

DAL MONDO DELLA STATISTICA APPLICATA

Nel 1995 è stata pubblicata su questa Rivista una mia recensione al volume di M.J. BAXTER *Exploratory Multivariate Analysis in Archaeology* (Edinburgh 1994, Edinburgh University Press), in cui l'Autore si soffermava in particolare sull'uso delle tecniche multidimensionali in archeologia e dichiarava apertamente di farlo secondo un'ottica specificamente applicativa, piuttosto che teorica, come ben si addice d'altronde a uno statistico che ha a lungo e attivamente operato in questo settore disciplinare. Un nuovo libro dello stesso Autore (*Statistics in Archaeology*, London 2003, Arnold) costituisce, a quasi dieci anni di distanza, l'appuntamento più importante del 2003 nel settore delle applicazioni statistiche in archeologia.

Come tiene a ribadire lo stesso Baxter, non si tratta di un testo introduttivo alla materia, quanto piuttosto di una guida a un approccio applicato, anche se non estremamente tecnico, inteso a porre in evidenza, tramite scelte selettive, aree di ricerca in cui le applicazioni sono più sporadiche, meno note e soprattutto meno citate nei riferimenti bibliografici. Il volume, composto da 21 capitoli e corredato da un ricco apparato bibliografico, può essere diviso orientativamente in tre parti, anche se esse risultano strettamente connesse l'una all'altra: la prima (cap. 2) descrive sia alcuni set di dati ampiamente utilizzati nel testo sia i problemi che essi pongono, in modo da far calare le diverse applicazioni nel loro specifico contesto archeologico; la seconda (capp. 3-14) è dedicata ad argomenti di carattere statistico, attraverso un esame di numerosi metodi di indagine e in particolare delle tecniche di analisi multidimensionale dei dati – non solo le più note (Analisi dei Componenti Principali, Cluster Analysis, Analisi Discriminante, Analisi delle Corrispondenze), ma anche quelle meno diffuse – dei metodi bayesiani e dell'analisi spaziale; la terza (capp. 15-21) è infine dedicata a specifici problemi archeologici, che hanno maggiormente catturato l'attenzione della statistica: le datazioni assolute e relative; la ricostruzione quantitativa delle testimonianze materiali; l'analisi isotopica e altre ancora.

L'impianto del volume segue l'ottica generale della collana (Arnold Applications of Statistics Series) che non ha uno scopo d'introduzione alla statistica applicata, bensì vuole consentire agli autori, e di conseguenza ai lettori, di concentrarsi sulle tecniche più avanzate e di maggiore rilevanza nei singoli campi di applicazione trattati. Come afferma lo stesso Baxter «the focus is on applications at an intermediate or advanced level». L'impostazione, dunque, cambia radicalmente rispetto al recente e fondamentale manuale di C. ORTON *Sampling in Archaeology* (Cambridge 2000, Cambridge University Press). Proprio in una recensione a quest'ultimo volume («Archeologia e Calcolatori», 11, 2000, 396-400) dicevamo, a proposito del tipo di atteggiamento utilizzato per proporre la materia, che Orton «relega le formule matema-

tiche in un'appendice a sé stante, ritenendo in tal modo di facilitare agli archeologi il compito della lettura e della comprensione». Baxter, invece, si rivolge direttamente con un linguaggio tecnico ai lettori, che possono essere sia archeologi interessati alle potenzialità offerte dalle applicazioni statistiche, sia statistici pronti a verificare in che modo la loro disciplina trova applicazione nella ricostruzione del passato.

E questo nuovo atteggiamento fiducioso nel rapporto tra archeologia e statistica trova conferma anche nella maggiore varietà tematica e nella vivacità che si va via via notando nei contributi inseriti nelle sezioni dedicate ai metodi quantitativi nell'ambito degli appuntamenti annuali dei convegni CAA. È interessante, ad esempio, confrontare anche i titoli attribuiti a tali sezioni: dal più vago e generale "Artefact, Analyses and Classification" dell'edizione del 2000 (Z. STANČIČ, T. VELJANOVSKI (eds.), *Computing Archaeology for Understanding the Past. CAA 2000*, BAR International Series 931, Oxford 2001), al più tecnico ma onnicomprensivo "Papers in the Database Applications and Statistics and Quantitative Methods" del 2001 (G. BURENHULT, *Archaeological Informatics: Pushing the Envelope. CAA 2001*, BAR International Series 1016, Oxford 2002), al più teorico "Reasoning, Statistical and Analytical Techniques" del 2002 (M. DOERR, A. SARRIS (eds.), *The Digital Heritage of Archaeology. CAA2002*, Athens 2003, Hellenic Ministry of Culture) e, infine, al più semplice ma chiaro "Statistics and Quantitative Methods", almeno come compare nel programma dell'edizione viennese del 2003, ancora in corso di pubblicazione. A tali titoli si può accostare anche quello, assai specifico, "The Application of Bayesian Methods in Archaeology" di uno dei Colloqui organizzati dalla Commissione IV "Data Management and Quantitative Methods in Archaeology" nell'ambito del XIV Congresso Internazionale UISPP tenutosi a Liegi nel 2001 (F. DJINDJIAN, P. MOSCATI (eds.), «Archeologia e Calcolatori», 13, 2002).

Tornando al volume di Baxter, il capitolo introduttivo risulta particolarmente utile per ottenere una visione generale del rapporto che si è instaurato nel tempo tra statistica e archeologia, a partire dagli inizi degli anni Cinquanta. In questa panoramica, in cui non si perde l'occasione per presentare anche i protagonisti del dialogo tra le due discipline e i principali strumenti bibliografici di riferimento, sono interessanti in particolare gli aspetti teorici trattati, anche se brevemente e, come dice lo stesso Autore, attraverso «a very simplified picture»: da un lato un cenno al dibattito sorto nel settore statistico intorno ai due principali approcci, cioè quello dell'inferenza classica e quello dell'inferenza bayesiana; dall'altro lato l'analisi del rapporto tra le due discipline in relazione agli sviluppi del dibattito teorico in archeologia e in particolare alla contrapposizione tra l'approccio processuale e quello post-processuale. Non nuova, ma sempre incisiva, è a tale proposito la nota rivolta alle difficoltà subite nell'affermarsi dell'approccio statistico in archeologia

a causa dell'automatica convinzione che l'applicazione di tali tecniche comportasse anche l'adesione a un approccio alla scienza proprio del positivismo.

Un altro punto fermo nell'esposizione di Baxter è la distinzione, spesso operata nelle applicazioni archeologiche, tra tecniche "semplici" e tecniche "complesse" e in particolare l'erroneo assunto che le prime devono precedere le seconde nell'analisi di una problematica archeologica (tra i volumi dedicati in modo specifico all'applicazione in campo archeologico della statistica "elementare", nel senso delle tecniche univariate e bivariate di analisi grafica e descrittiva, cfr. ad esempio: M. FLETCHER, G.R. LOCK (eds.), *Digging Numbers. Elementary Statistics for Archaeologists*, Oxford University Committee for Archaeology Monograph 33, Oxford 1991 e la relativa recensione in «Archeologia e Calcolatori», 4, 1993, 357-361). Poiché ormai i software a disposizione rendono semplice, cioè facile da eseguire, un'ampia serie di procedimenti, secondo Baxter il confine tra semplicità o meno delle diverse tecniche, se è proprio da ricercare, potrebbe essere eventualmente trovato nella distinzione, che non viene però del tutto esplicitata, tra metodi esplorativi per l'analisi dei dati e metodi basati su modelli.

Questa distinzione tra semplicità e complessità, anche se piuttosto generica rispetto a una terminologia specialistica, ha richiamato l'attenzione anche di altri studiosi. Ad esempio, C. ORTON (in *CAA 2001*, cit., 495-499), nel suo intervento dedicato alle modalità di produzione della ceramica romana nel sito di Highgate Wood, presso Londra, individua due diversi tipi di modelli statistici in archeologia: quelli informali, spesso impliciti, e quindi relativamente semplici, ovvero quelli formali e complessi, che egli lega all'uso di software specializzato, come nel caso della statistica bayesiana. Nell'indagine condotta, Orton fa uso di un semplice modello stocastico, basato sull'attribuzione di diversi valori alle variabili.

Come confermato anche da altri recenti interventi, il dibattito sulle applicazioni statistiche all'archeologia appare incentrato soprattutto sulle questioni relative alla fase iniziale della ricerca, cioè alla scelta dei dati da analizzare. Interessante a questo proposito è la posizione di A. BIETTI (in AA.VV., *I modelli nella ricerca archeologica. Il ruolo dell'informatica*, Roma 2003, Accademia Nazionale dei Lincei) sui limiti e i vincoli richiesti dall'applicazione delle diverse tecniche statistiche, oggi alla facile portata degli archeologi attraverso package che possono essere implementati su qualunque PC, ma che spesso vengono utilizzati «a scatola chiusa». Lo studioso, infatti, sostiene che, laddove subentri l'uso di tecniche inferenziali, oltre all'ovvia necessità di conoscere almeno i fondamenti basilari di probabilità e statistica, siano indispensabili ipotesi stringenti per effettuare stime sulla distribuzione delle popolazioni e sull'attendibilità dei campioni.

A tal fine è opportuno che le tecniche inferenziali vengano utilizzate solo dopo l'applicazione di una serie di tecniche esplorative, come ad esem-

pio la Cluster Analysis, l'Analisi dei Componenti Principali e l'Analisi delle Corrispondenze, e sulla base dell'acquisizione di risultati compatibili. L'analisi inferenziale assume così un compito di conferma dei risultati ottenuti dall'analisi esplorativa, purché sia effettuata su campioni indipendenti e rispondenti a criteri di rappresentatività e casualità rispetto alla popolazione di riferimento (S. CAMIZ, *From data collection to mathematical models: methodological pathways*, in *Scientific Background and Practical Tools for Ecological Expertise. 17th ISPM School*, Tbilisi 2001, in corso di stampa). Ciò risulta tanto più importante in una disciplina come l'archeologia in cui, a differenza di quanto avviene nelle scienze esatte, i dati provengono in genere da operazioni di scavo e quindi le ipotesi non possono essere verificate attraverso esperimenti ripetibili.

A questo dibattito sulle fasi iniziali della ricerca e sul processo decisionale, con tutti i dubbi e i problemi che esso implica, si legano l'adozione di un approccio bayesiano all'inferenza statistica, che ricordiamo ha riscontrato un successo in archeologia soprattutto a partire dagli anni Novanta (cfr. in particolare C.E. BUCK, W.G. CAVANAGH, C.D. LITTON, *Bayesian Approach to Interpreting Archaeological Data*, Chichester 1996, Wiley, e la relativa recensione in «Archeologia e Calcolatori», 8, 1997, 187-190) e l'applicazione della logica fuzzy (cfr. ad esempio S. HERMON, F. NICCOLUCCI, in questo volume, con bibliografia precedente), formalizzata nel 1965 da L.A. Zadeh, che non si muove più nell'ambito di una logica tradizionale bivalente (sì/no, vero/falso, 0/1), bensì si basa sul valore o grado di appartenenza di un elemento a un dato insieme, con una transizione non più brusca ma graduale tra appartenenza e non appartenenza. Ambedue questi approcci, benché non strettamente connessi, offrono secondo G.L. COWGILL «a nuanced way of dealing with uncertainty» (in DJINDJIAN, MOSCATI, cit., 189).

Veniamo, infine, alle problematiche archeologiche che in questi ultimi anni sono state con maggiore frequenza oggetto di studio e ad alcune evoluzioni nelle tecniche utilizzate per analizzarle statisticamente. Anche se in numero quantitativamente non rilevante, sussistono sempre applicazioni che potremmo definire tradizionali, rivolte cioè alla classificazione tipologica dei materiali e allo studio della distribuzione spaziale dei manufatti, con un certo privilegio per le classi ceramiche. Lo studio della ceramica, d'altronde, costituisce un punto di riferimento anche per altri tipi di applicazioni, connessi soprattutto con l'uso di tecniche statistiche per l'elaborazione dei risultati provenienti dalle analisi archeometriche. In quest'ultimo caso, è interessante ad esempio la proposta di analizzare contemporaneamente diversi tipi di dati, sia geochimici sia mineralogici (C.C. BEARDAH *et al.*, in *CAA 2001*, cit., 441-445 e in *CAA 2002*, cit., 261-265).

Anche il settore della statistica applicata all'archeologia non è rimasto avulso dall'evoluzione delle tecniche di image processing, in fase sia di acqui-

sizione delle informazioni, ad esempio tramite laser scanner, sia di elaborazione (image analysis). Mi riferisco in particolare alla comparazione delle immagini e all'analisi delle forme (shape analysis) per giungere a una classificazione del campione preso in esame. A tale proposito possiamo citare due esempi: l'analisi e il confronto di immagini microscopiche di campioni relativi a strumenti litici e lignei, attraverso una descrizione quantitativa e non qualitativa della tessitura e un confronto tra i diversi valori della luminanza (M. ADÁN *et al.*, in *CAA 2002*, cit., 253-259) e lo studio tipologico di 200 scarabei mediante un'analisi automatizzata delle loro caratteristiche formali (S. ANDRENUCCI, M. ANDRENUCCI, in AA.VV., *Proceedings of the 14th Round Table Computer Aided Egyptology*, Pisa 2002, Università di Pisa - Consorzio Pisa Ricerche).

PAOLA MOSCATI