indici di correlazione tra le coppie di indicatori considerati in Tabella 2 per le prime 30 riviste del settore "STATISTICS & PROBABILITY" secondo IF

Nella Tabella 2 vengono anche evidenziati i dati riferiti a sei riviste che sono ritenute di maggiore prestigio; si tratta delle seguenti riviste: *Annals of Statistics*, *Biometrics*, *Biometrika*, *Econometrica*, *Journal of the American Statistical Association*, *Journal of the Royal Statistical Society - series B*. L'individuazione di queste riviste è basata, oltre che su considerazioni personali dell'autore della presente nota, sui risultati dei lavori di alcune

commissioni SIS e di indagini svolte dalla società medesima tramite questionari somministrati ai suoi soci.¹ A questo riguardo si ricorda, in particolare, che la Commissione Frosini² aveva individuato un insieme di nove riviste valutate come eccellenti; rispetto alle sei riviste di cui sopra questo gruppo non contiene *Biometrics*, mentre contiene in aggiunta *Annals of Probability*, *Annals of Applied Probability*, *Journal of Econometrics* e *Probability Theory and Related Fields*. Inoltre, dall'indagine SIS sulla percezione delle riviste nelle scienze statistiche³ è emerso che le cinque riviste considerate di maggiore prestigio per i settori scientifico-disciplinari SECS-S01 e SECS-S02 sono *Annals of Statistics*, *Biometrics*, *Biometrika*, *Journal of the American Statistical Association*, *Journal of the Royal Statistical Society - series B*, mentre per il settore SEC-S03 sono *Biometrika*, *Econometrica*, *Journal of Econometrics*, *Journal of the American Statistical Association*, *Journal of the Royal Statistical Society - series B*.

Dall'esame della Tabella 2 emerge che, in termini di IF, alcune delle riviste di maggiore prestigio sono superate da altre riviste che sono sicuramente meno conosciute anche perché di sola recente istituzione. Il caso emblematico si presenta per *Biometrics* e *Biometrika* che, nonostante la loro indubbia reputazione, si collocano rispettivamente al 26° e 27° posto in termini di IF. La situazione è solo marginalmente diversa se si considera l'IF a 5 anni e l'AIS; in questo caso l'ultima rivista tra le sei riviste di maggiore pregio è *Biometrics* che si colloca al 24° posto (secondo IF a 5 anni) e al 20° posto (secondo AIS). Tuttavia, è evidente una concordanza molto maggiore con la classifica basata sull'h-index a 5 anni fornito da GS. Secondo questo indice, infatti, le sei riviste di maggiore pregio si collocano dal 1° al 10° posto tra le 30 riviste riportate in Tabella 2. Infine, è interessante notare che il criterio tra quelli considerati che sembra essere in maggiore accordo con i risultati attesi è basato sul numero complessivo di citazioni. In questo caso le sei riviste in questione occupano le prime sette posizioni. L'unica rivista che si inserisce tra queste sei è *Statistics in Medicine* che è terza come numero complessivo di citazioni.

¹ Si rinvia alla pagina web <http://new.sis-statistica.org/sis/valutazione-della-ricerca-nelle-scienze-statistiche/> per una descrizione dettagliata

² <http://old.sis-statistica.org/files/pdf/2009/Valutazione.pdf>

³ http://old.sis-statistica.org/files/pdf/2011/indagine_sis_percezione_delle_riviste_di_maggior_prestigio.pdf

n.	Titolo abbreviato	IF	IF5		Citazioni		Articoli (2013)		AIS		h-index (5 anni)	
			valore	pos.	numero	pos.	numero	pos.	valore	pos.	valore	pos.
1	J R STAT SOC B	5,721	6,016	2	14568	6	34	26	5,932	2	37	8
2	J STAT SOFTW	3,801	6,538	1	3717	11	55	17	3,730	4	45	4
3	ECONOMETRICA	3,504	5,111	3	22304	2	67	14	9,847	1	75	1
4	MULTIVAR BEHAV RES	2,971	2,464	14	3348	14	31	28	1,518	19	20	27
5	STAT METHODS MED RES	2,957	3,155	8	2302	19	39	24	2,045	14	25	22
6	STOCH ENV RES RISK A	2,673	2,462	15	1653	23	149	3	0,595	30	28	18
7	BAYESIAN ANAL	2,576	2,321	18	716	30	32	27	1,887	15	25	22
8	ANN STAT	2,442	4,209	4	13953	7	108	7	4,762	3	61	2
9	CHEMOMETR INTELL LAB	2,381	2,533	13	5322	9	139	4	0,643	27	31	14
10	J BUS ECON STAT	2,319	2,641	12	3441	13	37	25	3,725	5	40	7
11	BIostatISTICS	2,236	3,170	7	2742	16	58	15	2,503	9	30	15
12	J AM STAT ASSOC	2,114	3,008	9	23305	1	113	6	3,362	6	47	3
13	STAT MED	2,037	2,828	11	17469	3	390	1	1,585	18	42	6
14	FUZZY SET SYST J	1,880	2,263	19	11823	8	177	2	0,639	28	45	4
15	CHEMOMETR	1,803	2,064	21	2822	15	50	20	0,648	26	24	24
16	TECHNOMETRICS	1,790	2,164	20	5291	10	40	23	1,596	17	24	24
17	STAT COMPUT	1,746	2,028	22	2218	21	57	16	1,719	16	28	18
18	STAT SCI	1,690	3,411	6	3503	12	31	28	3,356	7	28	18
19	ANN APPL STAT	1,689	2,442	16	1224	26	99	9	2,249	10	34	10
20	J COMPUT BIOL	1,670	1,575	27	2615	17	76	13	0,897	22	32	13
21	STATA J	1,646	3,554	5	1939	22	51	19	2,159	12	27	21
22	STATISTICS	1,594	1,322	30	843	29	92	10	0,748	24	13	30
23	J R STAT SOC A STAT	1,573	2,832	10	2222	20	44	22	2,208	11	29	17
24	IEEE ACM T COMPUT BI	1,536	1,497	28	1076	28	139	4	0,613	29	30	15
25	BRIT J MATH STAT PSY	1,527	1,974	23	1096	27	27	30	1,292	21	18	29
26	BIOMETRICS	1,521	1,943	24	15966	4	108	7	1,494	20	34	10
27	BIOMETRIKA	1,506	2,373	17	15358	5	77	12	2,648	8	35	9
28	ENVIRONMETRICS	1,486	1,391	29	1283	25	46	21	0,716	25	20	27
29	PROBABILIST ENG MECH	1,460	1,700	25	1295	24	53	18	0,860	23	23	26
30	PROBAB THEORY REL	1,458	1,586	26	2335	18	83	11	2,148	13	34	10

Tabella 2: Lista delle prime 30 riviste del settore “STATISTICS & PROBABILITY” secondo WoS in termini di IF al 2013, con indicazione dell’IF a 5 anni, numero di citazioni, numero di articoli (pubblicati nel 2013), AIS e h-index a 5 anni, insieme alla rispettiva posizione in graduatoria (pos.) secondo questi indici; in evidenza le sei riviste ritenute di maggiore prestigio.

L’aspetto principale che emerge dall’analisi precedente è che IF, IF a 5 anni e AIS possono dare dei risultati inattesi per le riviste ritenute di maggiore prestigio, collocando queste riviste al di sotto di altre ritenute di prestigio più basso. Una prima spiegazione di questo fenomeno è che questi indici misurano un aspetto che non è necessariamente collegato al prestigio di una rivista in quanto si tratta di indici di impatto o diffusione. Non a caso, le riviste che superano ampiamente alcune delle riviste di maggiore prestigio sono caratterizzabili come riviste trasversali, e quindi possono raccogliere facilmente citazioni da altri settori quali economia o medicina, o si occupano di aspetti computazioni e software, così da proporre contenuti di grande visibilità in quanto riguardano aspetti pratici, o di applicabilità di metodi statistici, con una ricaduta immediata. Invece, articoli che propongono metodi statistici spesso di una elevata complessità hanno tipicamente una ricaduta a livello di comunità scientifica che richiede un periodo piuttosto lungo. Non a caso, i risultati che sembrano meno soddisfacenti dal confronto di cui sopra sono quelli in termini di IF che, a differenza

degli altri indici, valuta le citazioni su un orizzonte di soli due anni.

Pur avendo stabilito che il prestigio di una rivista non coincide pienamente con la sua diffusione misurata in termini di IF, IF a 5 anni e AIS, occorre capire perché questa caratteristica sembra essere misurata in modo migliore dall’h-index e dal numero complessivo di citazioni, come è stato evidenziato nell’analisi di cui sopra. Innanzitutto IF (e indici simili) e h-index, pur utilizzando il numero di citazioni, si basano su due procedimenti di calcolo molto diversi. Ciò è confermato dai bassi livelli di correlazione tra questi indici che sono riportati in Tabella 1. In particolare, IF premia le riviste che per cui la diffusione media di un articolo è elevata, mentre h-index premia riviste che contengono un alto numero di articoli che hanno un impatto molto elevato, a prescindere dall’impatto dei rimanenti articoli (Bartolucci, 2012). In altri termini, a livello di politica editoriale, entrambi gli indicatori possono essere aumentati pubblicando articoli che hanno una elevata visibilità e quindi possono raccogliere molte citazioni. Tuttavia, mentre l’IF viene penalizzato dal fatto di pubblicare anche articoli con

bassa visibilità, l'h-index non ne risulta influenzato. Non a caso, la correlazione tra il numero di articoli pubblicati nel 2013 e IF, IF5 e AIS è negativa, mentre questa correlazione è positiva con l'h-index (Tabella 1).

Un'illustrazione del fenomeno descritto sopra la si può avere andando a confrontare due riviste che hanno tipologie di lettori piuttosto simili, *Journal of the Royal Statistical Society - series B* e *Journal of the American Statistical Association*, che però pubblicano annualmente un numero molto diverso di articoli, con la rivista inglese che ha pubblicato meno di un terzo degli articoli della rivista americana nel 2013. Come emerge dalla Tabella 2, la prima rivista surclassa la seconda in termini di IF (5,721 contro 2,114), in termini di IF5 (6,016 contro 3,008) e di AIS (5,932 contro 3,362). Tuttavia, la seconda risulta superiore in termini di h-index rispetto alla prima con un valore di 47 per il *Journal of the American Statistical Association* e di 37 per *Journal of the Royal Statistical Society - series B*.

L'altro aspetto di interesse è la forte concordanza, già evidenziata in precedenza, tra prestigio di una rivista e il numero complessivo di citazioni. Considerando anche questo aspetto si può quindi concludere che una rivista è ritenuta di prestigio se ha avuto storicamente una elevata diffusione (numero di citazioni complessivo) e tende a pubblicare periodicamente un grande numero di articoli con elevato impatto (misurato da h-index) a prescindere dal fatto che vengano pubblicati o meno anche articoli con limitata diffusione. Rimane tuttavia un aspetto da investigare. Secondo il parere dell'autore della presente nota, il prestigio è direttamente collegato al livello di difficoltà nel pubblicare in una certa rivista. Tuttavia, misurare questo aspetto tramite un semplice tasso di accettazione, ottenuto dividendo il numero di articoli pubblicati in un certo periodo per il numero di articoli sottoposti, sarebbe semplicistico in quanto non tiene conto della autoselezione nell'invio di articoli che viene attuata dagli autori. Pertanto, nella sezione seguente, viene proposto un metodo per instaurare una banca dati che permetta di ottenere in modo opportuno una valutazione del livello di difficoltà in questione.

3. Stima del livello di difficoltà

Una banca dati utilizzabile per stimare in modo opportuno il livello di difficoltà nel pubblicare in una certa rivista potrebbe essere creata andando a considerare la "storia" di ogni articolo pubblicato

in un certo periodo di tempo. In particolare, per ognuno di tali articoli l'informazione di maggiore rilievo è quella inerente la lista di riviste a cui l'articolo (nella stessa forma o in una forma sostanzialmente equivalente) è stato sottoposto (e quindi rifiutato) in precedenza. Un archivio di questo tipo potrebbe essere formato richiedendo agli editori delle riviste nello stesso settore di comunicare periodicamente allo stesso ente la lista degli articoli che sono stati sottoposti a tali riviste con indicazione dell'esito, essenzialmente accettazione o rifiuto. Questi dati andrebbero incrociati poi in modo opportuno andando a creare dei record che, articolo per articolo, ne descrivono il percorso editoriale.

Per fare un semplice esempio, un record come quello descritto in precedenza potrebbe contenere delle informazioni riferite allo stesso articolo del tipo: "rifiutato dalla rivista A", "rifiutato dalla rivista B", "accettato dalla rivista C". Il punto cruciale è come elaborare dati di questo tipo. Una possibilità interessante a questo riguardo è offerta dal modello di Bradley-Terry (Bradley, Terry, 1952) che è noto in ambito di competizioni sportive ed è stato già applicato per la valutazione delle riviste sulla base delle citazioni incrociate, come illustrato in Turner, Firth (2012). Infatti, un ipotetico record del tipo di cui sopra può essere interpretato come il risultato di tre "match" tra le riviste A, B, C. Il primo è tra le riviste A e B e si conclude con un pareggio (entrambe rifiutano l'articolo); il secondo "match" è tra le riviste A e C e vede A vincere (rivista tra le due su cui è più difficile pubblicare); il terzo "match" è tra le riviste B e C e vede B vincere. Ovviamente, oltre a un'applicazione statistica tramite il modello di Bradley-Terry, semplici indici di difficoltà posso essere calcolati al fine di rendere più comprensibile il procedimento da parte di non esperti di materie statistiche.

Una banca dati della tipologia descritta in precedenza non risulta ancora esistere anche se si riscontra l'esistenza di indagini, svolte da certe case editrici internazionali, in merito al percorso editoriale di certi articoli. Inoltre, occorre considerare che alcune riviste propongono meccanismi di revisione per cui i rapporti dei Referee su un articolo rifiutato possono essere direttamente passati a una nuova rivista, permettendo quindi di ricostruire informazioni analoghe a quelle descritte in precedenza. Si noti, in particolare, che rispetto a banche dati come quelle messe a disposizione da WoS o da GS, la banca dati in questione conterrebbe informazioni non solo sugli articoli pubblicati, ma anche su rifiuti degli stessi articoli da parte di altre riviste.

È ovvio che la costruzione della banca dati può presentare delle difficoltà connesse con la dispo-

nibilità degli editori nel divulgare tutte le informazioni necessarie. Inoltre, opportuni meccanismi devono essere implementati per salvaguardare la privacy e per evitare che le scelte editoriali di un editore vengano influenzate da quelle effettuate dagli editori che hanno esaminato lo stesso articolo in precedenza. Tuttavia, si ritiene che una banca dati del tipo di cui si auspica la costruzione possa fornire delle informazioni importanti che permettano in modo praticamente automatico di valutare la difficoltà nel pubblicare in una rivista senza dover ricorrere a indagini basate su interviste dei ricercatori operanti in un certo settore. Queste ultime, infatti, oltre a comportare costi considerevoli, possono anche presentare distorsioni dovute, tra l'altro, a un elevato tasso di mancate risposte da parte di particolari tipologie di ricercatore. Ovviamente, la disponibilità della banca dati basata su dati forniti direttamente dagli editori permetterebbe poi di studiare in che misura il prestigio di una rivista è collegato alla difficoltà di pubblicarvi. In ogni caso, i risultati della elaborazione di questi dati potrebbero proficuamente essere utilizzati, in aggiunta ad

indicatori già utilizzati, come IF e h-index, nella costruzione di liste di riviste a fini valutativi.

Riferimenti bibliografici

- Bartolucci, F., 2012, *On a possible decomposition of the h-index*. Letter to the Editor of the Journal of the American Society for Information Science and Technology, 63, 2126-2127.
- Bartolucci, F., Dardanoni, V., Peracchi, F., 2013, *Ranking scientific journals via latent class models for polytomous item response data*, EIEF working paper series, 13.
- Bradley, R.A., Terry, M.E., 1952, *Rank analysis of incomplete block designs I: The method of paired comparisons*, Biometrika, 39, 324-345.
- Hirsch, J.E., 2005, *An index to quantify an individual's scientific research output*, Proceedings of the National Academy of Science, 102, 16569-16572.
- Turner, H., Firth, D., 2012, *Bradley-Terry models in R: The BradleyTerry2 package*, Journal of Statistical Software, 48.

La VQR di area13: una riflessione di sintesi

Alberto Baccini

Dipartimento di Economia Politica e Statistica
Università degli Studi di Siena

baccini@unisi.it

Sommario

L'esercizio di Valutazione della Qualità della Ricerca 2004-2010 (VQR) condotto dall'Agenzia Nazionale di Valutazione dell'Università e della Ricerca (ANVUR) è stato accolto da critiche che ne hanno sottolineato l'inadeguatezza rispetto agli standard con cui analoghi esercizi sono stati condotti in altri paesi. Nel lavoro si mostra che l'Area 13 è stata uno degli snodi critici più importanti della VQR, dove emergono in modo drammatico problemi metodologici e tecnici tali da rendere i risultati sistematicamente distorti.

1. Introduzione

In molti paesi OECD vengono condotti esercizi nazionali di valutazione della ricerca condotti direttamente dal governo o da altre istituzioni, con lo scopo di verificare lo stato di salute dei sistemi nazionali di ricerca, e permettere l'adozione di politiche informate di finanziamento statale. La pratica fu adottata per la prima volta dal governo britannico di Margaret Thatcher che nel 1986 condusse il primo RAE (Research Assessment Exercise). L'Italia svolse il primo esercizio di valutazione triennale della ricerca (VTR) della ricerca ben 15 anni dopo (A. Baccini, 2013b), ispirandosi proprio a quel modello, e dando luogo ad un'esperienza condotta secondo standard ragionevoli, costosa (A. Baccini et al., 2013), con alcuni difetti che avrebbero dovuto essere superati nelle applicazioni successive (Emanuela Reale, 2008). Solo nel luglio 2011 quell'esperienza ha avuto un seguito con l'inizio della Valutazione della Qualità della Ricerca 2004-2010 (da ora in poi VQR) guidata dalla allora neonata Agenzia Nazionale di Valutazione della Ricerca (ANVUR). La VQR è apparsa fin da subito come profondamente diversa sia rispetto al precedente esercizio di valutazione, che rispetto alle metodologie adottate in ambito internazionale. Queste innovazioni sono state accolte da rilievi critici che ne hanno sottolineato l'inadeguatezza rispetto agli standard prevalenti a livello internazionale (Baccini, 2012a, Banfi, 2012, Coin, 2012). Nel mio intervento intendo riprendere e sottolineare alcune di queste critiche. In particolare intendo mostrare che la valutazione di Area 13, cioè la valutazione delle scienze economiche e statistiche, è stata uno degli snodi critici più importanti della VQR, dove sono emersi in modo esemplare problemi metodologici

e tecnici tali da rendere i risultati sistematicamente distorti. Dopo aver discusso brevemente le più rilevanti criticità, raccogliendole in quattro categorie distinte, concludo con alcune proposte di revisione in grado di eliminare le principali distorsioni.

2. La composizione del GEV13: un problema di *fairness*

Il primo elemento critico della VQR riguarda in generale l'assenza di trasparenza nelle procedure di nomina del personale chiamato a svolgere l'esercizio di valutazione. In (Baccini, 2013a) sono discussi in dettaglio i problemi relativi alla *governance* complessiva dell'esercizio, per cui è possibile soffermarci qui sulla sola questione della composizione dei Gruppi di esperti di valutazione (GEV). I GEV, uno per ogni area di ricerca, sono lo snodo cruciale della VQR perché hanno svolto tutte le attività connesse direttamente alla valutazione della ricerca. In particolare hanno definito i criteri di valutazione bibliometrica, hanno definito le procedure per decidere quali lavori dovessero essere valutati con la bibliometria e quali con la *peer review*, hanno scelto i revisori, li hanno coordinati, hanno sintetizzato i rapporti di revisione, hanno valutato in molti casi direttamente i lavori sottoposti a valutazione. La delicatezza di tali attività avrebbe dovuto suggerire ad ANVUR di prendere molto sul serio la questione della nomina dei membri dei GEV. Se si prende il caso del RAE britannico la composizione dei panel di esperti è sottoposto ad uno scrutinio attento per valutarne l'equità (*fairness*) della composizione in termini di genere e competenze (RAE, 2004). La *fairness* nella composizione di un panel di esperti è funzionale

alla rappresentazione di punti di vista e approcci metodologici diversificati (Baccini and Ricciardi, 2012).

In Italia il Decreto ministeriale (DM MIUR n. 17 del 15/07/2011) che ha governato la VQR era molto laconico, limitandosi a prevedere che fosse il consiglio direttivo di ANVUR a nominare i membri dei GEV ed a scegliere tra loro il presidente. Il consiglio direttivo di ANVUR non ha seguito neanche questa regola ed ha nominato prima i presidenti dei GEV (il file con i nomi dei presidenti è datato 10 ottobre 2011 ed è stato reso pubblico intorno a quella data), e dopo quasi due mesi i membri dei GEV (12 dicembre 2011). I presidenti dei GEV, come dichiarato da uno dei membri del consiglio direttivo di ANVUR, sono stati consultati nella fase di stesura delle regole di funzionamento della VQR (Anonymous, 2011); è pertanto verosimile che abbiano anche avuto voce in capitolo nella scelta degli altri membri dei GEV.

Nel caso del GEV 13 non sono state rispettate regole elementari di fairness per quanto riguarda la composizione di genere, con appena il 16,7% di donne – che hanno poi scelto revisori donne nel solo 20% dei casi (ANVUR, 2013c). E soprattutto il GEV è risultato composto da studiosi strettamente collegati tra loro da rapporti di coautoraggio, come risulta dalla Figura 1. L'assenza di fairness nella composizione del panel ha probabilmente ridotto al minimo le voci di dissenso rispetto ai metodi e alle regole di valutazione adottate dal GEV.

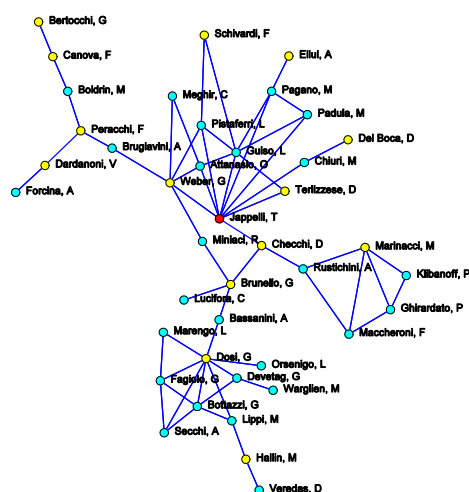


Figura 1. La rete di coautoraggio del GEV di Area 13. Fonte: (Baccini, 2011). Legenda: In rosso il presidente del GEV; in giallo i membri del GEV, in celeste coautori di membri del GEV che non fanno parte del GEV. Una linea che unisce due nomi indica che i due studiosi hanno firmato insieme almeno un articolo.

3. L'insostenibile inconsistenza della valutazione bibliometrica basata sulla classifica delle riviste

Il GEV di Area 13 è l'unico che, sulla base di una presunta specialità delle discipline economiche e statistiche, ha ritenuto di adottare una bibliometria *sui generis* che non trova riscontro in nessun altro esercizio di valutazione condotto a livello internazionale. In particolare la valutazione degli articoli su rivista è stata condotta sulla base di una classificazione delle riviste sviluppata dallo stesso GEV. Anche in questo caso va notata anzitutto l'anomalia procedurale di un gruppo di esperti che si sostituisce alla comunità scientifica nel definire la lista delle riviste e la loro classificazione in fasce di qualità. Di norma questo compito è riservato alle comunità scientifiche che attivano procedure più o meno complesse di consultazione degli studiosi, come avvenuto per esempio in Francia con le liste di riviste di economia sviluppate dal CNRS.

In secondo luogo va notato che l'uso di liste di riviste come surrogato della valutazione bibliometrica diretta degli articoli è una procedura che dà luogo a risultati che sono considerati generalmente privi di validità. Possiamo limitarci a ricordare a questo proposito che la San Francisco Declaration on Research Assessment (DORA, 2012), ad oggi sottoscritta da circa 550 organizzazioni scientifiche di tutto il mondo, stigmatizza l'uso improprio di misure di impatto delle riviste –principalmente l'Impact Factor– come surrogato della valutazione di impatto e qualità degli articoli e dei ricercatori. Una presa di posizione del tutto simile è stata adottata dal board dell'Institute of Electrical and Electronics Engineers (IEEE), la più grande organizzazione internazionale di studiosi di tecnologie, in una dichiarazione sull'uso corretto degli indicatori bibliometrici dove si legge che “*the use of the bibliometric index of a journal in which a researcher publishes (typically the Impact Factor (IF) as a proxy for the quality of his/her specific paper is a common example of a technically incorrect use of bibliometrics*” (Boards of Directors IEEE, 2013). Per una rassegna sintetica e aggiornata sul tema si rimanda a (Brembs et al., 2013; Brembs, 2013).

In Area13 il GEV ha addirittura messo a punto un'inedita metodologia di classificazione delle riviste, considerata statisticamente inconsistente (Torelli, 2014). È inoltre da notare che il GEV 13 ha fatto affidamento su Google Scholar generalmente considerato allo stato inutilizzabile a fini di valutazio-

ne perché “[it] is not built from structured records, that is from metadata fields. Rather than using the author, affiliation, reference etc. data provided by publishers, Google Scholar parses full-text to obtain its best guess for these items.” (Hicks and Wang, 2009). Sul punto si può vedere anche (Baccini, 2012b; Delgado López-Cózar et al., 2014; Gambardella, 2012; Labbé, 2010).

Last but not least, la classificazione delle riviste non adotta alcuna normalizzazione in relazione alle comunità di riferimento delle riviste. È ben noto che gli indicatori citazionali delle riviste sono sensibili alle modalità di comunicazione e citazione prevalenti nelle diverse comunità scientifiche (Zitt, 2010). Ciò rende particolarmente difficile operare comparazioni tra riviste di settori diversi: per esempio in area medica le riviste di oncologia hanno indicatori più elevati di quelle di epidemiologia che ha loro volta hanno indicatori migliori delle riviste interdisciplinari; analogamente in matematica le riviste che pubblicano articoli applicati sono molto più citate delle riviste di matematica pura che a loro volta sono molto più performanti delle riviste di storia della matematica. Per far fronte a questo tipo di problemi, e per permettere la comparazione tra riviste di campi diversi sono stati sviluppati indicatori bibliometrici specifici, tra i quali forse il più noto è il Source Normalized Impact per Paper (SNIP) (Waltman et al., 2013). Anche in questo caso anziché utilizzare la migliore tecnologia bibliometrica disponibile, il GEV 13 ha preferito una soluzione inedita (Maccabelli, 2012), adottando indicatori distorti settorialmente.

4. Bibliometria, *peer review* o entrambe? Un esperimento fallito

Gli esercizi di valutazione possono essere classificati secondo due modelli. Esercizi basati sulla *peer review*, come il RAE/REF britannico. Ed esercizi basati su un modelli bibliometrici puri, del tipo dell'ERA australiano. ANVUR ha ritenuto di perseguire un modello ibrido che usa sia *peer review* che bibliometria per valutare i singoli lavori di ricerca. Mentre le scienze umane, sociali e la giurisprudenza hanno fatto ricorso esclusivo alla valutazione *peer*, nelle scienze dure – le cosiddette aree bibliometriche corrispondenti alle aree CUN 1-9- e in Area 13 alcuni lavori sottoposti a valutazione sono stati valutati da pari, ed altri, la schiacciante maggioranza, con indicatori bibliometrici (ANVUR, 2013b).

Per giustificare questa scelta ANVUR ha altresì condotto un “esperimento”, sottoponendo il 5% circa degli articoli su rivista a una doppia valutazione, sia bibliometrica che attraverso *peer review*, per verificare la concordanza dei risultati tra le due tecniche. I risultati presentati in appendice al rapporto finale indicano chiaramente che i risultati che si ottengono con le due tecniche sono diversi (ANVUR, 2013a). In tutte le aree bibliometriche il grado di concordanza (kappa di Cohen) calcolato da ANVUR è compreso tra 0,17 e 0,35, che significa (Richard Landis and Koch, 1977) un grado di concordanza dei risultati delle due tecniche compreso tra limitato e moderato. Area 13 rappresenta un'eccezione con un indice di concordanza relativamente elevato (0,68).

Questo risultato anomalo di Area13 è frutto di un “esperimento” realizzato con modalità radicalmente difforni rispetto a quanto accaduto nelle altre aree bibliometriche. In estrema sintesi (per una analisi di dettaglio si veda (Baccini, 2014)):

1. Il campionamento degli articoli in Area 13, a differenza di quanto avvenuto nelle altre aree, non è stato casuale, poiché è stato aggiustato tenendo conto delle richieste di accedere alla *peer review* fatte da parte degli autori degli articoli;
2. I revisori di Area 13, a differenza dei revisori delle altre aree, erano nelle condizioni di sapere che stavano partecipando alla sperimentazione;
3. I revisori di Area 13, a differenza dei revisori delle altre aree, erano nelle condizioni di conoscere con estrema precisione il giudizio derivante dall'applicazione delle bibliometria (era sufficiente conoscere la sede di pubblicazione per avere immediatamente il giudizio bibliometrico);
4. In Area 13, a differenza di quanto avvenuto nelle altre aree, il giudizio dei due revisori non è stato sintetizzato sulla base di un algoritmo (media dei punteggi). È stato invece definito da un gruppo di consenso composto da (almeno) due membri del GEV, che ha considerato i due giudizi dei revisori come mere informazioni per definire il giudizio finale;
5. I membri del GEV che hanno costituito i gruppi di consenso sapevano che il loro giudizio finale sarebbe stato utilizzato per la sperimentazione;
6. Di fatto, adottando ipotesi molto conservative, si può calcolare che non meno del 53% dei giudizi derivanti da *peer review* non è stato espresso dai revisori, ma direttamente dai membri del GEV13.

Queste condizioni non solo hanno modificato l'esperimento in modo tale da alterarlo significativamente rispetto a quanto avvenuto nelle altre aree, ma sembra di poter affermare che sono state le procedure a determinare il risultato finale. Di fatto l'elevato indice di concordanza calcolato per Area13 indica che la valutazione basata sulla lista di riviste sviluppata dal GEV 13 si accorda con i giudizi che i membri del GEV 13 hanno dato quando essi hanno deciso le valutazioni nei Gruppi di consenso.

5. I risultati finali dipendono dal “mix valutativo”

Giuseppe De Nicolao (2014) ha mostrato che nelle aree bibliometriche, l'uso di indicatori bibliometrici differenziati per area e settore disciplinare ha reso i risultati dell'esercizio di valutazione non comparabili tra aree e tra settori disciplinari. Il diverso mix di peer review e bibliometria adottato nelle diverse aree, ha giocato un ruolo determinante nell'indurre distorsioni. In particolare, come riconosciuto dalla stessa ANVUR, i risultati delle valutazioni bibliometriche sono stati in media “più generosi” dei risultati raggiunti con la peer review. Per avere un'idea dell'entità della distorsione può essere utile confrontare i risultati VQR con quelli VTR 2001-2004. L'esercizio VTR utilizzava la sola peer review come strumento di valutazione dei lavori di ricerca, ed era ispirato da una logica completamente diversa (CIVR, Comitato di Indirizzo per la Valutazione della Ricerca and MIUR, Ministero dell'Istruzione dell'Università e della Ricerca, 2006; Reale, 2008). Venne infatti chiesto ad ogni struttura valutata di sottoporre a valutazione i migliori prodotti della ricerca. Per ogni struttura era definito il numero massimo di lavori da sottoporre a valutazione, ma non faceva alcuna differenza se i lavori provenivano da molti ricercatori o soltanto da uno. Nell'esercizio VQR è stato invece richiesto di inviare tre prodotti per ogni ricercatore che facesse parte della struttura valutata. La VQR è palesemente rivolta a stimare la “qualità media” della ricerca di una struttura, mentre la VTR era invece rivolta a valutare la presenza di ricerca eccellente nella struttura.

Per capire la differenza può essere utile confrontare le modalità di selezione da parte di un dipartimento dei lavori da sottoporre a valutazione nei due esercizi. Immaginiamo per semplicità un dipartimento tenuto ad inviare a VTR e VQR uno stesso numero di prodotti. Per la VTR il direttore agì in questo modo: ordinò secondo qualche criterio di qualità

tutti i prodotti di ricerca di tutti i docenti del dipartimento, dal migliore al peggiore. Scelse quindi i migliori della lista fino ad arrivare al numero richiesto. Nella VQR invece il direttore ha chiesto a ciascuno ricercatore del suo dipartimento di inviare i suoi tre lavori migliori. Se un ricercatore particolarmente bravo e produttivo nel periodo considerato avesse prodotto un numero di lavori eccellenti superiore a tre, il direttore non avrebbe avuto modo di sottoporre i lavori eccellenti eccedenti i tre alla valutazione, come accade invece per la VTR.

Da questo segue che, proprio a causa del disegno dei due diversi esercizi di valutazione, ci si dovrebbe aspettare, per ogni area scientifica, che la quota di lavori eccellenti rilevata con la VQR sia inferiore a quella rilevata con la VTR. La Tabella 2 confronta la quota di lavori valutati eccellenti nella VTR e nella VQR per area CUN.

Come si può notare nelle aree CUN 10, 11, 12 e 14, cioè nelle aree dove la valutazione è stata condotta con la sola *peer review*, la quota di prodotti eccellenti è drasticamente più bassa nella VQR che nella VTR. Nelle scienze giuridiche (Area12), per esempio, la quota di lavori giudicati eccellenti è scesa dal 20% della VTR all'11% nella VQR. Nella aree bibliometriche è successo invece l'opposto, con risultati eclatanti nell'area dell'ingegneria (Area 9) dove, a causa di ulteriori distorsioni nella procedura (Baccini and De Nicolao, 2012), i prodotti eccellenti sono passati dal 21% della VTR al 51,8% della VQR.

	Percentuale di lavori giudicati eccellenti nella VQR	Percentuale di lavori giudicati eccellenti nella VTR	Differenza tra percentuali VQR-VTR
Area1	44,8%	36,0%	8,8%
Area2	57,2%	52,0%	5,2%
Area3	56,8%	31,0%	25,8%
Area4	30,2%	34,0%	-3,8%
Area5	41,6%	33,0%	8,6%
Area6	38,4%	25,0%	13,4%
Area7	43,0%	10,0%	33,0%
Area8 a	43,2%	22,0%	21,2%
Area8 b	9,2%	22,0%	-12,8%
Area9	52,8%	21,0%	31,8%
Area10	24,2%	51,0%	-26,8%
Area11 a	16,0%	28,0%	-12,0%
Area11 b	35,8%	28,0%	7,8%
Area12	11,2%	20,0%	-8,8%
Area13	19,3%	17,0%	2,3%
Area14	8,9%	20,0%	-11,1%

Tabella 2. Prodotti giudicati eccellenti nella VQR e nella VTR. Un confronto.

Fonte: Elaborazioni su dati (ANVUR, 2013b) e (CIVR, Comitato di Indirizzo per la Valutazione della Ricerca, 2006)

Da queste considerazioni segue una conseguenza non certo secondaria sui risultati finali della VQR. Il voto finale attribuito alla struttura, all'area o ai

settori scientifici valutati è stato sistematicamente più basso tutte le volte che si è fatto ricorso relativamente più forte alla peer review. Di fatto i risultati medi di strutture/aree/settori riflettono semplicemente il diverso mix di bibliometria/peer review che è stato usato nella valutazione dei prodotti di strutture/aree/settori. Le strutture/aree/settori con un maggior numero di prodotti valutati attraverso peer review hanno ottenuto risultati medi sistematicamente più bassi.

In particolare in Area 13 i risultati in termini di voto medio conseguiti dai SSD è inversamente correlato (-0,77) alla quota di prodotti valutati con peer review, come si vede dal grafico 1.

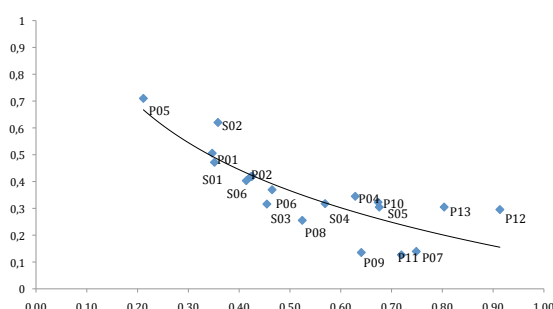


Grafico 1. Voto medio dei settori scientifico disciplinari (SSD) di Area 13 e quota di prodotti giudicati con peer review. Fonte: Elaborazioni su dati (ANVUR, 2013b)

6. Conclusioni

I risultati della VQR sono il frutto di una procedura macchinosa e complessa che ha introdotto tali e tanti elementi distorsivi, da renderli del tutto inutilizzabili a fini conoscitivi e di *policy*. Questo è vero per tutte le aree scientifiche in cui è stata organizzata la VQR, ed in particolare per l'Area 13. Per rimediare agli errori commessi non è possibile intervenire con semplici correttivi sulla struttura VQR, ma è necessario un ripensamento dell'intero modello.

L'adozione del modello britannico REF avrebbe alcuni indubbi vantaggi rispetto alla VQR. Il principale è che l'uso sistematico ed esclusivo della *peer review* insieme ad una corretta definizione di una unica scala di qualità dei prodotti, permetterebbero di salvaguardare la comparabilità dei risultati tra settori e aree diverse.

Nel modello REF è inoltre chiarissima la responsabilità della valutazione che è completamente in capo ai membri del panel. Ogni lavoro viene infatti valutato da un solo revisore membro del panel o designato direttamente e nominativamente da un

membro del panel.

Questo modo di procedere rende delicatissima la questione della composizione dei gruppi di esperti. Per la loro costituzione si dovrebbero seguire procedure pubbliche trasparenti. La composizione dei panel dovrebbe permettere la copertura di gran parte dei settori disciplinari di area. La composizione dei GEV dovrebbe altresì rispettare elementari principi di *fairness* per quanto riguarda genere, età e provenienza geografica dei membri del panel. Anche in questo caso sarebbe auspicabile l'adozione del modello di analisi britannico per la valutazione della corretta composizione dei panel.

Sarebbe altresì opportuno adottare qualche regola aggiuntiva che impedisca quanto documentato per il GEV di Area13. Si potrebbe per esempio adottare una semplice regola per cui non più del 10% dei membri di un panel possono essere co-autori tra loro o avere un co-autore comune. Si dovrebbe inoltre prevedere che nessun presidente GEV, e nessuno dei membri GEV della VQR 2004-2010 possa far parte dei GEV nel nuovo esercizio di valutazione.

Molte dei problemi illustrati in questo articolo erano stati sollevati prima dell'inizio della scorsa VQR, per cui non c'è da essere ottimisti sul futuro della valutazione della ricerca in Italia.

Riferimenti bibliografici

Anonymous, 2011, "La Ricerca Degli Atenei Italiani All'esame Dell'anvur," *Pisainforma.it*, Pisa, Comune di Pisa,

ANVUR, 2013a, "Rapporto Finale Vqr. Appendice B il Confronto Tra Valutazione Peer e Valutazione Bibliometrica," http://www.anvur.org/rapporto/files/Appendici/VQR2004-2010_AppendiceB.pdf.

_____. 2013b, "Rapporto Finale. Valutazione Della Qualità Della Ricerca 2004-2010", http://www.anvur.org/rapporto/files/VQR2004-2010_RapportoFinale_parteprima.pdf, Roma.

_____. 2013c, "Valutazione della Qualità della Ricerca 2004-2010 (Vqr 2004-2010), Rapporto Finale di Area, Gruppo di Esperti della Valutazione dell'Area 13 (Gev13)", Roma.

Baccini, A., 2013a, "Come e Perché Ridisegnare la Valutazione", Il Mulino, Bologna, 80-87.

_____. 2011. "Gli Esperti Di Valutazione All'italiana," <http://www.roars.it/online/gli-esperti-di-valutazione-allitaliana/>.

_____, 2012a, "La Valutazione Della Ricerca in Ita-

lia”, *Scuola democratica Learning for democracy*, 5, giugno, 5.

_____, 2013b, “La Valutazione Della Ricerca in Italia, in ritardo e tecnicamente inadeguata”, *Ricerca*, (5/6), 24-27.

_____, 2014, “Lo Strano Caso delle Concordanze della Vqr,” <http://www.roars.it/online/lo-strano-caso-delle-concordanze-della-vqr/>. www.roars.it.

_____, 2012b, “Primo Capitolo ed il suo H-Index”, <http://www.roars.it/online/primo-capitolo-e-il-suo-h-index/>.

Baccini, A., F. Coin and G. Sirilli, 2013, “Costi e Benefici della Valutazione della Ricerca e della Didattica”, *Paradoxa*, VII (2), 49-61.

Baccini, A. and G. De Nicolao, 2012, “Vqr: Tutte le Valutazioni sono Uguali, ma alcune sono più Uguali delle Altre,” [www.roars.it](http://www.roars.it/online/vqr-tutte-le-valutazioni-sono-uguali-ma-alcune-sono-piu-uguali-delle-altre/), <http://www.roars.it/online/vqr-tutte-le-valutazioni-sono-uguali-ma-alcune-sono-piu-uguali-delle-altre/>.

Baccini, A. and M. Ricciardi, 2012, “Vqr, la Composizione dei Gev ed una Questione di Fairness”, <http://www.roars.it/online/vqr-la-composizione-dei-gev-ed-una-questione-di-fairness/>.

Banfi, A. 2012. “Qualche Osservazione Su Valutazione Della Ricerca E Scienze Umane.” *Bullettino dell’Istituto di Diritto Romano “Vittorio Scialoja”*, 106(1), 437-62.

Brembs, B., K. Button and M. Munafò, 2013, “Deep Impact: Unintended Consequences of Journal Rank”, *Frontiers in Human Neuroscience*, 7.

Brembs, B., 2013, “What Ranking Journals Has in Common with Astrology”, *RT. A Journal on Research Policy and Evaluation*, 1(1).

CIVR, Comitato di Indirizzo per la Valutazione della Ricerca, 2006, “Relazione Finale Vtr 2001-2003,” Roma, CIVR-MIUR,

CIVR, Comitato di Indirizzo per la Valutazione della Ricerca and MIUR, Ministero dell’Istruzione dell’Università e della Ricerca, 2006, “Linee Guida Per La Valutazione Della Ricerca,” Roma.

Coin, F., 2012, “La Valutazione: da Ranking a Profilo di Qualità”, *Scuola democratica*, (5), 160-64.

De Nicolao, Giuseppe, 2014, “Vqr da Buttare? Persino Anvur Cestina i Voti Usati per l’assegnazione Ffo 2013”, <http://www.roars.it/online/vqr-da-buttare-persino-anvur-cestina-i-voti-usati-per-lassegnazione-ffo-2013/>.

Delgado López-Cózar, E., N. Robinson-García and D. Torres-Salinas, 2014, “The Google Scholar Experiment: How to Index False Papers and Manipulate Bibliometric Indicators”, *Journal of the Association for Information Science and Technology*, 65(3), 446-454.

DORA, 2012, “San Francisco Declaration on Research Assessment. Putting Science into the Assessment of Research”, <http://am.ascb.org/dora/>.

Gambardella, M., 2012, “Google Scholar: Soluzione O Problema?”, <http://www.roars.it/online/google-scholar-soluzione-o-problema/>.

Hicks, D. and J. Wang, 2009, “Towards a Bibliometric Database for the Social Sciences and Humanities”, Georgia Institute of Technology,

IEEE, Boards of Directors, 2013. “Appropriate Use of Bibliometric Indicators for the Assessment of Journals, Research Proposals, and Individuals”, http://www.ieee.org/publications_standards/publications/rights/ieee_bibliometric_statement_sept_2013.pdf.

Labbé, C., 2010, “Ike Antkare One of the Great Stars of the Scientific Firmament”, *ISSI newsletter*, 6(2), 48-52.

Landis, J. R. and G. G. Koch, 1977, “The Measurement of Observer Agreement for Categorical Data”, *Biometrics*, 33(1), 159-174.

Maccabelli, T., 2012, “L’inedita Bibliometria di Area 13: Valutazione e Ranking delle Riviste Economiche”, <http://www.roars.it/online/inedita-bibliometria-dellarea-13-valutazione-e-ranking-delle-riviste-economiche/>. www.roars.it.

RAE, 2004, “Units of Assessment and Recruitment of Panel Members”, London, HEFCE-RAE.

Reale E. (ed.), 2008, *La Valutazione della Ricerca Pubblica. Una Analisi della Valutazione Triennale della Ricerca*, Milano, Angeli.

Torelli, N., 2014, “Bibliometria: Fin Dove Arrivare? Alcune Riflessioni su Vqr e Asn per l’Area 13,” Roma, CUN, 6 giugno.

Waltman, L., N. J. van Eck, T. N. van Leeuwen and M. S. Visser, 2013, “Some Modifications to the Snip Journal Impact Indicator”, *Journal of Informetrics*, 7, 272-285.

Zitt, M., 2010, “Citing-Side Normalization of Journal Impact: a Robust Variant of the Audience Factor”, *Journal of Informetrics*, 4, 392-406.

Gli statistici nell'area 13 italiana e nei settori ERC

Daniela Cocchi

Dipartimento di Scienze Statistiche "Paolo Fortunati"
Università degli Studi di Bologna

daniela.cocchi@unibo.it

Sommario

Si illustrano le peculiarità della valutazione delle discipline statistiche, saldamente inserite nell'Area 13 del CUN, proponendo una strategia per ampliare la gamma delle riviste di interesse per gli statistici da trattare in modo quasi automatico per la valutazione. La classificazione degli argomenti di ricerca secondo l'European Research Council è uno schema che deve diventare sempre più popolare tra gli studiosi italiani.

La valutazione della ricerca in Italia ha una tradizione abbastanza lunga, e vede la statistica ormai consolidata nell'Area 13 (Scienze economiche e statistiche) del Consiglio Universitario Nazionale (CUN). Una discussione molto importante da affrontare è come costruire spazi più ampi per la statistica in quell'ambito, ben consapevoli che si tratta di un contesto in cui gli studiosi della nostra disciplina sono destinati a restare.

Riguardo alla valutazione della ricerca in ambito statistico ho la fortuna di un punto di vista privilegiato. Oltre a svolgere la mia ricerca in ambito statistico, ho condotto, dal maggio 2012, un'esperienza di due anni nel Comitato Nazionale dei Garanti della Ricerca (CNGR), formato da 7 esperti e attivo presso il Ministero dell'Istruzione, Università e Ricerca (MIUR). Come è noto, questo Comitato predispone la valutazione dei progetti di ricerca finanziati dallo stesso Ministero. Inoltre sono stata designata nella Commissione per la Valutazione della Ricerca di Ateneo (VRA) dell'Università di Bologna per il biennio 2014-2015, proprio per l'Area 13, e coordino la Commissione Ricerca del Dipartimento di Scienze Statistiche della mia Università.

Il peso dell'Area 13 rispetto al totale degli studiosi incardinati nelle università (al 10 marzo 2014) è abbastanza elevato a livello nazionale (4344 studiosi incardinati su 52103, pari all'8,03%), ed è addirittura più alto nell'Università di Bologna (268/2856, cioè il 9,38%). È ben noto che all'Area 13 afferiscono 19 settori scientifico-disciplinari che, grossolanamente, includono l'area economica e quella aziendale, oltre alla matematica per l'economia e alla storia economica.

La produzione scientifica dell'Area 13 necessita in parte di una valutazione bibliometrica e in parte di

valutazione non bibliometrica: parecchie delle riviste in cui pubblicano gli studiosi dell'Area hanno indicatori bibliometrici. Gli statistici, in particolare, tendono a pubblicare principalmente su riviste di questo tipo.

Queste affermazioni collimano con i risultati della prima Valutazione della Qualità della Ricerca 2004-2010 (VQR). Gli strascichi della VQR fanno emergere vivacemente: a) la difficoltà della definizione delle collocazioni editoriali "di prestigio"; b) l'auspicio di stabilire metodi di classificazione dei prodotti dei risultati della ricerca con modalità per quanto possibile automatiche; c) la necessità di limitare attività di *peer review* negli esercizi di valutazione su popolazioni ampie come quella di tutti i ricercatori di un paese.

Noi statistici abbiamo due ordini di difficoltà: da un lato condividiamo con l'Area 13 i problemi del confronto con le altre aree, dall'altro scontiamo una indubbia diversità rispetto ad altri sotto-settori dell'area.

Il Gruppo di esperti della Valutazione dell'Area 13 (GEV 13) della VQR, infatti, ha considerato, e classificato in 4 classi, 1902 riviste, una parte delle quali è di maggiore interesse per gli statistici. Riguardo all'Area 13, inoltre, l'ANVUR ha arricchito questa lista, introducendone, nella classe più alta, altre 235 fino al 2013.

La classificazione delle riviste è un argomento di discussione e, purtroppo, di contestazione molto appassionante. Preferisco però notare, come tanti di noi hanno già fatto, come, ancor prima di passare alla classificazione, la produzione scientifica degli statistici non sia adeguatamente contemplata dalle due liste ufficiali appena menzionate. La copertura della produzione di un dipartimento con una buona presenza di statistici tramite tali liste non è elevata.

La commissione ricerca del Dipartimento di Scienze Statistiche dell'Università di Bologna ha affrontato la difficoltà aumentando il numero delle liste di riviste. L'arricchimento della popolazione di riviste da considerare è avvenuto utilizzando le liste di due società scientifiche (Società Italiana di Statistica, SIS e Associazione per la Matematica Applicata alle Scienze Economiche e Sociali, AMASES). L'unione delle 4 liste ha dato luogo a un insieme finale di 2786 titoli. L'aggiunta di ulteriori liste non ha portato a incrementi degni di nota al numero di titoli classificati dalle 4 organizzazioni. Il Dipartimento di Scienze Statistiche di Bologna ha classificato 64 ulteriori titoli di riviste, su 312 che hanno visto pubblicazioni degli aderenti al Dipartimento dal 2004 al 2012.

Non tutti gli studiosi sono fiduciosi riguardo alle indicazioni delle società scientifiche; tra le riserve che vengono addotte, c'è il timore che queste indicazioni nascondano, e nemmeno troppo velatamente, atteggiamenti opportunistici di tipo protezionistico da parte delle società. Nel Dipartimento a cui appartengo, si è ritenuto che le società scientifiche che hanno prodotto classificazioni di riviste godano di una buona reputazione e abbiano un atteggiamento propositivo e collaborativo riguardo alla valutazione.

L'aumento delle liste di riferimento, e quindi l'aumento del numero di riviste i cui lavori possono essere classificati in modo quasi automatico, riduce il numero di casi nuovi che devono essere affrontati individualmente, ma porta ad un aumento del numero delle discordanze nelle classificazioni. Tali discordanze devono essere risolte. La commissione ricerca del dipartimento di scienze statistiche ha scelto di non contestare il risultato della classificazione ministeriale (VQR-GEV 13 e classificazione ulteriore nella classe A da parte di ANVUR ai fini dell'Abilitazione Scientifica Nazionale, ASN), ma di calcolare un ulteriore indice nei casi in cui la classificazione delle società scientifiche fosse migliore di quella del ministero e di utilizzare tale indice per le riviste non considerate dal ministero, anche se prese in esame dalle società scientifiche. In questo momento non è il caso di dare risalto alle discrepanze nella classificazione tra quanto concluso dal ministero e l'indicazione fornita dall'indice. L'indice del Dipartimento di Scienze Statistiche di Bologna è basato su indici di tipo bibliometrico, e normalizza rispetto alle categorie WOS o Scopus (si veda Cocchi et al., 2014 per maggiori dettagli). Per dare un'idea dell'estensione degli esiti scientifici dei contributi degli statistici si noti che i titoli delle 4 liste considerate toccano 200 delle 273 categorie WOS.

Questo sforzo ulteriore ha permesso di classificare molti più lavori degli aderenti al Dipartimento di Scienze Statistiche dell'università di Bologna rispetto alla copertura delle liste prodotte dal ministero. Dei 64 lavori pubblicati nei 7 anni di riferimento in riviste non appartenenti all'insieme individuato dalle 4 organizzazioni, 39 ulteriori sono stati classificati tramite l'indice. I titoli rimanenti sono stati alla fine egualmente classificati con una procedura che identificava il vicino più simile.

Ho quindi descritto sommariamente una proposta praticabile per la integrazione di esigenze nuove o locali nell'alveo di un impianto consolidato e centralizzato. Si tratta soltanto di un esempio, che può essere migliorato o reindirizzato.

Il lavoro svolto nel 2012-2013 della commissione ricerca del dipartimento di scienze statistiche dell'università di Bologna e che ho appena schematizzato appare utile per il lavoro che il panel dell'area 13 dell'Università di Bologna sta svolgendo per la commissione VRA.

Due aspetti ulteriori sono molto importanti: a) la definizione di quanta parte di un articolo sia da considerare equivalente ad un contributo a nome singolo (che poi sconfina anche nella questione della attribuzione della proprietà intellettuale dei lavori a seconda di quali siano le istituzioni in cui gli autori lavorano); b) la costruzione di un sistema di riferimento che sia dinamico ma che non disorienti le scelte di ricerca dei singoli, né li spinga a comportamenti opportunistici.

Comunque venga considerata, la questione del punto a) rischia di essere risolta in modo da non favorire la collaborazione tra i gruppi e la interdisciplinarietà, che per gli statistici dovrebbe invece essere un aspetto cardine. La questione del punto b) si collega a sua volta alla stabilità e alla tempistica di rinnovo delle liste di riviste considerate per le classificazioni.

Gli studiosi della disciplina statistica devono anche prestare attenzione ai paradigmi e alle classificazioni che si stanno consolidando a livello internazionale, indipendentemente dalla adesione ai settori scientifico-disciplinari e alle aree CUN dell'accademia italiana. Prendendo come esempio le parole chiave dell'European Research Council (ERC, 2011), le prospettive di ricerca europea in prima approssimazione si configurano in modo non positivo, perché il termine statistica è confinato in pochi sotto-settori. Ma l'assetto configurato nell'ambito dell'ERC deve invece essere percepito come incoraggiante, perché i temi di ricerca peculiari della statistica possono svilupparsi in ambito molto esteso.

Ragionando in modo analogo a quello che porta

le discipline statistiche ad essere riferite all'Area 13, la statistica può essere associata direttamente al sottosettore SH1_4: Econometrics, statistical methods, che è uno dei 14 casi di SH1: Individuals, institutions and markets: economics, finance and management, a sua volta uno dei 6 sottogruppi di SH: Social Sciences and Humanities. La demografia rientra nel terzo sottogruppo SH3: Environment, space and population: environmental studies, demography, social geography, urban and regional studies. Ma la statistica è menzionata esplicitamente anche nel macrosettore PE: Physical Sciences and Engineering, sia nel settore PE1: Mathematics: all areas of mathematics, pure and applied, plus mathematical foundations of computer science, mathematical physics and statistics (in particolare PE1_13 Probability and PE1_14 Statistics) sia nel settore PE6: Computer science and informatics: informatics and information systems, computer science, scientific computing, intelligent systems, in cui appare il sottosettore PE6_11 Machine learning, statistical data processing and applications using signal processing (eg. Speech, image, video).

Ma anche nel macrosettore Life Sciences LS, ci sono ambiti in cui gli statistici sono attivi, si tratta dei settori LS2: Genetics, genomics, bioinformatics and systems biology: genetics, population genetics, molecular genetics, genomics, transcriptomics, proteomics, metabolomics, bioinformatics, computational biology, biostatistics, che nello

specifico è LS2_12 (biological modelling and simulation, system biology, genetic epidemiology) e LS7 (diagnostic tools, therapies and public health: aetiology, diagnosis and treatment of diseases, public health and epidemiology, che sono in sintesi LS7_9, LS7_10 e LS7_11: pharmacology, clinical medicine, regenerative medicine, medical ethics).

La collocazione della ricerca scientifica nelle 14 aree CUN ha quindi bisogno di un confronto e una armonizzazione con la situazione internazionale. La classificazione ERC, seppure discutibile, seppure in evoluzione, esiste e guida l'orientamento della Ricerca scientifica Europea, verso cui vogliamo e dobbiamo tendere, sia per i finanziamenti, sia per la direzione dell'impegno di studio.

Anche il riscontro di come i nostri settori scientifico-disciplinari (SSD) possano essere associati ai settori ERC dovrebbe essere effettuato; a mio avviso si tratta di uno dei passi da affrontare al più presto.

Riferimenti bibliografici

Cocchi D., Cavaliere G., Freo M., Giannerini S., Mazzocchi M., Trivisano C., Viroli C., 2014, *A support for classifying scientific papers in a University Department*, Procedia Economics and Finance, SciVerse Science Direct, Elsevier

ERC, 2011, *ERC Grant schemes. Guide for Applicants for the Synergy Grant 2012*.