

ACCADEMIA NAZIONALE DEI LINCEI

ANNO CCCLXXIX - 1982

CONTRIBUTI DEL
CENTRO LINCEO INTERDISCIPLINARE
DI SCIENZE MATEMATICHE E LORO APPLICAZIONI
N. 61

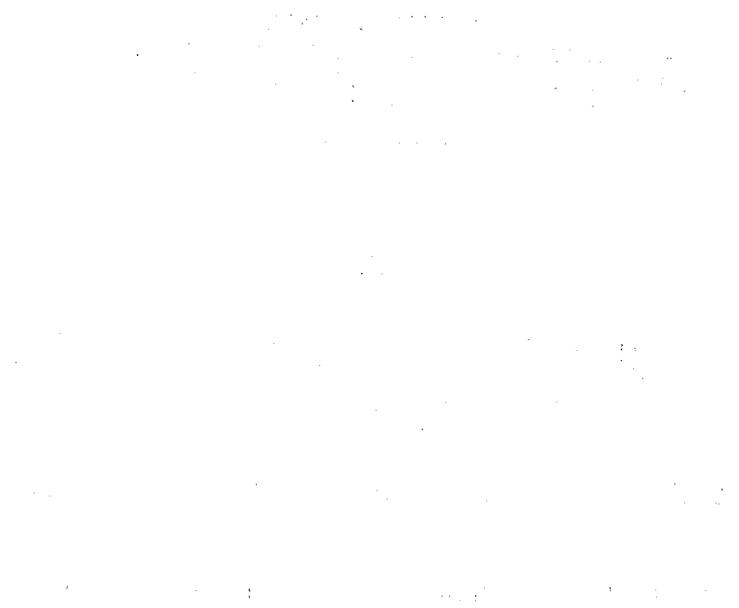
AMILCARE BIETTI

TECNICHE MATEMATICHE
NELL'ANALISI
DEI DATI ARCHEOLOGICI

(Ciclo di tre conferenze tenuto nel dicembre 1980)



ROMA
ACCADEMIA NAZIONALE DEI LINCEI
1982



Introduzione

In questa serie di lezioni cercherò di dare una panoramica generale dell'uso dei metodi matematici e statistici nell'analisi dei dati archeologici. In questo senso si tratta di uno sviluppo naturale della conferenza già tenuta presso l'Accademia nel giugno 1978 (Bietti, 1979).

Anche se cercherò di fornire una prospettiva più articolata, vi saranno in queste lezioni più o meno le stesse limitazioni di quella conferenza, ed in particolare il tempo limitato a disposizione (basti pensare al fatto che su questi argomenti J.E. Doran e F.R. Hodson (1975) hanno dedicato un intero libro!) ed in più la circostanza che il mio campo di ricerca è quello dell'archeologia preistorica del Paleolitico e del Mesolitico, e che pertanto gli esempi forniti riguarderanno in gran parte quelle problematiche.

Inoltre anche qui, l'aspetto di carattere più "informativo" e cioè quello collegato con il problema delle "banche dei dati" archeologici o più in generale di interesse culturale o artistico non viene affrontato: si tratta di ben altro campo di indagine, mentre queste lezioni riguardano solo le metodologie matematiche (o meglio solo alcune di esse) che si impiegano nell'*analisi* dei dati archeologici.

A questo scopo, come è stato già più volte sottolineato, l'impiego dell'elaboratore elettronico è quasi d'obbligo, non perché abbia in sé proprietà miracolose, ma solo perché permette di risparmiare una notevole quantità di tempo, specialmente quando la quantità dei dati da analizzare diventa particolarmente elevata.

Anche qui, come nella conferenza del 1978, le lezioni saranno suddivise - almeno in prima approssimazione - in tre livelli più o meno tradizionali: il primo riguarda il problema

della classificazione di oggetti provenienti dagli scavi, mediante tecniche abbastanza elementari, ed in particolare la ricognizione dei *tipi* (litici, ceramici, ecc.) a partire da attributi *metrici* (e cioè misurabili, come larghezza, lunghezza, ecc.) o *nominali* (presenza o assenza di caratteristiche qualitative come tipo di ritocco, di decorazione, ecc.).

Il secondo livello sposta l'attenzione allo studio comparativo dei complessi industriali, non solo mediante tecniche più o meno qualitative, ma tramite metodi più raffinati che impiegano le tecniche statistiche di analisi multivariata.

Il terzo livello dovrebbe essere più di interesse teorico e interpretativo globale: ricognizione di "strutture" e processi culturali sia nello spazio che nel tempo, ed in particolare la formulazione di *modelli* matematici.

In queste lezioni però, proprio per il loro carattere più spiccatamente didattico, porrò principalmente l'accento sui metodi matematici, procedendo da quelli più semplici verso quelli più complicati. Accadrà così che alcuni problemi di tipologia, che tradizionalmente rientrerebbero nel primo livello di analisi, verranno descritti nella seconda sezione dato che necessitano di statistica multivariata. D'altra parte nella terza sezione, ad esempio, oltre ad accennare a tecniche di ricognizione di "clustering" spaziale dei dati dello scavo, relativamente complesse, come la "nearest neighbour analysis", si farà riferimento, per lo stesso genere di problemi, anche a metodi più semplici, come l'analisi della varianza o il test del χ^2 .

Spero quindi che il lettore possa capire come ad ognuno dei livelli suesposti esistano diversi gradi di "matematizzazione" e, a seconda del tipo particolare di problema da trattare, sarà l'archeologo stesso, in base al proprio buon senso ed alla conoscenza dei materiali, a scegliere il metodo matematico opportuno (o eventualmente a non impiegarlo affatto, se non sarà necessario!).

La bibliografia citata è sostanzialmente archeologica. Per gli argomenti più strettamente matematici o statistici ci si riferisce a testi standard. Per citarne solo qualcuno, dai più semplici ai più complessi, si possono consultare, ad esempio Young (1962), Spiegel (1961), il classico Fisher (1950), Cramèr (1946), Yule e Kendall (1953), Blalock (1960), almeno per quanto riguarda le statistiche elementari. In ogni caso un'ampia bibliografia è presente nel già citato volume di J.E. Doran e F.R. Hodson (1975).

I - Metodi basilari di classificazione dei dati archeologici mediante statistiche elementari

In questo primo livello figurano per esempio tutte le analisi elementari di attributi allo scopo di determinare un *tipo*, inteso come entità progettuale e prodotto finito derivante dalla correlazione intenzionale dei diversi attributi (Bordes, 1967), che possono essere a buon diritto considerati come le unità elementari ed "irriducibili" per la costruzione dei tipi.

Questi attributi si dividono convenzionalmente in due categorie: *a*) attributi *metrici* (cioè misurabili quantitativamente), come ad esempio la lunghezza, la larghezza o lo spessore di uno strumento di selce, il diametro dell'orlo di un vaso di ceramica, percentuali dei vari elementi chimici in un oggetto di bronzo, ecc., e *b*) attributi *nominali*, come ad esempio il tipo di ritocco su di uno strumento di selce, o il tipo di decorazione su uno strumento ceramico, il tipo di ansa, ecc. e che pertanto non sono misurabili, o meglio tali che una loro definizione quantitativa comporta delle imprecisioni notevoli (v. per esempio la misura dell'angolazione di un ritocco su uno strumento litico), certamente maggiori od eguali a quella dovuta alla variabilità naturale dell'artefice preistorico.

Fino a pochi anni fa, come è noto, la ricognizione dei tipi avveniva in modo essenzialmente intuitivo: l'esame di centinaia o migliaia di pezzi da parte dei ricercatori di lunga esperienza ha portato ad individuare tipi ben distinti, almeno in grandi linee, proprio perché il cervello umano ha in sé la capacità di riconoscere spesso differenze e somiglianze, anche se dipendenti da diversi parametri, almeno nei casi più vistosi.

Abbiamo così, per esempio, le celebri liste di tipi della scuola di Bordeaux per il Paleolitico inferiore e medio (Bordes,

1961) e per il Paleolitico superiore (Sonneville-Bordes e Perrot, 1954-56).

Per periodi posteriori bastera' citare, ad esempio, la tipologia dei corredi delle necropoli laziali dell'Eta' del Ferro di H. Muller-Karpe (1962).

Se si vogliono eseguire analisi quantitative, per definire meglio l'intenzionalita' dell'artefice preistorico nei riguardi di certe quantita', come ad esempio le dimensioni di strumenti litici o di oggetti ceramici, i metodi piu' semplici e piu' elementari, per gli attributi metrici, sono ovviamente quelli derivati dalla statistica tradizionale in una variabile.

Se, per esempio, da una prima analisi qualitativa, c'e' ragione di credere che un tipo particolare di strumento litico, in uno o piu' giacimenti, abbia una lunghezza predeterminata, e cioe' che viene costruito secondo uno "standard" di lunghezza prescelto, si puo' disegnare un istogramma di frequenza, su un totale di N strumenti, ad esempio, e cioe' un grafico in cui venga rappresentato il numero di oggetti al variare della lunghezza.

Si calcolano quindi, sulla base del grafico, quantita' che danno indicazioni sulla *tendenza centrale* e sulla *dispersione* della distribuzione di oggetti in esame.

Le misure piu' comuni di tendenza centrale sono la media definita come

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (1.1)$$

(nel nostro caso gli X_i sono le lunghezze degli N strumenti litici considerati) e la moda, che e' definita come il valore X_M per cui si ha il numero massimo di oggetti (il valore di X di massima frequenza). Naturalmente una distribuzione, anche se ha una sola media, puo' avere piu' mode e cioe' piu' picchi di frequenza, che in linea di principio potrebbero corrispondere a piu' scelte intenzionali della grandezza considerata da parte

dell'artefice.

Bisogna comunque osservare che la moda è strettamente dipendente dal "passo" dell'istogramma, e cioè dall'ampiezza di ogni singolo intervallo di lunghezza (nel nostro caso). Infatti, a titolo di esempio, possiamo considerare gli istogrammi delle larghezze delle lame rinvenute nei tagli 4-5 del giacimento Epigravettiano di Taurisano (Lecce) datati a circa 15.500 anni da oggi (Bietti, 1979a). In fig.1 abbiamo due di questi istogrammi ricavati tramite un programma (TIPOMT) impiegato sull'elaboratore UNIVAC 1100/22 dell'Università di Roma, in collaborazione con la dr.ssa L. Zanello dell'Istituto di Fisica dell'Università di Roma.

Come si può vedere, il primo istogramma, con passo 1 mm, sembra mostrare una distribuzione bimodale e cioè con due picchi di frequenza rispettivamente intorno a 6 mm e a 10 mm. Se si aumenta il passo a 2 mm (secondo istogramma) la curva diventa invece sostanzialmente unimodale, con un picco intorno a 7 mm, non molto diverso dalla media della distribuzione, che risulta essere 7,8 mm.

Siccome non esistono ragioni particolari per pensare che i cacciatori dell'Epigravettiano di Taurisano riuscissero a staccare dai nuclei di selce lamelle con la precisione di un millimetro, l'unica conclusione ovvia è che in grande maggioranza (istogramma 2) venivano prodotte lamelle di larghezza piuttosto piccola (il limite convenzionale di larghezza tra lame e lamelle è generalmente fissato a 12 mm (v. Tixier, 1963)).

La misura di dispersione più comune è la *deviazione standard* o scarto quadratico medio, definito da

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\bar{X} - X_i)^2} \quad (1.2)$$

dove \bar{X} è la media sopra definita. L'importanza di questa quantità è data dal fatto che per una distribuzione normale, dove

cioè la moda coincide con la media (la Gaussiana, che è appunto una curva a "campana"), il 68,3% degli oggetti è compreso entro una deviazione standard (a destra e a sinistra) dalla media, e il 95,4% entro due deviazioni standard. La variabilità del campione in esame è quindi tanto più piccola quanto più piccolo è il valore di σ rispetto ad \bar{X} .

Nel caso della fig.1 ad esempio, si ha $\sigma = 3,2$ mm e quindi una variabilità piuttosto ampia, dato che la media è di 7,8 mm (la moda è circa 7 mm come abbiamo già detto, così che la distribuzione è approssimativamente normale), risultato perfettamente ragionevole in un contesto del Paleolitico superiore come quello di Taurisano.

Se vi fosse stata una effettiva bimodalità nel senso, per così dire, che si aveva un picco a 7 mm ed un altro picco (anche se non così alto), a 2 cm (per esempio) si poteva ricorrere a "test" di significanza nella separazione dei due picchi semplicemente separando la distribuzione in due distribuzioni centrate intorno alle due mode, calcolando quindi due medie e due deviazioni standard e pertanto stabilire di che entità era la sovrapposizione in area tra le due distribuzioni. Se tale sovrapposizione fosse minore o eguale a due deviazioni standard (da una parte e dall'altra) potremmo a buon diritto dire che almeno al 95% di "livello di confidenza" c'è una chiara indicazione di due distinte larghezze nella fabbricazione delle lame. Torneremo più tardi su questo problema del confronto tra due distribuzioni.

Un altro esempio interessante è quello della distribuzione dei pesi degli oggetti di bronzo, e dei loro frammenti rinvenuti nei ripostigli dell'Eta' del Ferro italiana. A questo proposito è stata sollevata l'ipotesi (Peroni, 1966) di una struttura pre-monetale per le asce e per le fibule ad esempio, sia per i pezzi interi che per i loro frammenti (frammentazione intenzionale). In quest'ottica un istogramma dei pesi dovrebbe mostrare un andamento chiaramente multimodale, in corrispondenza

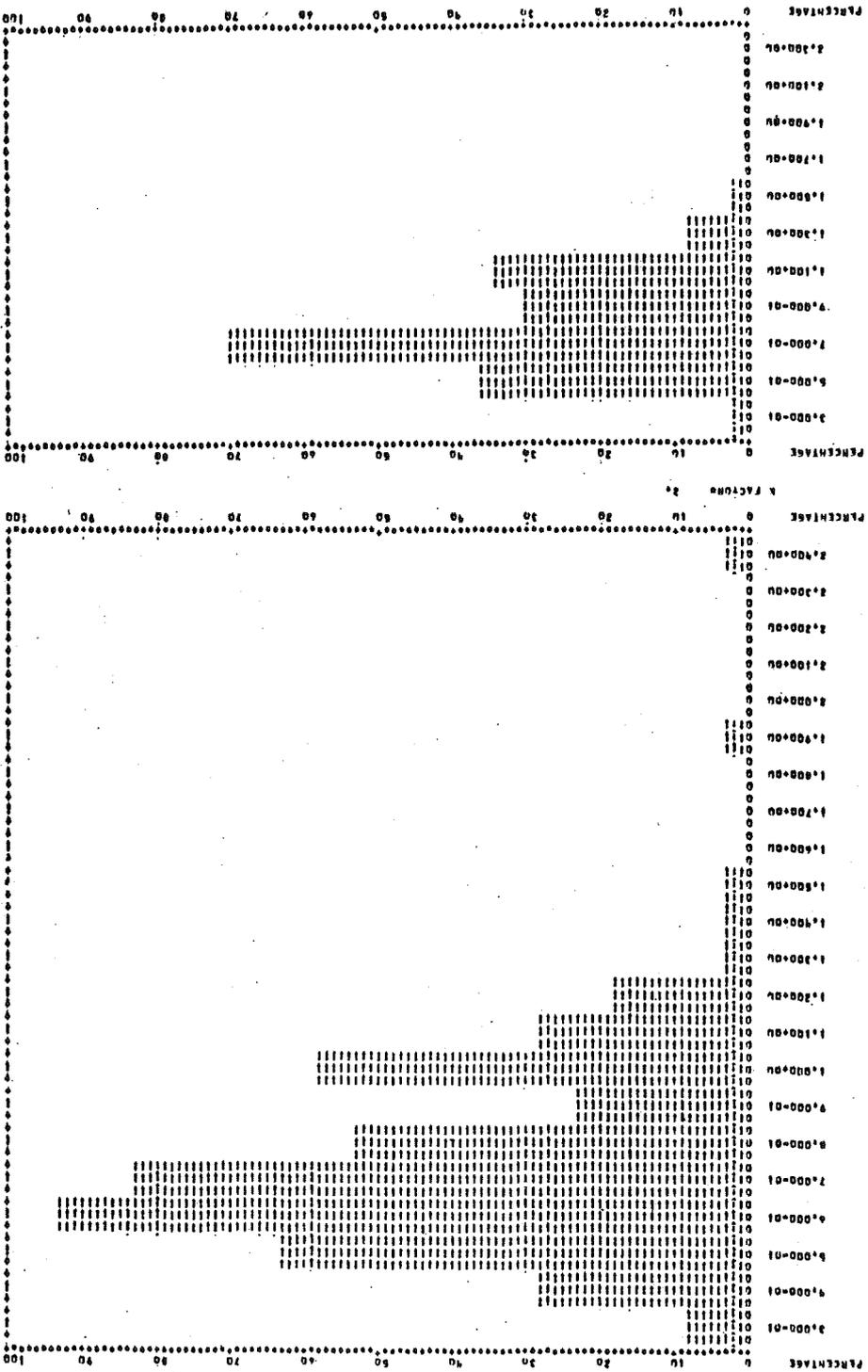


Fig. 1

Istogrammi delle larghezze delle lame e lamelle dei tl. 4-5 del giacimento Epigravettiano di Taurisano (Lecce).

cospicui del Villanoviano italiano, i cui pesi, a partire da circa 120 g fino a circa 700 g, mi sono stati gentilmente messi a disposizione da parte della dr.ssa S. Sorda dell'Istituto Italiano di Numismatica (Sorda, 1975).

In fig.2 sono rappresentati gli istogrammi con passo rispettivamente di 5 e di 10 g (mi è sembrato che passi minori fossero fuori di discussione dal punto di vista della significanza statistica) ottenuti, mediante il "package" statistico HBOOK, dall'elaboratore UNIVAC 1100/22 dell'Università di Roma, con la collaborazione delle dr.sse A. Rambaldi e L. Zanello dell'Istituto di Fisica.

Come si vede, in entrambi gli istogrammi non si può vedere alcuna "scansione regolare" nei picchi della distribuzione che è certamente multimodale. Il picco a 170-180 gr, ad esempio nel grafico con passo 10 gr, non trova nessun multiplo effettivo per pesi successivi: le mode massime sono infatti a gr 410, 440 e 480. E' quindi difficile stabilire una ripartizione "numerale" dei pesi e bisogna anche notare che la statistica è bassa: il massimo è solo di 9 asce!

Un altro esempio dove è possibile invece vedere una distribuzione bimodale è lo studio delle percentuali dell'arsenico e dello stagno nei bronzi del vicino Oriente dalla II^a metà del IV millennio A.C. fino alla fine del III millennio, in corso di studio da parte di M. Frangipane, A. Palmieri dell'Istituto di Paleontologia dell'Università di Roma e di A.M. Palmieri del C.N.R.

In fig.3 vediamo l'istogramma (sempre ottenuto tramite il programma HBOOK già citato, a cura delle dr.sse A. Rambaldi e L. Zanello dell'Istituto di Fisica) della percentuale di arsenico nei bronzi del vicino Oriente del periodo antico e cioè corrispondente a quello che viene chiamato Tardo-Calcolitico e Bronzo Antico I^o (intorno alla seconda metà del IV^o millennio A.C.). Come si vede nella lega c'è un picco intorno all'1% di arsenico ed un altro chiaramente separato, intorno al 3%.

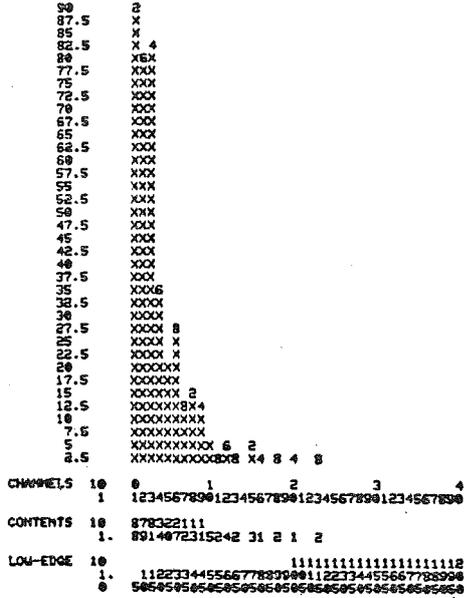


Fig. 4

Istogramma delle percentuali di arsenico nelle leghe dei bronzi del Bronzo Antico II e III del vicino Oriente.

ditori. E' interessante peraltro notare che le leghe con l'arsenico intorno all'1,5% e al 3% continuano a persistere in alcune zone, come si vede, secondo gli Autori, da istogrammi in cui gli oggetti con stagno sono esclusi, mentre nel Bronzo medio (dopo la fine del III° millennio), in seguito probabilmente alla diffusione su larga scala dello stagno in tutte le aree, le leghe con il solo arsenico scompaiono. Ulteriori risultati su questi problemi sono in corso di pubblicazione (Palmeri, 1980; Caneva et al., 1980).

Sempre nell'ambito delle analisi statistiche in una sola variabile e' interessante, ad esempio, il confronto fra grandezze di oggetti simili tipologicamente ma provenienti da giacimenti differenti, nello spirito di identificare piu' correttamente affinita' tipologiche fra questi oggetti.

Come esempio, si puo' considerare uno strumento particola-

re del Paleolitico superiore italiano, la punta a tacca o "à cran", quantitativamente importante nei giacimenti Epigravet-

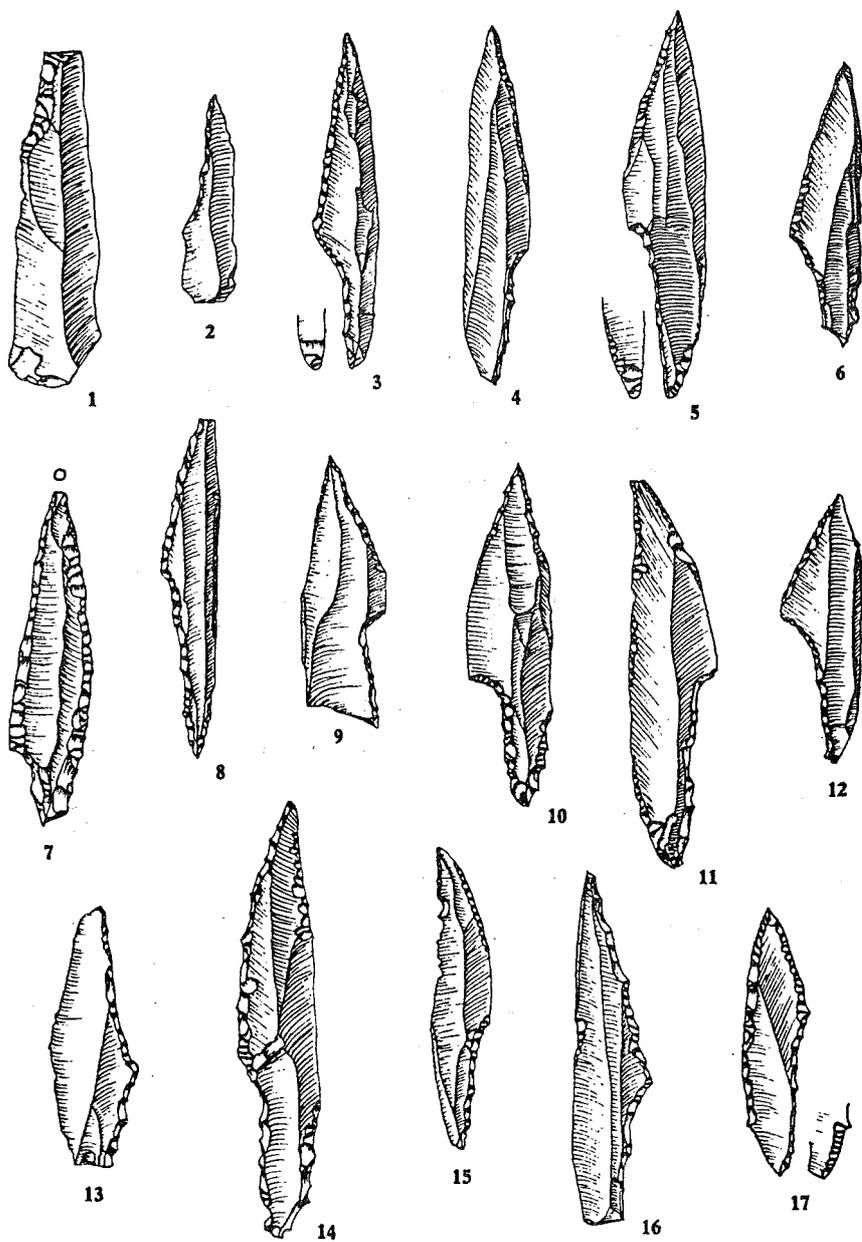


Fig. 5

Strumenti "à cran" dei tagli 17-10, di Grotta Paglicci (Foggia).
(da Mezzena e Palma di Cesnola, 1967) (5/6 gr. Nat.).

tiani dell'Italia sud-orientale, ed in particolare a Taurisano di cui abbiamo già parlato, e a Grotta Paglicci, nel Gargano (Mezzena e Palma di Cesnola, 1967). Gli strati di questo giacimento contenenti tali strumenti in percentuali significative sono più o meno contemporanei di quelli di Taurisano (strati 6-23, circa 16.000 anni da oggi) o leggermente più antichi (strati 17-10 di Paglicci, lo strato 18 è datato a circa 20.000 anni da oggi ed è attribuito al Gravettiano (v. Palma di Cesnola, 1975).

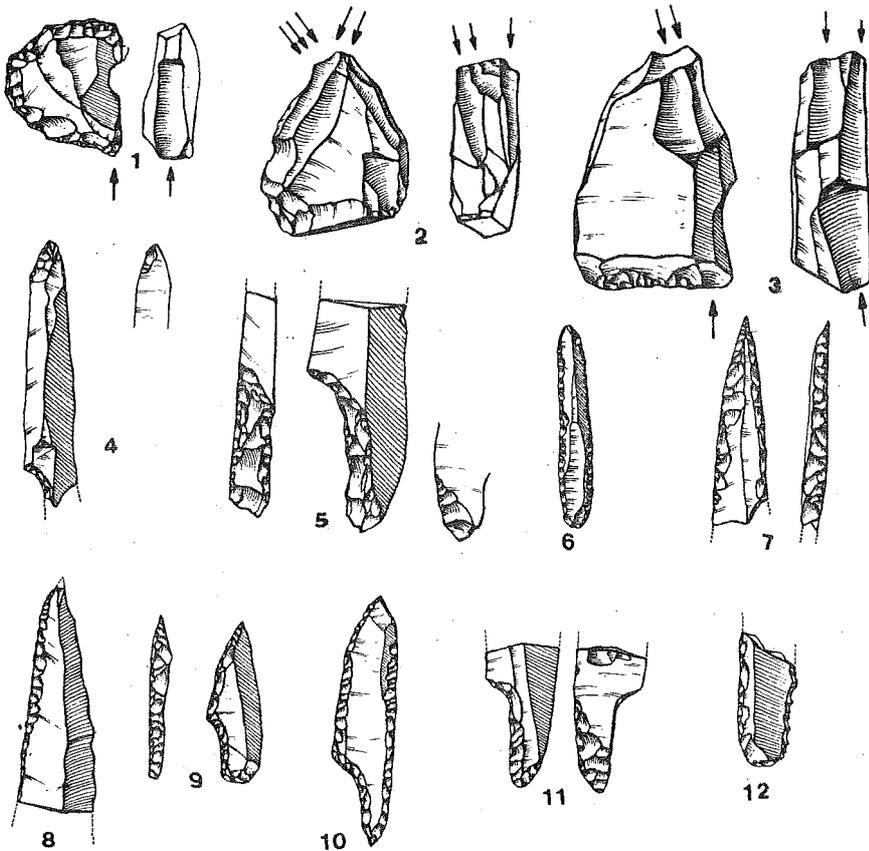


Fig.6

Industria litica dei tagli 9-7 del giacimento Epigravettiano di Taurisano (Lecce). Gli strumenti "à cran" sono i nn. 4, 5, 9, 10 e 11.

Ho potuto personalmente esaminare e misurare, grazie alla cortesia del prof. A. Palma di Cesnola nonché del prof. S. Ruffo, direttore del Museo di Storia Naturale di Verona, dove gran parte del materiale è conservata, quasi un centinaio di questi strumenti "à cran" di Grotta Paglicci, mentre quelli provenienti da Taurisano sono circa una trentina.

A parte la somiglianza che appare anche qualitativamente tra gli strumenti dei due giacimenti (v. per esempio figg. 5 e 6) è interessante osservare gli istogrammi delle larghezze mostrati in fig. 7 derivati dal programma TIPOMT già citato. A sinistra abbiamo quello degli strumenti "à cran" di Grotta Paglicci e a destra quello per Taurisano: le medie e le deviazioni standard sono rispettivamente 1,11 cm, 0.35 cm e 0.95 cm e 0.24 cm. Come si vede le larghezze di questi strumenti sono praticamente le stesse per entrambi i giacimenti entro una deviazione standard: si tratta di lamelle in entrambi i casi, risultato di grande interesse come intenzionalità di fabbricazione, visto che a Taurisano nel Salento la selce non è particolarmente abbondante e non si riviene in noduli di grandi dimensioni, mentre a Paglicci, nel Gargano, questo tipo di materia prima, com'è noto, è molto abbondante, con noduli di notevoli dimensioni.

È anche interessante osservare, in fig. 8, gli istogrammi per i due giacimenti della grandezza spessore/larghezza che dà un'idea della forma di questi strumenti. In questo caso la media è esattamente la stessa (0.33) mentre le deviazioni standard sono .08 per Paglicci e .09 per Taurisano.

Come vedremo anche successivamente, queste notevoli concordanze inducono a pensare che questo tipo di strumento fosse caratteristico di quel periodo nella zona medio e bassa adriatica ed inoltre, date le sue dimensioni ridotte (le lunghezze massime sono 6.4 cm per Paglicci e 3.8 cm per Taurisano), si può sostenere il loro uso come *punte di freccia*, piuttosto che di giavelotto e di conseguenza ipotizzare l'esistenza dell'ar-

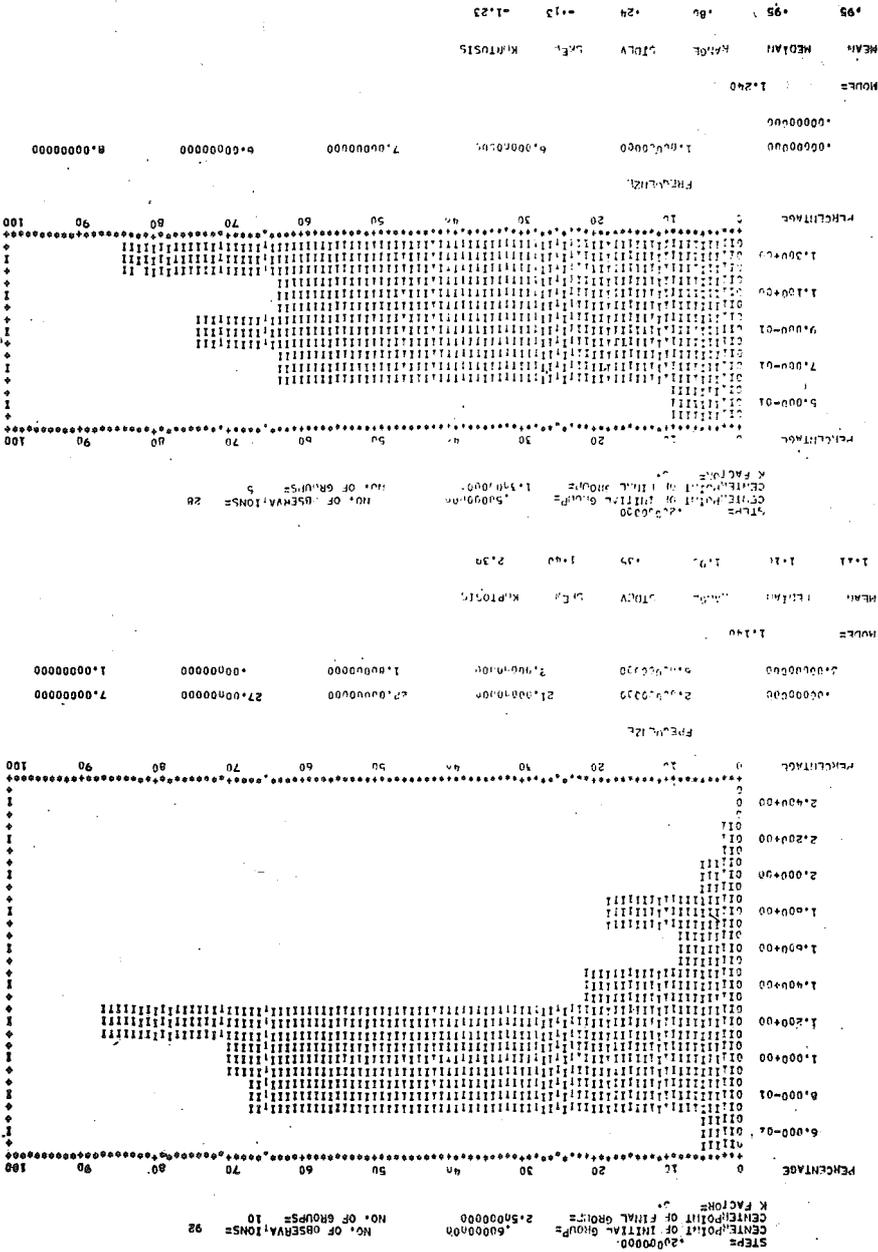


Fig. 7 Istogrammi delle larghezze degli strumenti "à cran" di Grotta Paglicci (a sinistra) e di Taurisano (a destra).

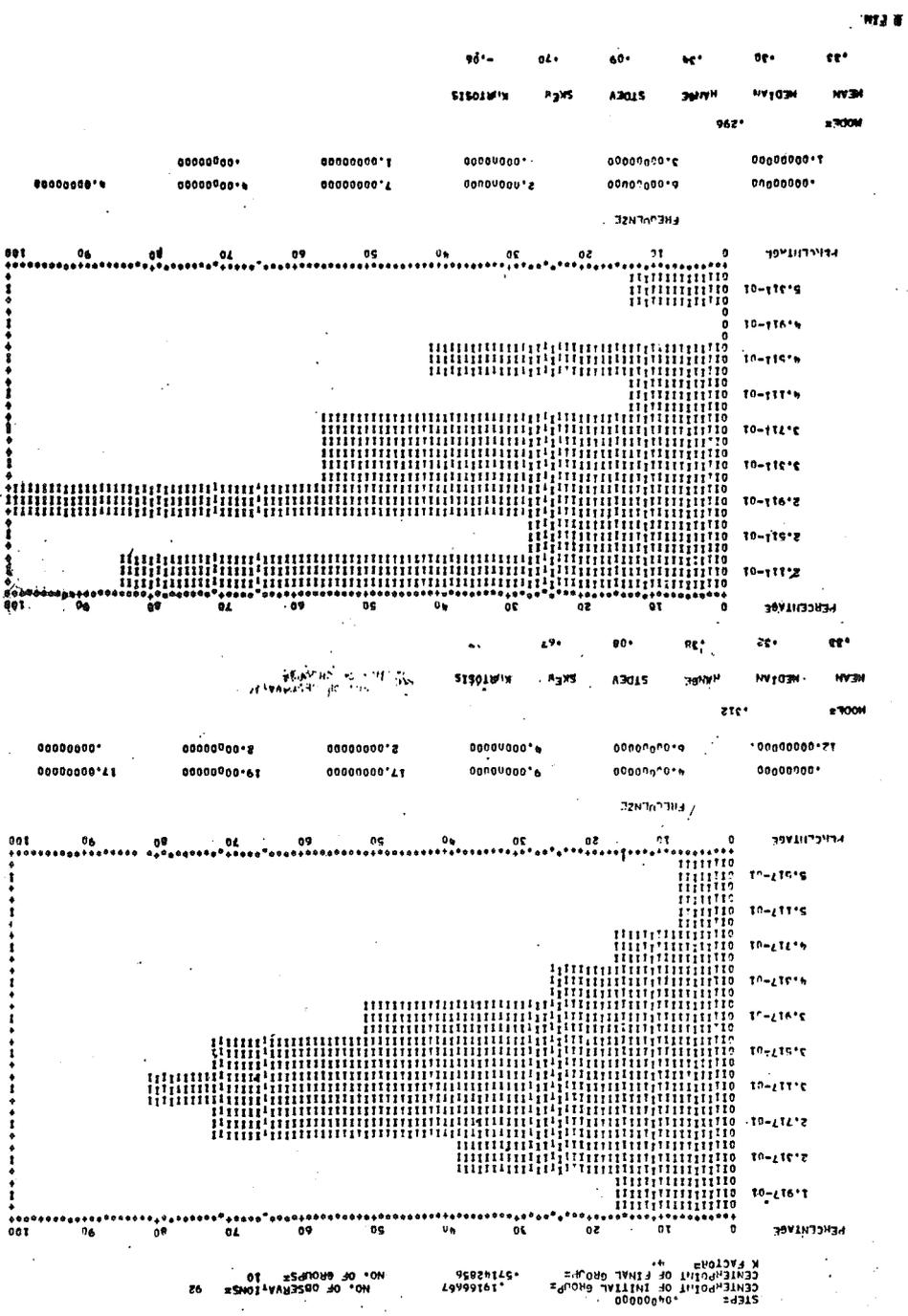


Fig. 8
Istogrammi del rapporto spessore/larghezza degli strumenti "a cran" di Grotta Paglicci (a sinistra) e di ~~Castelluccio~~ (a destra).

co, almeno in quella regione, in periodi abbastanza antichi.

Senza entrare in maggiori dettagli sui metodi statistici qualcuno potrebbe in ogni caso obiettare che nella comparazione tra gli strumenti di due giacimenti, ed in particolare tra le medie di alcune grandezze, si dovrebbe applicare piu' correttamente la teoria della "estimazione statistica" delle differenze o somiglianze tra due "popolazioni", delle quali i giacimenti considerati sarebbero dei "campioni". In questo caso, secondo la teoria, le deviazioni standard non sarebbero piu' σ ma σ/\sqrt{N} dove N e' la dimensione del campione.

A parte il fatto che un "test" statistico di questo genere applicato, ad esempio alla differenza delle medie della grandezza spessore/larghezza delle punte "à cran" dei due giacimenti darebbe il massimo di probabilita' di eguaglianza tra le due popolazioni (dato che le medie sono uguali) bisogna in ogni caso osservare che l'applicazione rigorosa dell'analisi della varianza alla stima e decisione statistica tra "popolazioni", in base alla "sampling theory", presuppone che i campioni (i "samples" appunto) siano assolutamente "random" e non credo che sia questo il caso per quasi nessun campione proveniente da uno scavo archeologico.

* * *

Dopo questa breve illustrazione dell'uso delle statistiche elementari in una dimensione, e' forse il caso di passare allo studio di piu' attributi metrici, ed in particolare alla *correlazione* statisticamente significativa che puo' essere stabilita tra due di essi.

In fig.9 vediamo una *nuvola di punti* o "scatterplot" dove sono rappresentate le punte a dorso ("Gravettes" e "Microgravettes") del giacimento del Perigardiano superiore di Corbiac in Dordogna (Bordes, 1967), in un grafico dove in ascisse abbiamo le larghezze, e in ordinata le larghezze. Ogni punto del grafico rappresenta una di queste punte a dorso.

Come si nota subito i punti sono chiaramente distinti in due gruppi: uno (triangoli) che rappresenta le Microgravettes l'altro, piu' sparso (punti) che rappresenta le Gravettes. Se si facesse la proiezione sull'asse delle larghezze, e cioe' l'istogramma delle larghezze di questi strumento avremmo quindi una distribuzione bimodale, con due picchi ben separati (non altrettanto avviene per le lunghezze).

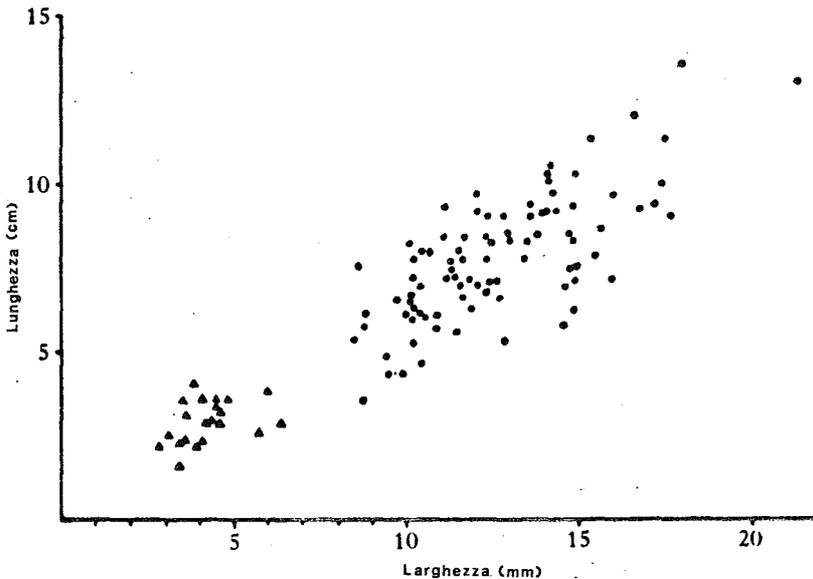


Fig.9

"Scatter diagram" delle punte a dorso (Gravettes e Microgravettes) del giacimento del Perigordiano Superiore di Corbiac (Dordogna) (da Bordes, 1967).

Inoltre la figura mostra che tutti i punti sono sostanzialmente allineati su di una retta, e cioe' c'e' una correlazione abbastanza precisa tra la lunghezza e la larghezza sia delle Gravettes che delle Microgravettes. Questa retta si puo' calcolare mediante il metodo dei minimi quadrati e si chiama *retta di regressione*. Piu' precisamente possiamo scrivere

$$y = a + bx \quad (1.3)$$

dove ad esempio y rappresenta la lunghezza e x la larghezza e i coefficienti a e b si calcolano con il metodo dei minimi quadrati e cioè imponendo che la quantità

$$\sum_{i=1} (y_i - a - bx_i)^2$$

sia minima, dove y_i e x_i sono i valori effettivamente misurati delle lunghezze e delle larghezze rispettivamente.

In modo analogo si potrebbe cercare la miglior retta in cui la lunghezza sia la variabile dipendente e cioè

$$x = a' + b'x \quad (1.4)$$

La quantità

$$r = \sqrt{bb'} \quad (1.5)$$

si chiama coefficiente di correlazione e misura la correlazione tra le due variabili e si può anche definire come

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_k - \bar{y})^2}} \quad (1.6)$$

Se le due rette di regressione fossero perfettamente uguali avremmo ovviamente che $b = \frac{1}{b'}$ e pertanto $r = 1$ (correlazione massima). Nel caso della fig.9 si calcola $r = 0.88$ che è un valore notevolmente alto.

In fig.10 abbiamo lo "scatter plot" delle larghezze e degli spessori delle punte "à cran" di Paglicci precedentemente citate, ottenuto con il solito programma TIPOMT (i numeri indicano il numero dei pezzi che hanno le stesse misure, con la approssimazione di un millimetro per le larghezze e di 0.2 millimetri per lo spessore).

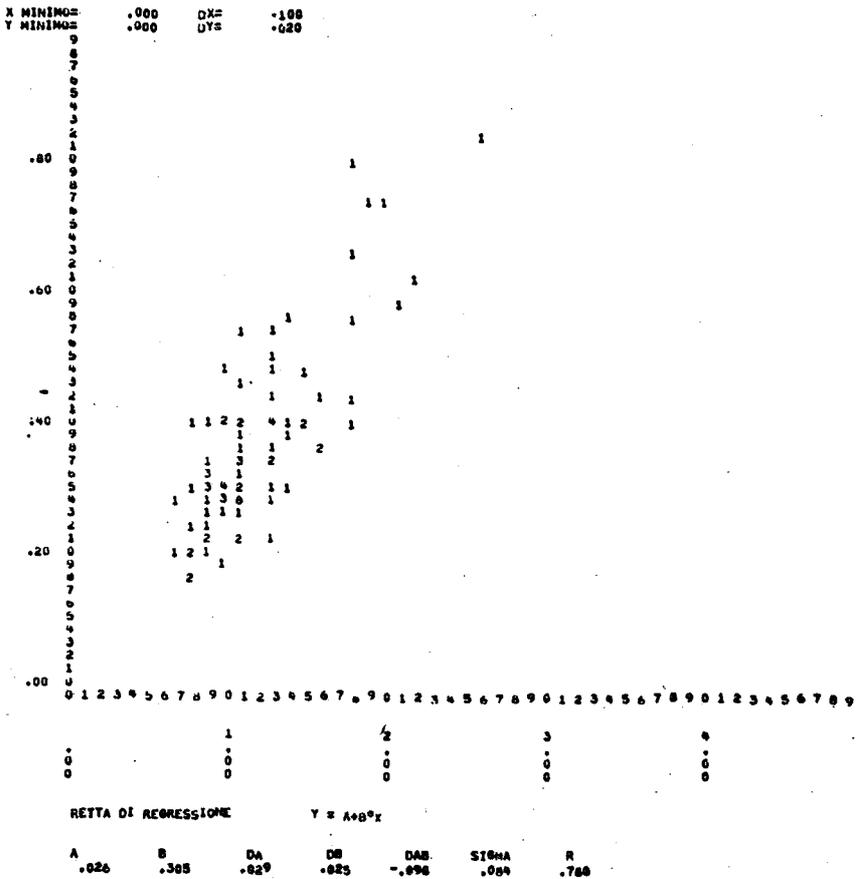


Fig.10

"Scatter diagram" delle larghezze e degli spessori delle punte "à cran" di Grotta Paglicci.

Vengono calcolati anche i coefficienti della retta di regressione e il coefficiente di correlazione e' piuttosto alto ($r = 0.788$).

In fig.11 abbiamo invece la stessa nuvola di punti per gli strumenti "à cran" del giacimento di Taurisano: la correlazione e' buona ($r = 0.51$) ma non cosi' alta come a Paglicci.

Naturalmente piu' i punti sono sparsi nel piano, piu' grande e' l'incertezza sui parametri della retta di regressione e

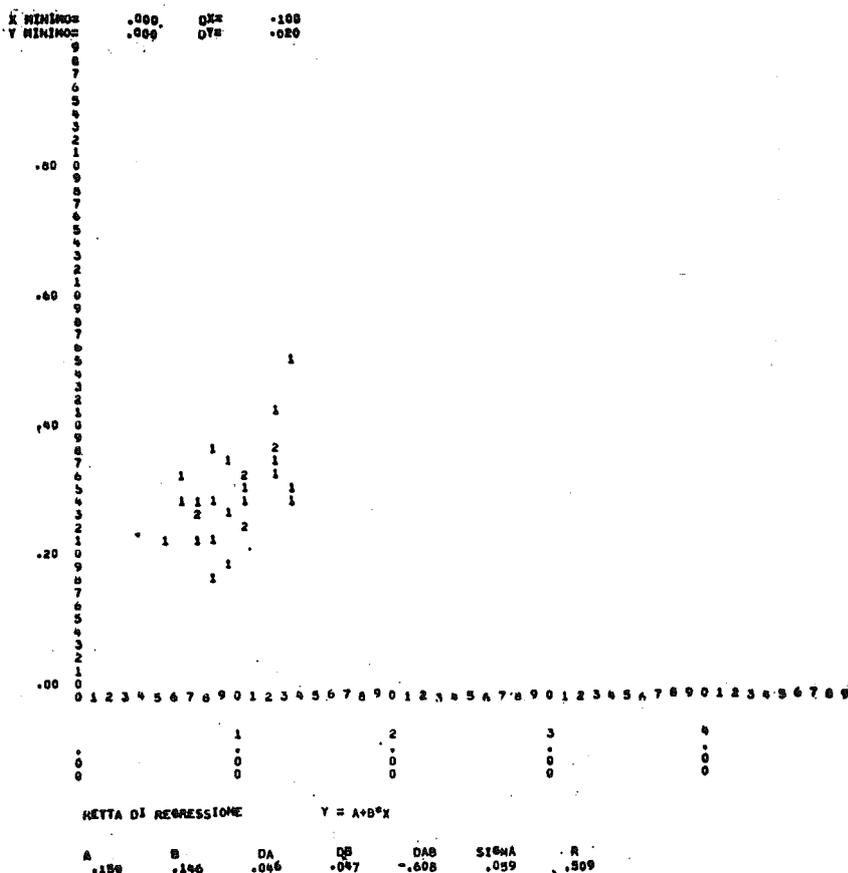


Fig.11

"Scatter diagram" delle larghezze e degli spessori delle punte "à cran" del giacimento di Taurisano.

più piccola è la correlazione. In fig.12, ad esempio, vediamo lo "scatter plot" delle larghezze e delle lunghezze del "cran" per le punte di Paglicci. La distribuzione è piuttosto sparsa, ed infatti il coefficiente di correlazione è soltanto $r=0.197$.

Evidentemente non vi è per questi strumenti una "forma" predeterminata del "cran", mentre vi era una forma predeterminata di tutto lo strumento così come si vedeva dalla correlazione tra larghezza e spessore della fig.10.

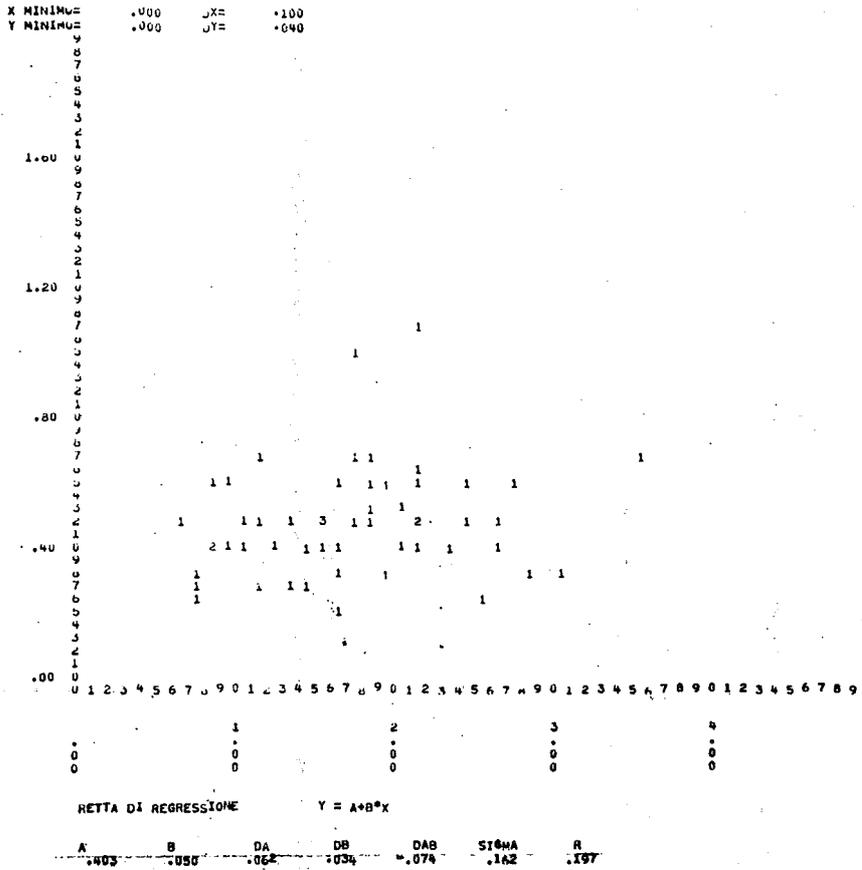


Fig.12

"Scatter diagram" delle larghezze e lunghezze dei "cran" degli strumenti di Grotta Paglicci.

Prima di passare al problema degli attributi non metrici vorrei solo brevemente fare delle considerazioni sulle correlazioni tra un numero di attributi metrici maggiore di due.

Nello spirito della retta di regressione, si usano sistemi di *regressione multipla*, e cioè in più dimensioni: la formula sarà, per il caso di regressione lineare:

$$y = \sum_{k=0}^n a_k x_k \quad (1.7)$$

con $x_0 = 1$.

Si possono anche pensare formule di regressione multipla *non lineari* e cioè quelle dove le varie variabili "dipendenti" x_K non compaiono alla prima potenza, ma con potenze qualsiasi (in particolare, nel caso di due dimensioni, potremo avere espressioni polinomiali, funzioni particolari, ecc.). In ogni caso le tecniche di regressione multipla mostrano le correlazioni di una variabile "privilegiata", per così dire (la y dell'equazione 1.7), rispetto a tutte le altre variabili indipendenti x_K .

Se invece si vogliono trovare correlazioni significative tra attributi metrici, in termini *politetici*, e cioè dove questi attributi sono considerati sullo stesso piano, bisogna usare metodi di analisi *multivariate* che saranno accennate nel prossimo capitolo data la loro importanza per i problemi di comparazione tra più giacimenti. Avremo allora anche l'occasione di parlare di analisi multivariate per problemi di tipologia, nonché di possibili sistemi *monotetici* o "gerarchici" in contrasto con sistemi di classificazione puramente politetici.

* * *

Per quanto riguarda gli attributi "nominali" è ovvio che la ricognizione di una intenzionalità di ritocco su uno strumento di selce, di un particolare tipo di decorazione su un frammento ceramico, ecc., ed in particolare la correlazione fra questi attributi è stata spesso in passato oggetto di studio essenzialmente qualitativo. Come al solito i risultati in questo campo sono largamente soddisfacenti, soprattutto in quei paesi dove c'è una lunga tradizione di studi tipologici e soprattutto un lungo e paziente studio dei materiali, come ad esempio per il Paleolitico francese. In altre regioni, come è già stato osservato (Bietti, 1978a), è possibile che tecniche più quantitative possano aiutare notevolmente lo studio delle correlazioni intenzionali tra attributi per la identificazione

di tipi, sempre intesi da un punto di vista *culturale*, come si è già detto precedentemente.

Il metodo più usato per la ricognizione di correlazioni statisticamente significative è quello delle tabelle di contingenza e del χ^2 (non parametrico), largamente usato in scienze umane, come la biologia e la psicologia, e che con tutta probabilità è stato introdotto nell'archeologia per la prima volta da A.C. Spaulding (1953), e successivamente applicato da numerosi Autori (ad esempio Binford (1963), Sackett (1966)).

Per una illustrazione elementare del metodo, ripeto qui l'esempio già presentato nella mia breve conferenza del 1978.

Supponiamo di voler stabilire se, in un determinato contesto culturale, un tipo di decorazione (a fasce dipinte, ad esempio) è correlato con una forma ceramica particolare (una tazza rispetto ad una brocca, ad esempio). Come esempio concreto, supponiamo di avere 70 tazze, di cui 45 decorate e 25 non decorate, e 40 brocche, di cui 10 decorate e 30 non decorate. Possiamo così costruire una *tabella di contingenza* (tab.1) che riporta questi numeri, insieme ai cosiddetti *totali marginali* (e cioè al numero totale di tazze (70), di brocche (40), di vasellame decorato (55) e non decorato (55)), nonché, ovviamente, il numero totale di oggetti (110). La tabella 1 mostra quella che si chiama la distribuzione *osservata*: per vedere se vi sono correlazioni statisticamente significative, bisogna confrontarla con un'analogha tabella, quella dei valori *aspettati*, calcolati nell'ipotesi che la presenza o assenza di decorazione sia egualmente distribuita tra tazze e brocche, tenendo fissi i totali marginali. Questi valori aspettati vengono ottenuti considerando per ogni casella il prodotto dei due totali marginali corrispondenti (ad esempio 55×70 per la prima casella) e dividendo per il numero totale di pezzi (110 nel nostro caso). Abbiamo così la tabella 2. La significanza statistica si trova con il metodo del χ^2 , e cioè si calcola per ogni casella la quantità

$$\chi^2 = \frac{(\text{freq. osserv.} - \text{freq. aspett.})^2}{\text{freq. aspett.}}$$

Tabella 1

Tabella 2

Tabella 3

	dec.	non dec.	Totale	dec.	non dec.	Totale	dec.	non dec.	Totale
Tazze	45	25	70	35	35	70	2.9	2.9	5.8
Brocche	10	30	40	20	20	40	5	5	10
Totali	55	55	110	55	55	110	7.9	7.9	15.8

ottenendo così la tabella 3. Più grande è il valore del χ^2 e più grande è la deviazione dei dati osservati da quelli aspettati. Nel nostro caso è la distribuzione della decorazione tra le brocche che si discosta di più dall'ipotesi di distribuzione "random".

Per avere una definizione probabilistica di questa deviazione dall'ipotesi "random", si somma su tutte le caselle della tabella 3 per ottenere il valore totale del χ^2 . Si calcola poi il numero di "gradi di libertà" ν , e cioè il numero di parametri liberi del problema, che nel nostro caso (tabella 2×2) è 1, perché anche se le frequenze sono 4, vi sono tre quantità prefissate, e cioè i quattro totali marginali meno però il totale generale che li vincola. Ottenuto il valore di ν ed il χ^2 totale, esistono opportune tabelle, su tutti i testi di statistica, che danno la probabilità di indipendenza (o alternativamente di associazione). Nel nostro caso $\chi^2_{\text{tot.}} = 15.8$ e $\nu = 1$, e per questi valori le tabelle forniscono un valore di probabilità molto più piccolo dello 0.5% (a questo valore corrisponderebbe un $\chi^2_{\text{tot.}} = 7.88$) che la distribuzione osservata sia spiegabile come fluttuazione statistica di quella aspettata in base all'ipotesi "random".

La correlazione presenza di decorazione-tazza e assenza

di decorazione-brocca nel nostro esempio e' pertanto statisticamente significativa.

Il metodo si puo' generalizzare a piu' attributi (per esempio tre tipi di decorazione, o piu' forme ceramiche) nel qual caso otteniamo tabelle di contingenza a piu' righe e piu' colonne, e non piu' 2×2 come nel nostro esempio; o addirittura a piu' classi di attributi (per esempio tipi di decorazione, forme ceramiche, tipi di anse, ecc.).

E' chiaro che in questi casi i calcoli sono piu' laboriosi e l'uso di un calcolatore elettronico diventa particolarmente comodo. E' disponibile, ad esempio, un programma (CNTG) che effettua analisi di contingenza e del χ^2 fino a 5 classi di at-

	Microg. dist.	Microg. prox.	Backed blad.	Totals
No flat inv. ret.	81	50	71	202
Flat prox. inv. ret.	11	27	16	54
Flat dist. inv. ret.	36	1	6	43
Totals	128	78	93	299

86.475	52.696	62.829
23.117	14.087	16.796
18.408	11.217	13.375

a) b)

	0.347	0.138	1.063	
$\chi^2_{tot} = 49.97$	6.351	11.85	0.037	$C = 0.381$
$V = T = 0.3$	16.80	9.32	4.06	$C_{max} = 0.82$

c)

Fig. 13

Tabelle di contingenza per Microgravettes e lamelle a dorso del giacimento Epigravettiano di Palidoro (Roma) (da Bietti, 1974-75).

tributi, con 7 attributi per ogni classe (Bietti e Zanello, 1975).

Come esempio reale di correlazione statisticamente significativa voglio mostrare (fig.13) quella tra punte a dorso microlitiche (Microgravettes) distali e prossimali (e cioè appuntite dalla parte della punta o del tallone) e lamelle a dorso da un lato, e dall'altro la presenza o assenza di assottigliamento inverso distale o prossimale su questi strumenti, provenienti dal giacimento Epigravettiano di Palidoro, presso Roma, datato tra circa 14.000 e 16.000 anni da oggi (Bietti, 1974-1975). In fig.14 diamo un esempio dell'industria rinvenuta.

Come si vede dalla fig.13, e' una tabella 3×3 , e il numero di gradi di liberta', calcolato secondo la formula $\nu = (r-1)(c-1)$, dove r e c sono il numero di righe e di colonne della tabella di contingenza, e' 4 (da notare che, per $r = c = 2$, come era l'esempio precedente delle tazze e delle brocche, la formula da' proprio $\nu = 1$).

Il valore totale del χ^2 , come si vede in fig.13 e' 49.97, che per $\nu = 4$, da' un valore di probabilita' molto piu' piccolo dello 0.01%. In figura vengono dati i valori dei coefficienti di associazione

$$V = \sqrt{\frac{\chi_{tot.}^2}{N(q-1)}} \quad C = \sqrt{\frac{\chi_{tot.}^2}{\chi_{tot.}^2 + N}} \quad (1.8)$$

dove N e' il numero totale degli oggetti, e q e' il piu' piccolo tra c ed r , abbastanza alti ($V = 0.3$ e $C = 0.381$), tenuto conto del fatto che i loro valori massimi sono rispettivamente 1 e 0.82.

La correlazione significativa, come si puo' vedere facilmente dai valori parziali del χ^2 nella tabella c) della figura 13 e' data dall'assottigliamento inverso preferenziale all'estremita' distale delle Microgravettes distali e all'estremita' prossimale per quelle prossimali.

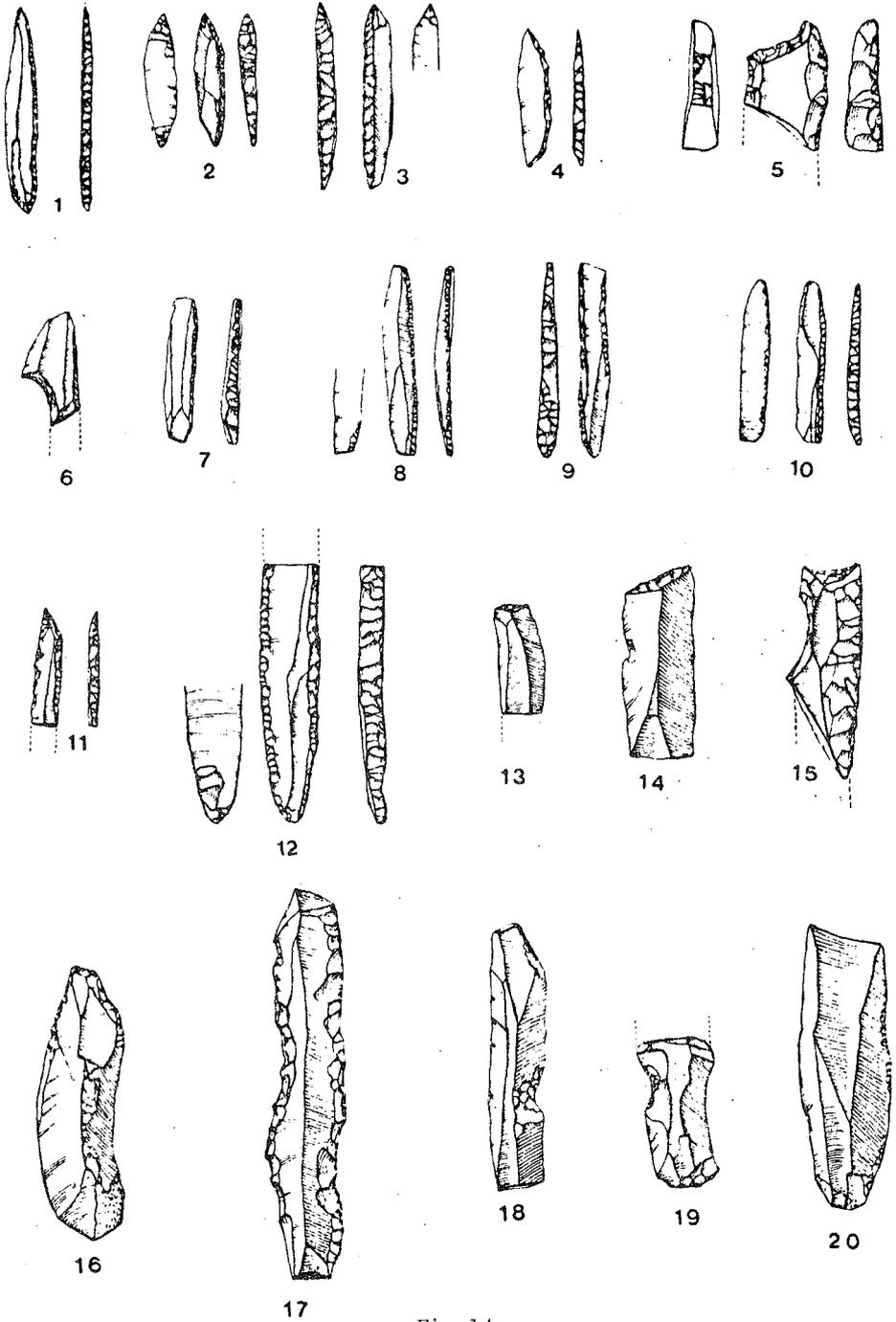


Fig.14

Industria litica del taglio 5 del giacimento Epigravettiano di Palidoro (Roma) (da Bietti, 1976-77).

In fig.15 presentiamo invece un esempio di sostanziale assenza di correlazione: si tratta di stabilire su di un campione piuttosto cospicuo (1474) di nuclei del Paleolitico superiore della grotta Polesini, presso Tivoli (Radmilli, 1974), conservati presso il Museo Pigorini a Roma, se vi è correlazione tra la laminarità dei distacchi, considerata in 5 classi

IL PRIMO INDICE RAPPRESENTA IL TIPO DI NUCLEO = A 1 PIANO
E A 2 PIANI OPPOSTI. 2L SECONDO INDICE RAPPRESENTA LA LAMINARITA'
DEI DISTACCHI A PARTE DA LAMINARITA' NULLA.

NO. OF INDEXES MAX. VALUES FOR THE INDEXES 1974	2		2	5
	INDEXES	OBS.		
1 1	32.000	28.229	.50+00	
1 2	283.000	287.493	.70-01	
1 3	395.000	401.153	.94-01	
1 4	361.000	357.324	.38-01	
1 5	24.000	20.801	.49+00	
2 1	6.000	9.771	.15+01	
2 2	104.000	99.507	.20+00	
2 3	145.000	136.847	.27+00	
2 4	120.000	123.676	.11+00	
2 5	4.000	7.199	.14+01	
CHI SQUARE=	.466+01			
N.O.F. 4	PROB. .324+00			
CRAMER COEFF.=	.056			

Fig.15

Tabelle di contingenza per nuclei prismatici del giacimento Epi-gravettiano di grotta Polesini (Tivoli, Roma).

(da nulla, e cioè senza tracce di distacco di lamelle, fino a totale, e cioè dove tutte le tracce di distacchi sono lamellari) e i tipi di nuclei prismatici: a uno o a due piani di percussione. La figura mostra l'uscita di elaboratore elettronico del già citato programma CNTG (Bietti e Zanello, 1975), e, come si vede, la probabilità che la distribuzione osservata sia spiegata da quella "random" è singolarmente alta: 32.4%. Inoltre il coefficiente di Cramer (V nell'eq. (2)) è anche molto basso (0.056). Non c'è quindi traccia di correlazione preferenziale tra il grado di laminarità dei distacchi e i due tipi di nuclei considerati.

Questo risultato sui nuclei di grotta Polesini, fa parte di uno studio tipologico sui materiali di questo giacimento e

che è in corso di collaborazione con la dr.ssa D. Zampetti dell'Istituto di Paleontologia dell'Università di Roma e del dr. S. Antonelli, che si è laureato appunto sull'analisi quantitativa dei nuclei del giacimento.

È bene ricordare che l'uso delle tabelle di contingenza e del χ^2 dovrebbe essere limitato a correlazioni tra attributi *nominali*, la sua applicazione ad attributi metrici, opportunamente divisi in classi di grandezza (per esempio lamelle, lame strette, lame larghe, ecc.) è stata giustamente criticata recentemente da J.E. Doran e F.R. Hodson (1975).

È chiaro a questo punto, che il problema della ricognizione di "clusters" tipologici coinvolge contemporaneamente sia attributi metrici che nominali, e quindi, come abbiamo già detto, metodi di analisi multivariata di cui avremo occasione di parlare nel prossimo capitolo.

II - Analisi e confronti tra complessi archeologici. Cenni a metodi di analisi multivariata

Secondo quanto già detto nell'introduzione, in questa parte parleremo del livello successivo d'analisi e cioè dei complessi archeologici: giacimenti diversi o strati diversi dello stesso giacimento.

Anche se, in generale, questo tipo di confronto dovrebbe tener conto di tutte le variabili del giacimento: faune, analisi sedimentologiche o paleobotaniche, industrie, ecc., tradizionalmente quest'analisi viene spesso eseguita principalmente al livello delle sole industrie, in particolare per i giacimenti del Paleolitico e Mesolitico.

Il dato di partenza è una lista di tipi ben definiti, secondo i criteri esposti nel capitolo precedente (eventualmente con le tecniche matematiche descritte), e in seguito un confronto tra le percentuali di questi tipi nei diversi giacimen-

ti o strati dello stesso giacimento.

L'uso delle percentuali di tipi, e' meglio precisarlo subito, e' intuitivamente ovvio, perche' serve a "normalizzare" in qualche modo i dati, e cioe' a poter confrontare anche insieme con un numero differente di reperti effettivi.

E' chiaro d'altronde, che l'uso delle percentuali dovrebbe presupporre, cosa non sempre riconosciuta dagli archeologi, una situazione *simile* tra i giacimenti (o strati dello stesso giacimento): molto spesso la superficie scavata e' molto limitata, cosi' che si verifica un "sampling bias" sostanziale nel confronto e pertanto vi possono essere diverse specializzazioni di attivita' nei giacimenti che si vogliono paragonare. Bisogna in piu' osservare che l'uso di percentuali di tipi e' finora sostanzialmente limitato all'archeologia preistorica del Paleolitico e del Mesolitico: per periodi piu' recenti, anche se esistono tipologie abbastanza ben caratterizzate sembra che, almeno in Italia, sia piu' importante la presenza o assenza dei tipi piuttosto che la loro effettiva consistenza percentuale.

Una volta definite le percentuali dei tipi, il metodo piu' comune di confronto tra diversi giacimenti e' quello di costruire *diagrammi cumulativi*, *poligoni di frequenza*, o *diagrammi a blocchi* ("block diagrams" oppure "bar charts") sulla base della lista tipologica.

Il diagramma cumulativo, introdotto dalla scuola di Bordeaux (Bordes, 1953), e' costruito aggiungendo alla percentuale di un tipo quella del tipo successivo nella lista: e' un po' come l'*ogiva* della statistica elementare: si parte cosi' da un valore della percentuale che corrisponde a quello del primo tipo della lista, per arrivare al 100% con l'ultimo tipo.

In fig.16, in alto, sono rappresentati, per esempio, i diagrammi cumulativi, basati su di una lista di 92 tipi, di due giacimenti dell'Aurigraziano francese: Lartet (linea piena) e Poisson (linea tratteggiata) (Sonneville-Bordes e Perrot, 1953). Nella stessa figura (in basso) le stesse percentuali sono in-

vece visualizzate mediante poligoni di frequenze, mentre i diagrammi a blocchi (l'analogo degli istogrammi delle variabili continue in statistica elementare) nella parte superiore della figura sono costruiti a partire da gruppi dei 92 tipi della lista: grattatoi, bulini diedri, bulini su troncatura, ecc.

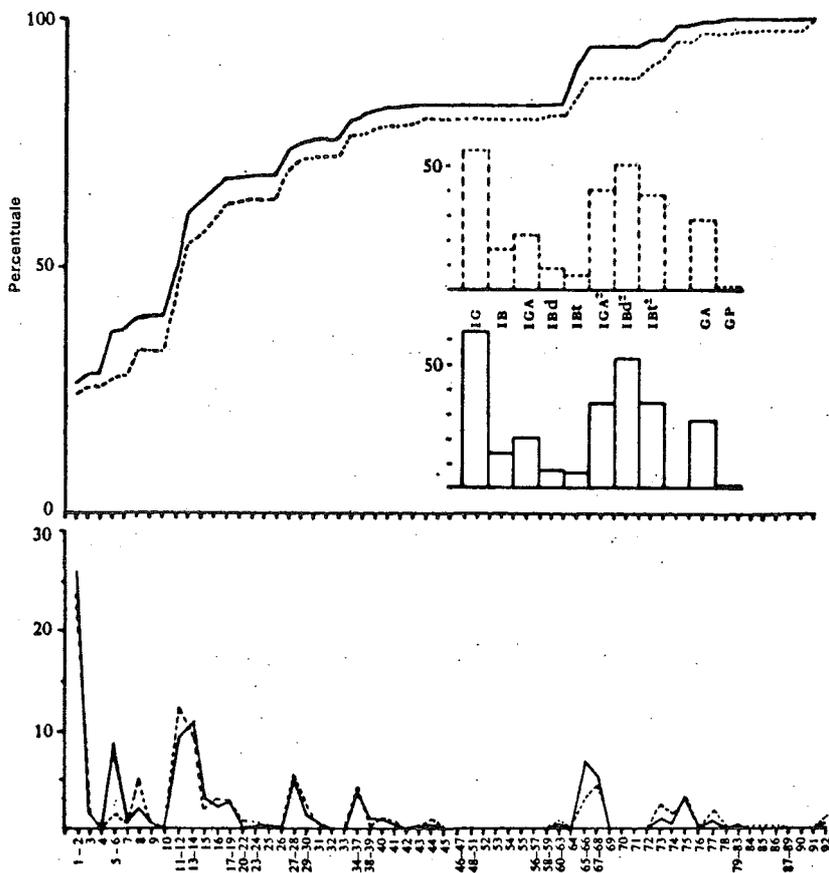


Fig. 16

Diagrammi cumulativi, diagrammi a blocchi e poligoni di frequenza per i giacimenti Aurignaciani di Lartet e Poisson (Dordogna) (da Sonnevile-Bordes e Perrot, 1953).

Per quanto riguarda i diagrammi a blocchi, è bene ricordare che è il metodo di rappresentazione grafica più comunemente usato da parte degli studiosi del Paleolitico italiano, ed

in particolare quelli che seguono il metodo tipologico di G. Laplace (1964).

A questo livello "zero" di matematizzazione, la rappresentazione grafica mediante il calcolatore elettronico fa risparmiare tempo, e non presenta nessuna difficoltà: in fig.17 vediamo, ad esempio, i diagrammi cumulativi di tre livelli del già citato giacimento Epigravettiano di Taurisano (Lecce) (Bietti, 1979a), eseguiti mediante l'elaboratore UNIVAC dell'Università di Roma.

L'uso dei diagrammi cumulativi è stato oggetto di varie critiche (vedi ad esempio Kerrich e Clarke, 1967). Secondo queste obiezioni l'ordinamento dei tipi è fatto in modo tale da mascherare oggettive differenze percentuali: vi sono dei raggruppamenti tipologici "a priori", per esempio, gruppo dei grattatoi Aurignaziani, gruppo di strumenti Solutreani, ecc. (per la lista del Paleolitico superiore).

In realtà, secondo me, questa circostanza ha solo un'influenza *qualitativamente visiva*: è chiaro che anche un raggruppamento diverso dei tipi, analizzato in modo accurato, può rivelare esattamente le stesse differenze in percentuali, che si possono riscontrare in grafici cumulativi con "a priori diagnostics".

Le critiche più profonde che possono essere fatte al metodo dei diagrammi cumulativi sono invece, a mio avviso, sostanzialmente due. In primo luogo non è possibile fare ipotesi di similarità o differenza tra giacimenti diversi, se si cerca di confrontarne più di 5 o 6.

A titolo di esempio in fig.18 viene mostrato il confronto tra diagrammi cumulativi del Musteriano di due strati di Grotta Guattari (2 e 4) (Taschini, 1979) con quello dei livelli 17 e 22 di Combe-Grenal in Dordogna (Bordes et al., 1966) e del giacimento di Chinchon nella Gironde (Sireix e Bordes, 1972).

Come si vede, la figura è già abbastanza confusa anche se si vede chiaramente un'affinità notevole dell'industria di Grot-

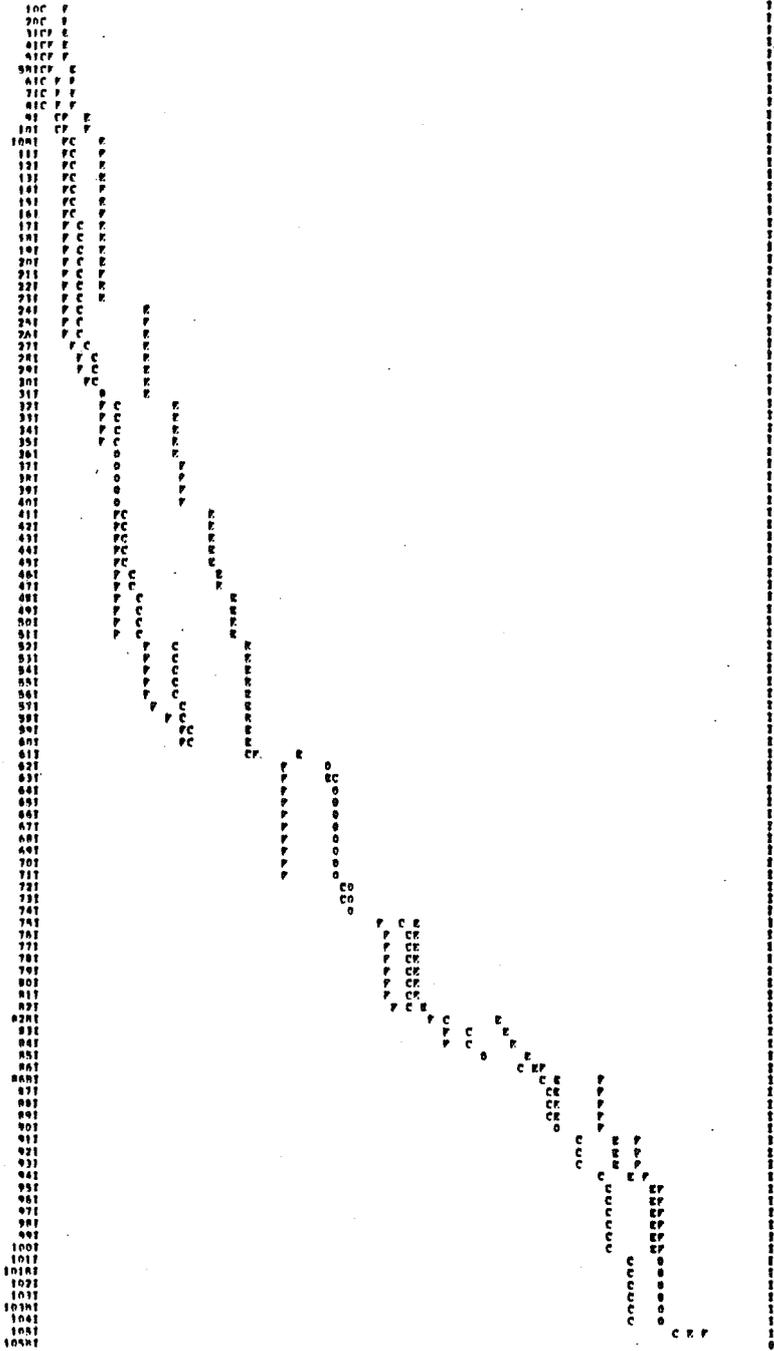


Fig. 17
Diagrammi cumulativi di tre livelli del giacimento di Taurisano (Lecce).

ta Guattari (che è quella dove è stato rinvenuto il famoso cranio del Circeo), ed in particolare quella dello strato 2, con Chinchon e con lo strato 22 di Combe-Grenal, che rappresentano un Musteriano di tipo Quina molto caratteristico.

Questa figura ha anche un interesse, per così dire "storico", perché è la prima volta, per quello che mi risulta, che viene presentato un confronto tra diagrammi cumulativi di giacimenti classici del Musteriano francese, e giacimenti italia-

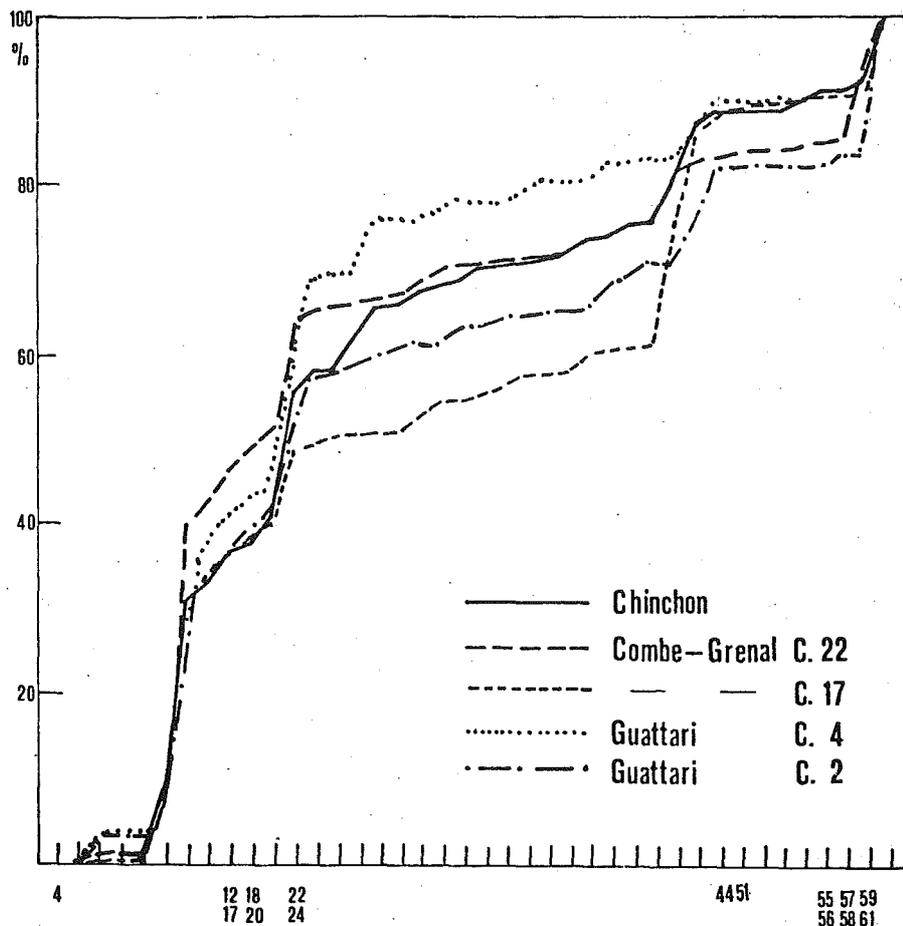


Fig. 18

Grafici cumulativi di giacimenti mustერიანი: Grotta Guattari (Latina) strati 4 e 2, Combe-Grenal (Dordogne), Chinchon (Gironde). (da Taschini, 1979).

ni, per lo meno dell'area centro-meridionale.

L'altra limitazione dell'uso dei diagrammi cumulativi, risiede, a mio giudizio, nell'impossibilita' di dare stime di *significanza statistica* nel confronto percentuale tra due (o piu') giacimenti. Nel confronto tra le percentuali di strumenti di due giacimenti l'idea piu' semplice e' proprio quella di considerare, per il numero di oggetti rappresentativi di ogni singolo tipo, un "errore" dato dalla distribuzione di Poisson, che si riferisce appunto a numeri interi. In questo caso la frequenza (o probabilita' secondo i concetti della statistica elementare per grandi numeri) con cui si puo' trovare un numero n di conteggi, se il numero medio e' \bar{n} e' data da

$$f(n) = \frac{n^{\bar{n}} e^{-\bar{n}}}{n!} \quad (2.1)$$

dove ovviamente $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1)n$. L'"errore" e' cioe' la deviazione standard per una distribuzione di questo tipo e' dato da

$$\sigma = \sqrt{\bar{n}} \quad (2.2)$$

A partire dalle valutazioni di σ sul numero di strumenti per ogni singolo tipo (n) e per il totale N di tipi, si puo', attraverso la formula di propagazione degli errori, calcolare l'errore sulla percentuale ΔP , secondo l'espressione

$$\Delta P = \frac{1}{N^2} \sqrt{k^2 (\Delta n)^2 + n^2 (\Delta k)^2} \quad (2.3)$$

dove $k = N - n$ e' ovviamente $\Delta n = \sqrt{\bar{n}}$ e $\Delta k = \sqrt{k}$.

E' chiaro che dato il carattere "additivo" di un diagramma cumulativo la rappresentazione di errori sul grafico e' praticamente impossibile, mentre puo' essere agevolmente eseguita per diagrammi a blocchi ("bar charts").

In fig. 19 sono mostrati proprio questi diagrammi a blocchi

per il giacimento già citato del Paleolitico superiore di Taurisano (Bietti, 1979a), tagli 4-5 e 7-9, e per il taglio 5 del giacimento, praticamente contemporaneo, di Palidoro, nel territorio di Roma (Bietti, 1976-77). In a) è mostrato il diagramma a blocchi per gruppi di tipi: grattatoi (G), strumenti composti (Com), punteruoli e becchi (Bc), bulini (Bu), dorsi grandi (Gravettes, ecc., Do), troncature (T), lame ritoccate (L.R.), raschiatoi e denticolati (P.V.), geometrici (Ge), microliti (M.), punte (P), strumenti "ésquillés" e diversi (ES.D.).

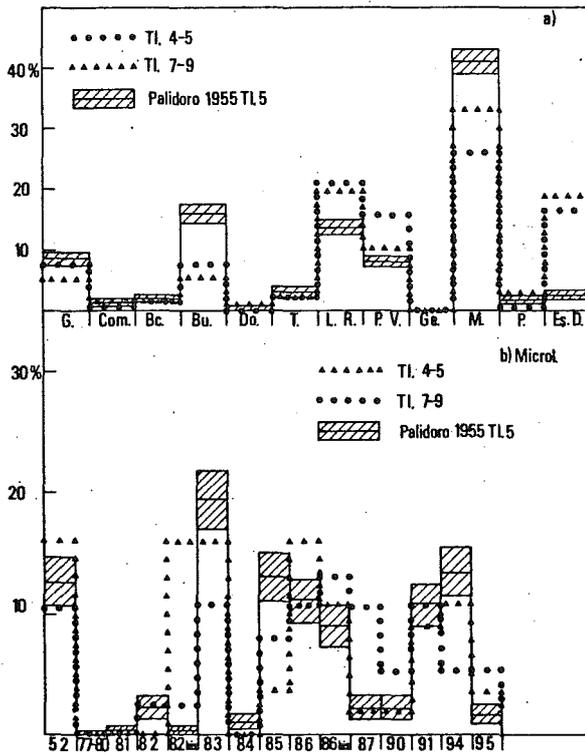


Fig. 19

Diagrammi a blocchi per i giacimenti di Palidoro (Roma) taglio 5, e Taurisano (Lecce) tagli 4-5 e 7-9 (da Bietti, 1979a).

Le percentuali di Palidoro (tl. 5) sono appunto segnate da un'area tratteggiata che è sostanzialmente $\pm \Delta P$ (v. formula

(2.3)).

Come si vede le discrepanze tra i due giacimenti sono essenzialmente nella distribuzione dei bulini, dei microliti e dei diversi ed "esquillés".

Nella stessa figura 19 b) è mostrata la percentuale "di gruppo" dei microliti (comprese le Microgravettes (n. 52)): i numeri si riferiscono appunto alla lista tipologica del Paleolitico superiore di Bordeaux aggiornata (v. ad esempio Bietti, 1976-77 e 1979a): è interessante notare, ad esempio, come il tipo 82bis (microlito "à cran") di cui ci siamo già occupati nel I° capitolo, è molto importante per i tagli 7-9 di Taurisano, mentre è sostanzialmente assente nel taglio 5 di Palidoro, ben al di là di due deviazioni standard, in sostanziale accordo con l'ipotesi che questo tipo, a quell'epoca, come abbiamo già detto, è caratteristico della zona medio-bassa adriatica, ed è praticamente assente sulla costa tirrenica.

In fig. 20 abbiamo invece i diagrammi a blocchi per gruppi di strumenti di diversi tagli di uno stesso giacimento Epigravettiano (Taurisano, Bietti, 1979a). Questa volta le deviazioni standard secondo la formula (2.3) sono calcolate per i tagli statisticamente più ricchi: 4-5 in a) e b) e 7-9 in c), per ragioni di migliore consistenza statistica.

La figura mostra una sostanziale omogeneità tra i tagli 1-3, 4-5 e 6, e tra i tagli 7-9, 13-17 e 18-23, entro i limiti degli errori statistici. Deviazioni più grandi si hanno invece tra i tagli 4-5, 10-12 e 18-23, in particolare per le lame ritoccate ed i microliti.

Queste deviazioni scompaiono sostanzialmente, però, nell'ambito di due deviazioni standard per le percentuali di ogni gruppo di tagli. Si può giustamente osservare che l'applicazione di distribuzioni statistiche teoriche, come quella di Poisson, che presuppone campioni *effettivamente* rappresentativi di popolazioni, ad insiemi archeologici, che sono generalmente ben lontani dal soddisfare a queste condizioni, non ha molto senso.

Resta però il fatto che, solo sulla base di differenze percentuali, è procedura comune da parte degli studiosi tradizionali, riconoscere *differenze culturali* sostanziali tra insiemi industriali, indipendentemente da qualsiasi considerazione di

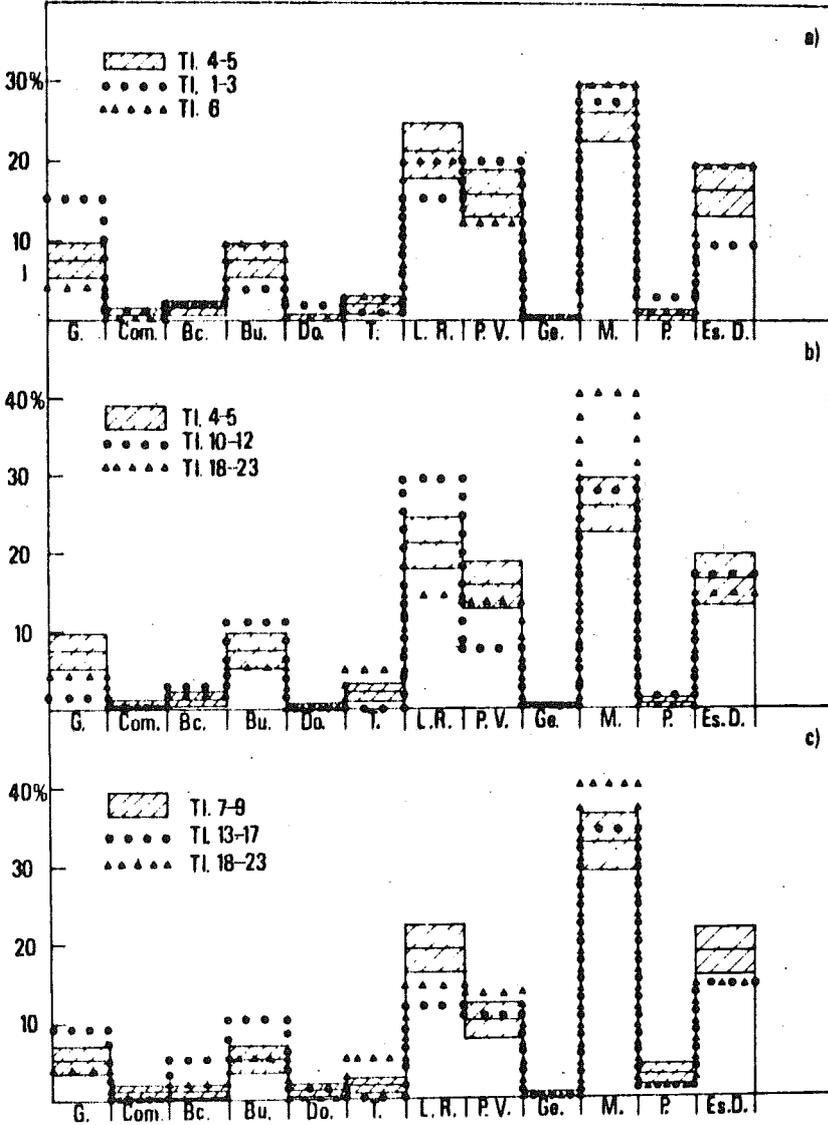


Fig. 20

Diagrammi a blocchi per tre diversi gruppi di tagli del giacimento di Taurisano (Lecce) (da Bietti, 1979a).

"buon senso" sulle deviazioni "sistematiche" insite nei dati: giacimenti mal scavati o scavati solo in parte, strati o depositi considerati come rappresentativi, anche in assenza di correlazioni più generalmente ecologico-ambientali, della *totalità* della cultura del gruppo in esame, ecc.

E' chiaro pertanto che in questa situazione il confronto tra insiemi archeologici, tenendo conto delle deviazioni standard date da distribuzioni teoriche, ha almeno il valore di con-

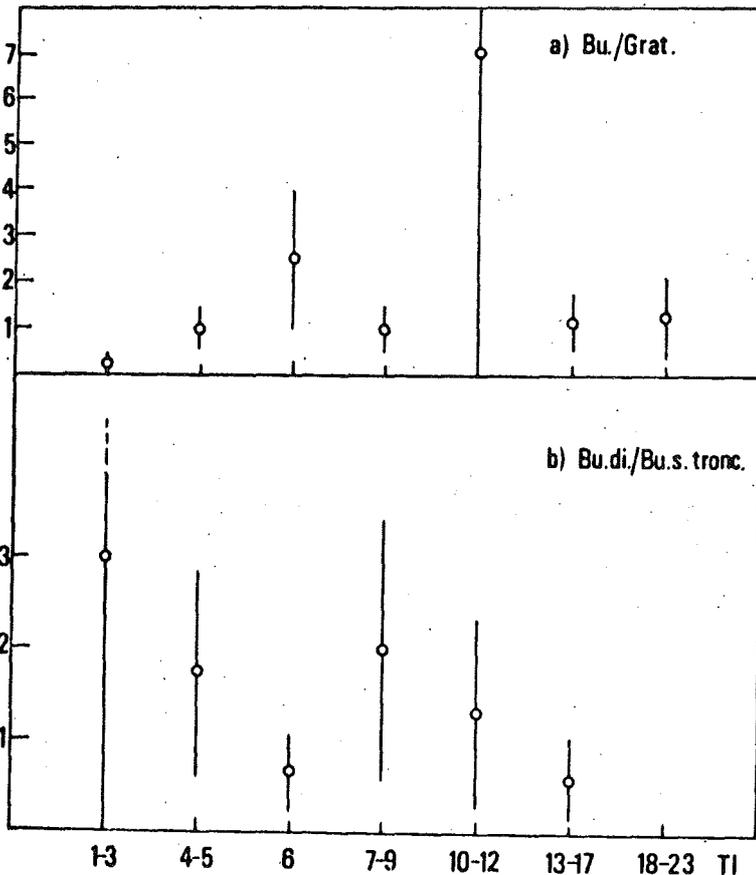


Fig. 21

a) Rapporto tra i bulini ed i grattatoi, b) rapporto tra i bulini diedri ed i bulini su troncatura per sette gruppi di tagli del giacimento di Taurisano (Lecce) (da Bietti, 1979a).

fermare, o quanto meno di non smentire, l'*ipotesi di zero* ("null hypothesis"), e che cioè non si può affermare che due insiemi archeologici sono diversi, se le percentuali dei tipi che li compongono coincidono entro due deviazioni standard.

In fig. 21, ad esempio, si possono vedere i rapporti tra bulini e grattatoi (fig. 21a) e tra bulini diedri e bulini su troncatura (fig. 21b) per sette gruppi di tagli (in ascisse) sempre del giacimento di Taurisano, rappresentati ± 1 deviazione standard, calcolata tramite la formula di propagazione degli errori e la solita distribuzione di Poisson data da

$$\Delta R = \frac{1}{B} \sqrt{(\Delta A)^2 + \frac{A^2}{B^2} (\Delta B)^2} \quad (2.4)$$

con $\Delta A = \sqrt{A}$ e $\Delta B = \sqrt{B}$, $R = \frac{A}{B}$.

Come si vede, la variabilità è piuttosto larga, e non vi è pertanto nessuna indicazione di deviazione dall'*ipotesi di zero*, e cioè della sostanziale omogeneità dei rapporti considerati nei vari gruppi di tagli.

Invece di considerare le percentuali di tipi per ogni giacimento, si possono considerare i loro numeri effettivi e poi paragonarli tramite il metodo delle tabelle di contingenza, descritto nel capitolo precedente. In questo caso le righe, per esempio, della tabella sono date da vari strati o giacimenti e le colonne dai vari tipi considerati. Questo metodo è stato usato ad esempio da G. Laplace (1974), e non differisce molto dal precedente, basato sulle percentuali e le deviazioni standard, dato che l'interpretazione probabilistica con il metodo del χ^2 per le tabelle di contingenza è sempre basato su distribuzioni teoriche del tipo Poisson.

A titolo di esempio in fig. 23 mostriamo l'applicazione del metodo del χ^2 ai raschiatoi (che sono gli strumenti predominanti in entrambi i giacimenti) per lo strato 2 di Grotta Guattari (Circeo) con industria "pontiniana" (Taschini, 1979) (di

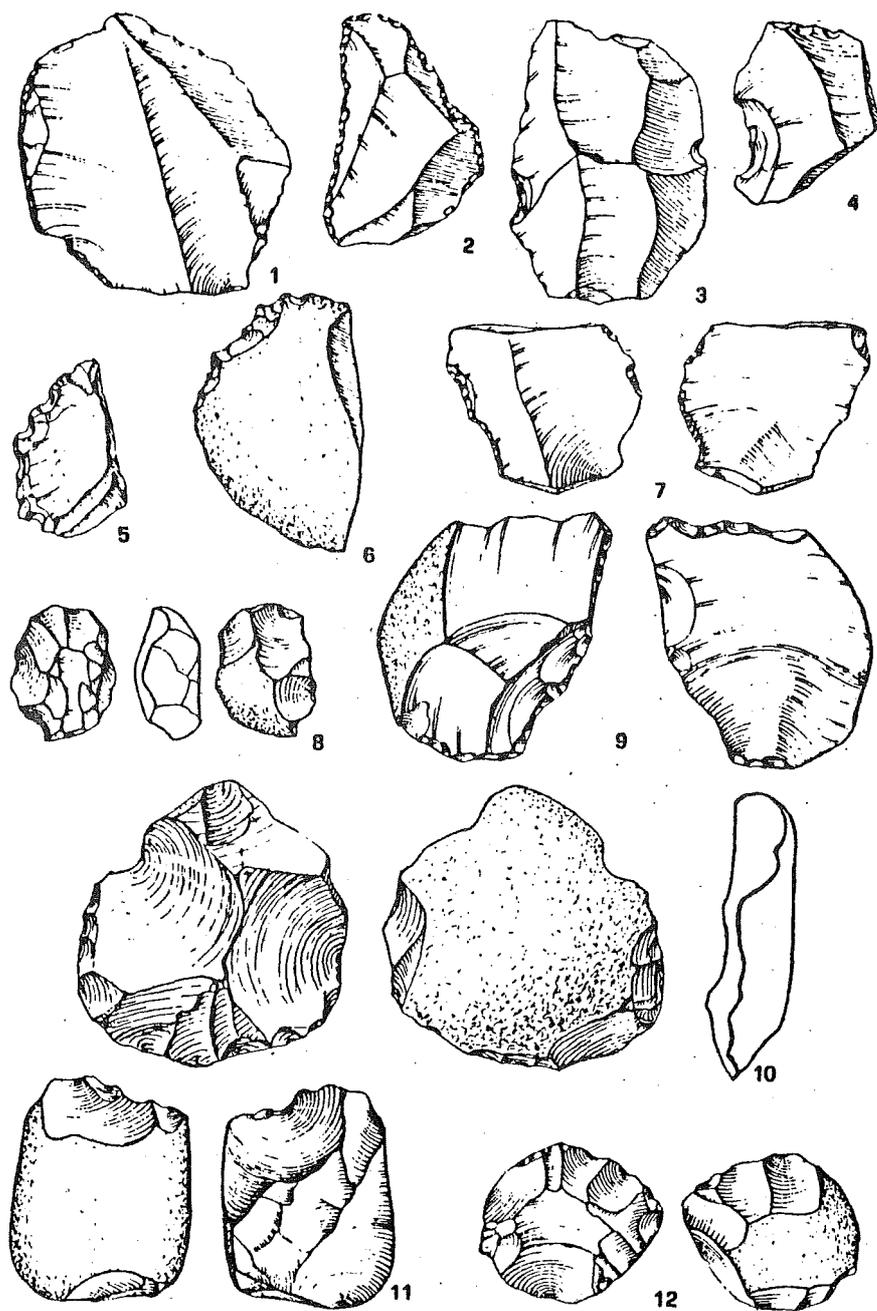


Fig.22

Industria dello strato 2 di Grotta Guattari (Latina) (da Taschini, 1979).

cui diamo un'illustrazione in fig. 22) e dello strato *d* del giacimento di Torre in Pietra, vicino Roma, con industria definita "premusteriana" (Piperno e Biddittu, 1978). La figura mostra un'uscita stampata da un terminale video interattivo (notare il perentorio messaggio degli operatori del Centro di Calcolo del-

CORRELATIONS BY THE METHOD OF CONTINGENCY TABLES

FIRST INDEX REPRESENTS SITES GUATTARI 2 AND TORRE IN PIETRA D
 SECOND INDEX REPRESENTS BORDES TYPES 9,10,11,12-17,18-20,21,22-24,25-29.

NTOT=	NO. OF INDEXES MAX. VALUES FOR THE	2	2	8	
	INDEXES		OBS.	EXP.	PART. CHI-SQUARE
	1 1		13.000	13.397	.12-01
	1 2		28.000	22.367	.11+01
	1 3		3.000	4.211	.35+00
*TM*PREGO LIBERARE IL NASTRO.					
	1 4		5.000	8.421	.14+01
	1 5		2.000	2.297	.38-01
	1 6		12.000	11.100	.73-01
	1 7		12.000	11.483	.23-01
	1 8		5.000	6.124	.21+00
	2 1		22.000	21.603	.73-02
	2 2		32.000	37.033	.68+00
	2 3		8.000	6.789	.22+00
	2 4		17.000	13.579	.86+00
	2 5		4.000	3.703	.24-01
	2 6		17.000	17.900	.45-01
	2 7		18.000	18.517	.14-01
	2 8		11.000	9.876	.13+00
CHI SQUARE=	.517+01				
N.D.F. =	7	PROB.	.639+00		
CRAMER COEFF.=			.157		

Fig. 23

Tabelle di contingenza per i raschiatoi musteriani dello strato 2 di Grotta Guattari (Latina) e per quelli dello strato *d* di Torre in Pietra (Roma).

l'Universita' di Roma: "prego liberare il nastro": sono gli inconvenienti che spesso accadono quando si lavora in interattivo e si e' poco esperti!) usando il gia' citato programma CNTG (Bietti e Zanella, 1975).

Il risultato interessante e' che la distribuzione dei vari tipi di raschiatoi (i numeri 9, 10, ecc. sono quelli della lista tipologica di F. Bordes (1961): semplici diritti, convessi, concavi, doppi, ecc.) e' praticamente la stessa per i due giacimenti: infatti la probabilita' che la distribuzione sia "random"

è ben del 64% e il coefficiente di Cramer, già definito nel capitolo precedente, è solo 0.157, tenendo conto del fatto che il numero totale di strumenti non è molto elevato.

Si potrebbe obiettare che alcuni valori aspettati come si vede dalla figura (per l'esattezza quelli con indici 1,3, 1,5 e 2,5) sono minori di 5 e pertanto bisognerebbe applicare delle correzioni al tradizionale metodo del χ^2 , ben note in tutti i testi di statistica. In ogni caso, il valore del 64% è talmente alto che l'influenza di queste correzioni non può alterare in modo sostanziale le conclusioni; il cui interesse deriva dal fatto che siano così somiglianti industrie tradizionalmente datate una al Tirreniano o Interglaciale Riss-Wurm (Torre in Pietra, liv. *d*) e l'altra al Wurm II (strato 2 di Grotta Guattari) e quindi separate da un intervallo di tempo dell'ordine di grandezza di 50.000 anni.

* * *

Per tornare ora alla prima difficoltà dell'uso dei diagrammi cumulativi, e cioè la difficoltà di effettuare valutazioni comparative quando il numero di insiemi sia superiore a 5 o 6 è chiaro che neanche il metodo delle tabelle di contingenza e del χ^2 , di cui si è appena parlato, serve allo scopo, poiché si limita ad osservare a quali strumenti è dovuta una differenza significativa tra i vari giacimenti o strati presi in esame, a seconda del valore del χ^2 parziale ottenuto.

Si spiega così per quale ragione ad un primo livello di "quantificazione" i confronti di insiemi archeologici, nonostante abbiano come punto di partenza le percentuali dei vari tipi, considerati quindi quantitativamente sullo stesso piano, tradizionalmente finiscono spesso per fare ricorso solo a particolari tipi considerati "diagnostici" o "fossili direttori" delle culture.

È chiaro che, se, ad esempio, i tipi considerati sono solo due, si possono paragonare quanti giacimenti o strati si vo-

gliono, mediante una nuvola di punti, come quelle considerate per gli attributi metrici del capitolo precedente in cui in ascisse e ordinate sono rappresentate le percentuali dei due tipi e i punti della nuvola sono i giacimenti o strati.

Se i tipi sono tre si può ad esempio, ricorrere ad un *diagramma triangolare* (molto usato dai geologi). In fig.24 abbiamo le proporzioni di tre tipi "diagnostici" del Perigordiano V

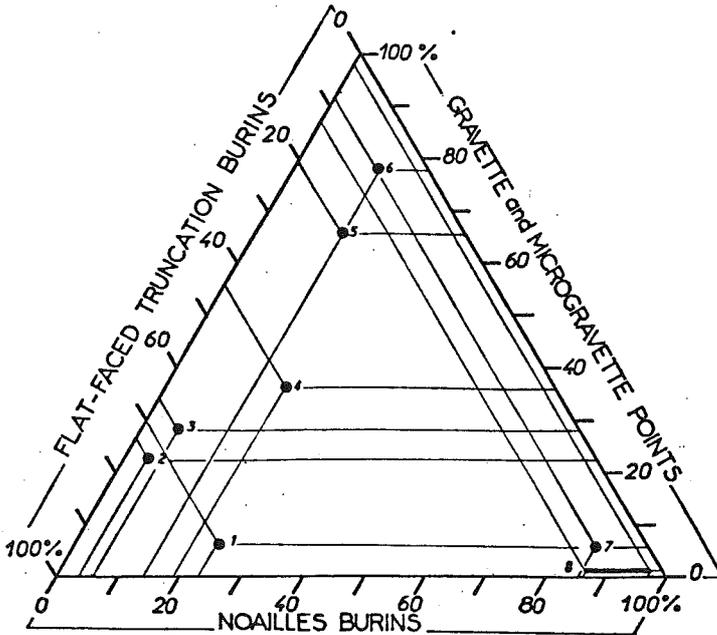


Fig.24

Diagramma triangolare per tre tipi importanti del Perigordiano V nel Perigord: 1-2; Flageolet I strati V-VI; 3-4, Les Jambes strati 3-2; 5-6, Flageolet I strati VII-VI; 7, Roc de Gavaudun strato II; 8, Abri du Facteur strati 10-11 (da Laville e Rigaud, 1973).

nella zona classica della Dordogna (Laville e Rigaud, 1973): le Gravettes e Microgravettes, i bulini piani su troncatura (tipo "Bassaler") e i bulini di Noailles.

Come si vede gli otto giacimenti (o strati dello stesso giacimento come il Flageolet) si raggruppano in modo abbastan-

za distinto.

Se il numero di "parametri" per giacimento e cioè dei tipi, aumenta, e' chiaro che le difficoltà di interpretazione crescono, ed il problema diventa quindi quello di tenere conto, per un numero abbastanza grande di unità da correlare, della interdipendenza su base "politetica" di tante variabili e le prime applicazioni in questo senso (v. per esempio Clarke, 1968 per un orientamento generale), impiegano metodi di analisi *multivariate*, già da tempo impiegate in psicologia e biologia.

I tipi di analisi più comuni, per insiemi archeologici, sono quelle comunemente chiamate di tipo Q e di tipo R (v. per esempio Tugby, 1969).

Il primo tipo di analisi cerca di riconoscere raggruppamenti significativi ("clusters") tra entità archeologiche (giacimenti, strati, necropoli preistoriche, ecc.) a partire dal loro contenuto di *caratteri* (strumenti litici dati in percentuali, corredi di una tomba in termini di presenza o assenza, ecc.). In questi casi i metodi più comunemente seguiti sono quelli della cosiddetta *Tassonomia Numerica* (Sokal e Sneath, 1963).

Nel modo R, invece, le variabili sono i giacimenti o le tombe, in altre parole le entità archeologiche e gli oggetti da correlare, tra cui si cercano associazioni, sono i tipi litici, ceramici, ecc. presenti in ogni entità archeologica. Si possono usare i metodi di tipo Q, ma più comuni sono, in questo caso, procedimenti come l'analisi delle componenti principali ("principal component analysis") che cercherò di illustrare tra breve.

Per incominciare con le analisi di tipo Q, vi sono diverse tecniche matematiche che si possono impiegare, ma il punto di partenza comune è costituito dal calcolo di *coefficienti di similarità* tra le entità archeologiche che si vogliono correlare. Ne elencheremo qui alcuni, tra quelli più usati nelle analisi correnti.

Uno dei più semplici è certamente il "simple matching

coefficient", o coefficiente di Sokal e Michener, molto utile nel caso di confronto tra tombe, e che si basa sul conteggio di presenze ed assenze contemporanee di oggetti di corredo nelle tombe. Prese in considerazione la tomba i -esima e la tomba j -esima, indicando con N_{++} il numero di oggetti di corredo contemporaneamente presenti nelle due tombe, con N_{--} il numero di quelli contemporaneamente assenti, con N_{+} e con N_{-} il numero di quelli presenti nella prima ed assenti nella seconda e viceversa, si definisce il coefficiente di Sokal e Michener per le tombe i e j semplicemente come

$$S_{sm}(i, j) = \frac{N_{++} + N_{--}}{N_{++} + N_{--} + N_{+-} + N_{-+}} \quad (2.5)$$

Questo coefficiente varia tra 0 (tombe totalmente differenti) e 1 (tombe perfettamente eguali).

Un analogo coefficiente, sempre basato sulle presenze ed assenze di caratteri, è quello di Jaccard, che differisce da quello di Sokal e Michener perché non tiene conto del numero di oggetti che sono contemporaneamente assenti, e cioè manca di N_{--} sia al numeratore che al denominatore dell'equazione (2.5).

È chiaro che la scelta tra questi due coefficienti dipende dal tipo di problema che si ha davanti: in un contesto archeologico le assenze di certi oggetti possono essere significative, mentre in un altro possono costituire semplicemente delle variabili ridondanti, ed è pertanto l'archeologo stesso che deve scegliere il tipo di coefficiente che considera più opportuno.

Un altro coefficiente di uso molto comune è quello di correlazione lineare, già introdotto nell'equazione (1.6): considerati due elementi da correlare, i e j , composti ognuno da n caratteri di valore $x_k^{(i)}$ e $x_k^{(j)}$ rispettivamente ($k=1, \dots, n$), con medie date pertanto da

$$\bar{x}^{(i)} = \frac{1}{n} \sum_1^n x_k^{(i)} \quad \bar{x}^{(j)} = \frac{1}{n} \sum_1^n x_k^{(j)}$$

il coefficiente (che varia tra -1 e +1) sarà dato da

$$r_{ij} = \frac{\sum_1^n (\bar{x}^{(i)} - x_k^{(i)}) (\bar{x}^{(j)} - x_k^{(j)})}{\sqrt{\sum_1^n (\bar{x}^{(i)} - x_k^{(i)})^2 \sum_1^n (\bar{x}^{(j)} - x_k^{(j)})^2}} \quad (2.6)$$

Un altro coefficiente che si può usare è quello di distanza (in questo caso tanto più è piccolo il coefficiente tanto più simili sono le entità da correlare), definito come

$$d_{ij} = \sqrt{\sum_1^n (x_k^{(i)} - x_k^{(j)})^2}$$

Una modificazione di questo coefficiente è particolarmente utile per confrontare complessi di industrie, ad esempio del Paleolitico, che, come abbiamo visto, sono solitamente espresse in termini in percentuali di strumenti. Se abbiamo due giacimenti i e j con percentuali per lo strumento k -esimo ($k = 1, \dots, n$, dove n è il totale di tipi considerati), $p_k^{(i)}$ e $p_k^{(j)}$, il coefficiente

$$p_{ij} = 1 - \sqrt{\frac{1}{2} \sum_1^n (p_k^{(i)} - p_k^{(j)})^2} \quad (2.7)$$

varia tra 0 ed 1 (giacimenti con percentuali perfettamente identiche).

Si potrebbero enumerare molti altri tipi di coefficienti di similarità, ma è meglio ora cercare di spiegare, in modo sommario, i procedimenti più comuni di "cluster analysis", una volta calcolati i coefficienti di similarità tra le entità archeo-

logiche a due a due. Se il numero di queste entita' e' N, si ottengono $N(N-1)/2$ coefficienti di similarita', che costituiscono una matrice. Un esempio e' dato in tabella 4 per sei elementi, con coefficienti di similarita' che variano tra 0 e 1.

Tabella 4

	1	2	3	4	5	6
1	1	0.3	0.5	0.9	0.6	0.7
2	0.3	1	0.2	0.5	0.9	0.2
3	0.5	0.2	1	0.1	0.1	0.8
4	0.9	0.5	0.1	1	0.8	0.6
5	0.6	0.9	0.1	0.8	1	0.5
6	0.7	0.2	0.8	0.6	0.5	1

Tabella 5

	2	5	4	1	6	3
2	1	0.9	0.5	0.3	0.2	0.2
5	0.9	1	0.8	0.6	0.5	0.1
4	0.5	0.8	1	0.9	0.6	0.1
1	0.3	0.6	0.9	1	0.7	0.5
6	0.2	0.5	0.6	0.7	1	0.8
3	0.2	0.1	0.1	0.5	0.8	1

Come si vede, la matrice ha tutti 1 sulla "diagonale principale" (coefficienti di similarita' di un elemento con se stesso) ed e' *simmetrica*, e cioe' i numeri a destra della diagonale principale sono uguali a quelli a sinistra, dato che il coefficiente di similarita' tra i e j e' ovviamente uguale a quello tra j e i .

Il problema e' ora quello di raggruppare tra di loro gli elementi con coefficiente di similarita' piu' alto, isolando cosi' dei "clusters".

Uno dei metodi tradizionalmente impiegati e' quello del riordinamento della matrice, molto in uso per i problemi di seriazione cronologica, e cioe' quando si vogliono raggruppare le entita' archeologiche (tombe di una necropoli, ad esempio), in modo da costituire sequenze cronologiche. Questo metodo e' stato probabilmente introdotto per la prima volta da G.W. Brainerd (1951) e W.S. Robinson (1951).

Si tratta sostanzialmente di cambiare l'ordine delle righe e delle colonne della matrice in modo da portare i coef-

ficienti di similarita' piu' alti il piu' vicini possibile alla diagonale principale.

La tabella 5 mostra appunto la matrice della tabella 4 riordinata ("sorted") in modo da soddisfare a questo requisito, e si vede subito che vi sono tre "clusters": (2,5), (1,4) e (6,3) ad un livello piu' basso di correlazione.

Con questo metodo e' pero' difficile visualizzare bene le connessioni successive tra i "clusters" che sono stati identificati al piu' alto livello di correlazione. Quest'operazione diventa tanto piu' complicata quanto piu' alto e' il numero N di entita' da correlare, nel qual caso anche il riordinamento della matrice di partenza e' piuttosto laborioso, dato il gran numero di righe e di colonne della matrice di similarita' (l'esempio mostrato nelle tabelle 4 e 5 e' ovviamente molto semplice).

Piu' rapido, e soprattutto piu' chiaro anche visivamente, fino ai piu' bassi livelli di correlazione e' il metodo dei *dendogrammi*, largamente usato da vari autori per problemi di tassonomia (per una illustrazione dettagliata v. ad esempio Sokal e Sneath, 1963).

Qui accennero' solo ad uno dei metodi piu' semplici per costruirli, e cioe' il "weighted pair group method" con l'"average linkage". Si isolano innanzitutto i gruppi di due elementi che hanno coefficiente mutuo di similarita' piu' alto (nell'esempio della tabella (2,5), (4,1) e (6,3)). Si calcolano poi nuovi coefficienti di similarita', per gli elementi rimasti scorrelati come media dei vecchi coefficienti: se il "cluster" A e' formato da due elementi i e j con i quali l'elemento k aveva coefficienti di similarita' rispettivamente c_{ik} e c_{jk} , il nuovo coefficiente sara' $c_{Ak} = (c_{ik} + c_{jk})/2$. Nel caso di due "clusters" A e B, formati dagli elementi i e j ed l ed m rispettivamente, avremo per il nuovo coefficiente di similarita' $C_{AB} = (c_{il} + c_{im} + c_{jl} + c_{jm})/4$.

Nell'esempio della tabella 4, dove non ci sono elementi singoli scorrelati (i primi "clusters" sono 3 e gli elementi di

partenza 6), avremo tre nuovi coefficienti tra i tre "clusters" isolati al primo ciclo: quello tra (2,5) e (1,4) che vale 0.55, quello tra (1,4) e (6,3) che vale 0.47 e quello tra (2,5) e (6,3) che vale 0.25. Ripetendo il procedimento di partenza si isola un solo "cluster" con coefficiente di similarità mutuo più alto, e cioè ((2,5)-(1,4)), rispetto al quale l'elemento residuo (6,3) ha un coefficiente di similarità nuovo (che è l'ultimo) che vale 0.36.

Si ottiene così un diagramma ad albero (*dendrogramma*), che è rappresentato in fig.25. Il processo che è stato ora descritto diventa naturalmente piuttosto lungo se il numero delle entità da correlare N diventa grande, e pertanto ci si deve servire dell'aiuto di un calcolatore elettronico.

Già diversi di questi programmi sono in circolazione da tempo, e recentemente ne abbiamo realizzato uno (DENDR, Bietti et al., 1978) che costruisce e disegna materialmente dendrogrammi fino ad un massimo di 100 elementi da correlare (per un numero più grande di elementi il tempo macchina e l'occupazione della memoria diventano troppo grandi, per un calcolatore di media capacità), nel quadro delle attività del gruppo di ricerche interdisciplinari di metodi matematici applicati all'archeologia e alla paleontologia della Facoltà di Scienze dell'Università di Roma.

Mediante questo programma voglio qui illustrare brevemente alcuni esempi di applicazione concreta ad alcuni problemi archeologici.

Come primo esempio voglio citare brevemente uno dei risultati di uno studio preliminare sui complessi musteriani italiani, ed in particolare quelli pontiniani del Lazio, presentato al II° Congresso della Società Europea di Antropologia tenutosi a Brno (Cecoslovacchia) nell'agosto 1980 (Bietti, 1980).

Come abbiamo già visto, vi sono forti analogie tra il Pontiniano di Grotta Guattari e il Musteriano di tipo Quina dei

giacimenti del sud-ovest della Francia, almeno al livello di indici e di diagrammi cumulativi (Taschini, 1979).

Inoltre l'analisi del χ^2 ha mostrato come anche il livel-

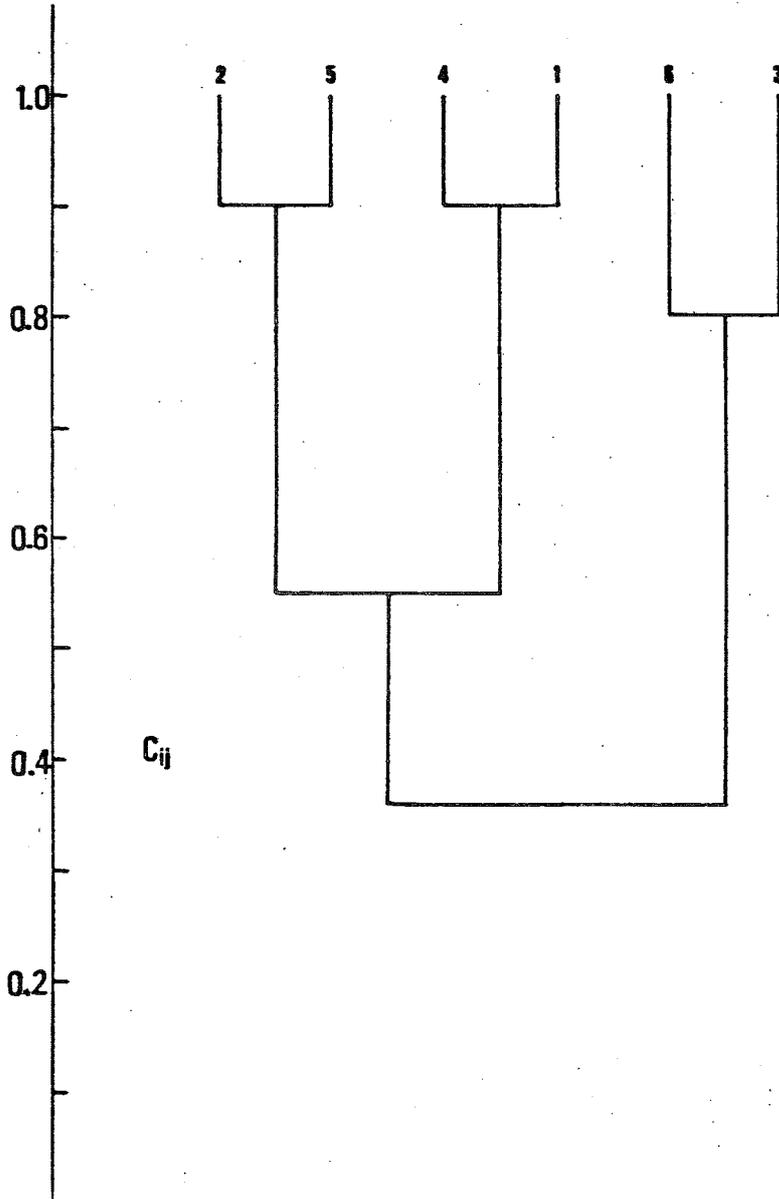


Fig.25

Dendrogramma ottenuto dai dati in tabella 4 (o tabella 5).

lo *d* del giacimento di Torre in Pietra, considerato dell'Interglaciale Riss-Wurm, presenti fortissime analogie con l'industria di Grotta Guattari (livello 2) (Fig.22).

Si è pertanto pensato di eseguire una "cluster analysis" tra questi giacimenti, per ora soltanto sulla base delle industrie, usando quindi un coefficiente di distanza percentuale (eq. 2.7), basato sulla lista "essenziale" di F. Bordes (1961) e cioè con schegge e punte Levallois aggiungendo inoltre quattro giacimenti italiani S. Agostino (Gaeta), Barma Grande ai Balzi Rossi in Liguria, Romanelli nel Salento, Torre Nave in Calabria) nonché il giacimento d'Erd, vicino Budapest, in Ungheria (Garbori-Csank, 1968).

In fig.26 si vede appunto il dendrogramma di tipo Q per questi giacimenti, ottenuto direttamente dal calcolatore tramite il già citato programma DENDR, e, più precisamente: 1) Guattari livello 1, 2) idem, livello 2, 3) idem, livello 4, 4) idem, livello 5, 5) S. Agostino livello A-A4, 6) Torre Nave livello 13, 7) Romanelli livello G3, 8) Barma Grande livello 2, 9) idem, livello 3, 10) idem, livello 4, 11) idem, livello 5, 12) Chinchon, 13) Erd e 14) Torre in Pietra livello *d*.

Come si vede, c'è un "cluster" Musteriano di tipo Quina-Pontiniano ben definito che consiste da una parte in una stretta associazione tra Guattari 2 (2) e Torre in Pietra *d* (14), già osservata con il metodo del χ^2 (ma sui soli raschiatoi!) ed a un livello più basso di correlazione con Romanelli G3, che è un Musteriano Quina su calcare (Pipernò, 1974), e dall'altra in un gruppo che comprende gli altri livelli di Guattari (1, 3, 4), S. Agostino (5), Chinchon (12) e l'Erd (13).

Ben separato da questo "cluster" abbiamo il complesso di Torre Nave (6) e Barma Grande livelli 3, 4, 5 (9, 10, 11), caratterizzato da un importante "debitage" Levallois (G.M. Bulgarelli, 1972; 1974) debolmente correlato con il livello 2 della Barma Grande (8) considerato Charentiano con debole "debitage" Levallois (G.M. Bulgarelli, 1974).

A parte il riconoscimento del carattere "Quina" del gruppo Pontiniano, già osservato da M. Taschini (1979), l'esame del dendrogramma in fig.26 pone, secondo me, interessanti problemi per l'interpretazione del Musteriano italiano.

Il secondo esempio riguarda un'analisi globale dei giacimenti mesolitici italiani, e cioè di quei giacimenti che, grosso

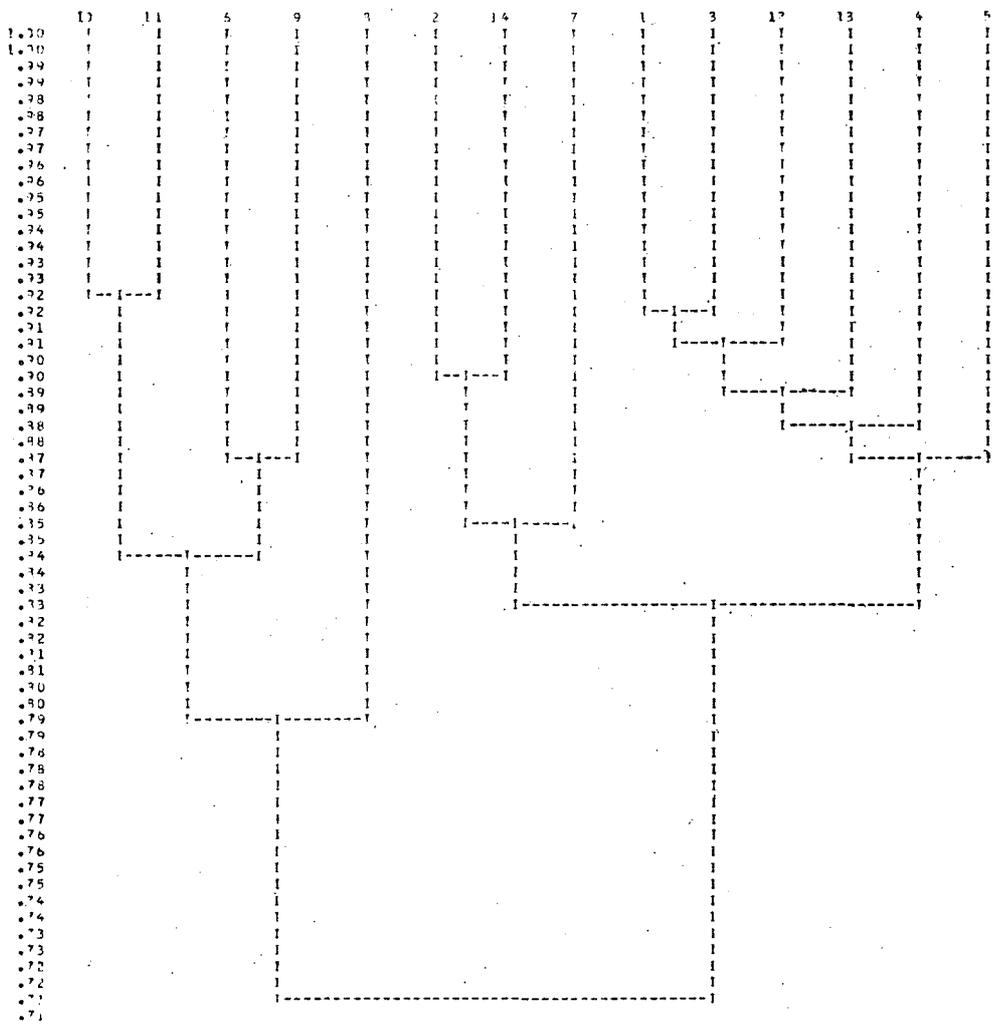


Fig. 26

Dendrogramma di tipo Q per giacimenti musteriani calcolato con le distanze percentuali (da Bietti, 1980).

1. Vatte 10-7
2. Romsignano III ACI
3. " " ABC
4. " " AB2-1
5. " " AA
6. Colbricon 1
7. Colbricon 8
8. Fienile Rossino
9. Provaglio
10. Benussi 6-5
11. " " 4
12. " " 3
13. Gr. Azzurra
14. Gr. Tartaruga
15. Arene Candide CIII
16. " " CII
17. " " CI
18. Arma Stefanin V-IV
19. Arma Nasino 13-11
20. Comunella
21. Lama Lite
22. Sammartina
23. Staz. Ortucchio I'
24. Gr. La Punta 27-21
25. Gr. Marizza IV-III
26. Capo d'Acqua
27. Ripoli 4
28. Feschio Ranaro
29. Rip. Blanc
30. Gr. Erica
31. Praia L 52-47
32. Praia I 46-44
33. Torre Testa
34. Gr. delle Hura C
35. Gr. Cavallo B2a
36. " " B1
37. Cipolliano 1
38. Gr. Corruggi
39. Levanzo
40. S.Basilio III

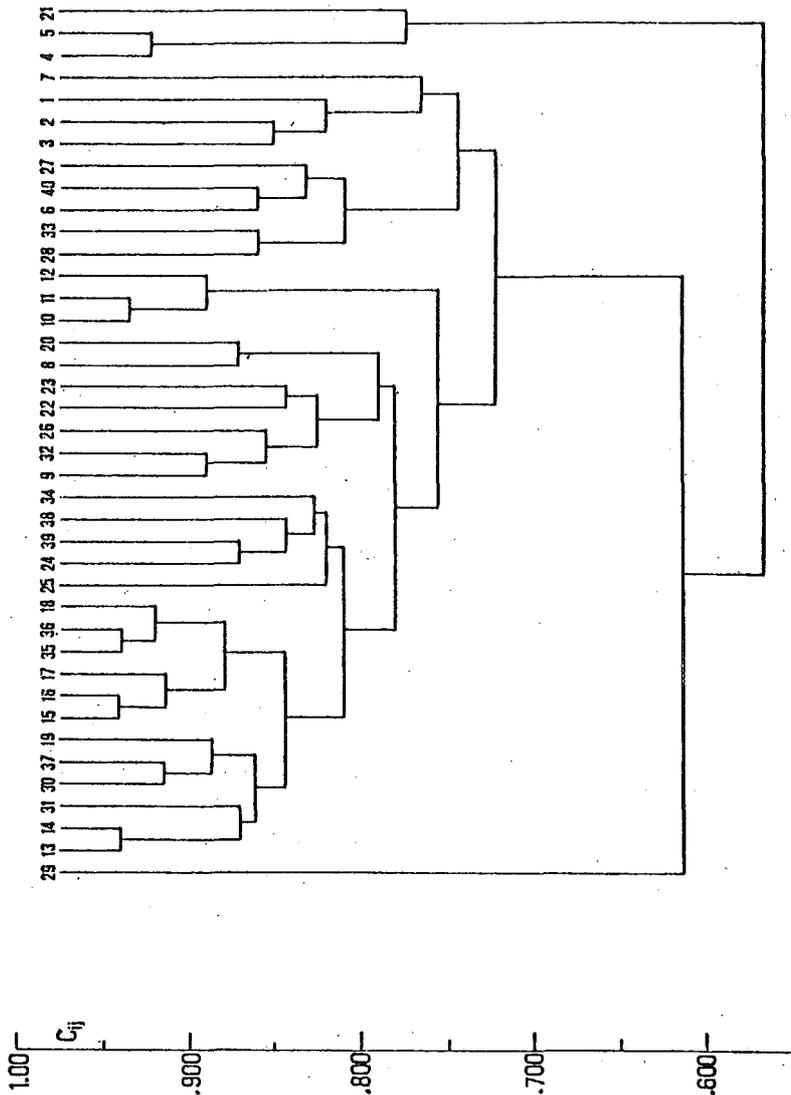


Fig.27

Dendrogramma di tipo Q per le industrie dei giacimenti Mesolitici Italiani, calcolato con le distanze percentuali (da Bietti, 1978).

modo, sono cronologicamente compresi tra la fine del Dryas III ed il Boreale, circa dal IX al VI millennio A.C. (Bietti, 1978).

Si è costruito innanzitutto un dendrogramma basato sulle sole industrie dei giacimenti, usando il coefficiente di distanza percentuale definito dall'eq. (2.7), ed il risultato è mostrato in fig.27. Come si vede, non c'è una chiara indicazione di "clustering" tra giacimenti geograficamente vicini. Grotta Azzurra (15) e Grotta della Tartaruga (14), che si trovano nel Carso Triestino, sono correlate con grotta della Madonna a Praia, in Calabria, (31), e poi con grotta Erica (30) a Positano, in Campania, Arma di Nasino (19), in Liguria, e le grotte Cipolliane 1 (37); nel Salento, ma non sono per niente correlate con il vicino sito di Grotta Benussi (10,11,12). Gli strati inferiori (10,11) di quest'ultimo giacimento, definiti "Sauveterroidi" (Broglia, 1971), non sono affatto correlati con il gruppo "Sauveterroide" della valle dell'Adige, formato da Romagnano III (2,3) e da Vatte di Zambana (1). Lo strato superiore (12) di grotta Benussi, definito "Tardenoide" (Broglia, 1971), non è per niente correlato con il complesso "Tardenoide" degli strati superiori di Romagnano III (4,5), che forma invece un gruppo totalmente a sé, insieme con il giacimento di alta montagna di Lama Lite nell'Appennino Tosco-emiliano.

Si è pertanto cercato di stabilire delle correlazioni tra i giacimenti anche da un punto di vista ecologico-ambientale: oltre all'industria si sono pertanto considerati come caratteri anche le faune dei giacimenti, divise in macromammiferi, micromammiferi, uccelli, pesci, molluschi d'acqua dolce, molluschi terrestri, marini di laguna salmastra e di scoglio; in più si sono considerati anche il tipo di sito (grotta, riparo, ecc.) e la sua posizione (costiera, interna, ecc.).

Impiegando questa volta il coefficiente di correlazione dato dall'eq. (2.6), si è ottenuto il dendrogramma in fig.28, dove si possono osservare innanzitutto alcuni "clusters" (considerando tali quelli con $r_{ij} > 0.4$) geografici, come quello del-

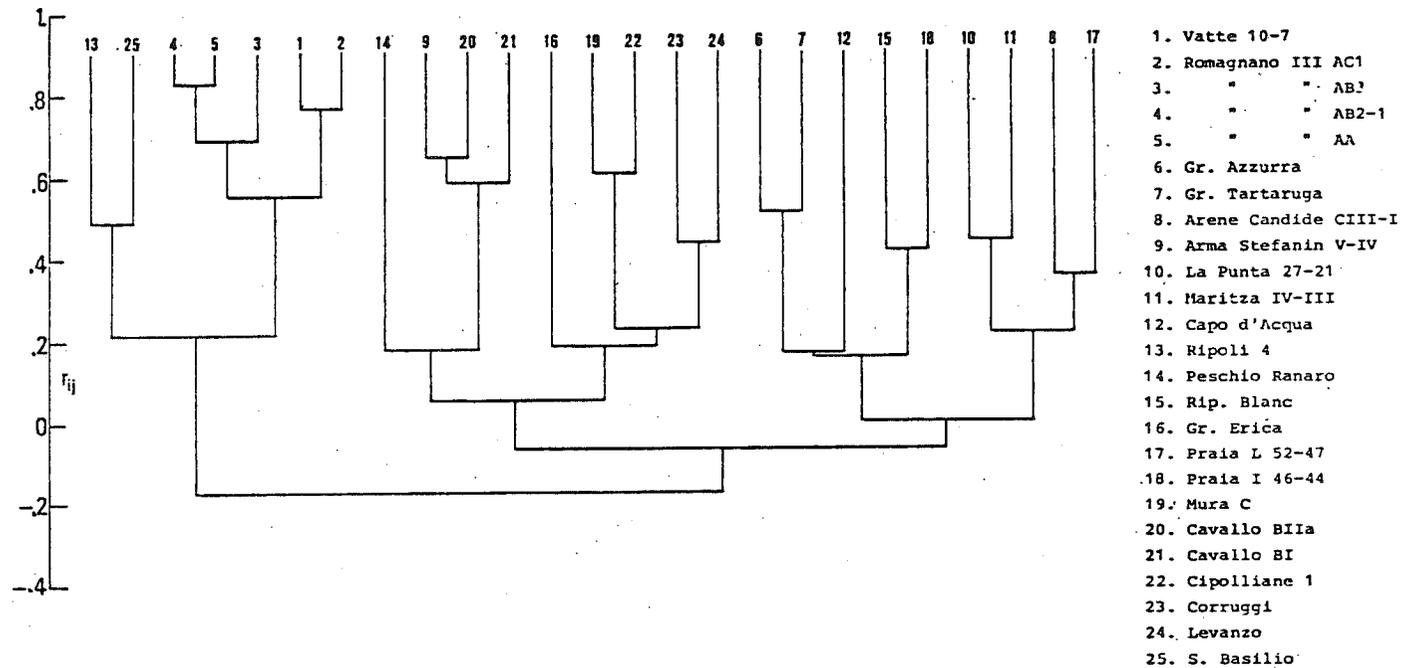


Fig. 28

Dendrogramma di tipo Q per i giacimenti Mesolitici Italiani, ottenuto mediante il coefficiente di correlazione lineare (da Bietti, 1978).

la valle dell'Adige (1,5), quello del Carso triestino (6,7), quello del bacino del Fucino in Abruzzo (10,11), quello della costa pugliese (19,22), ed infine un "cluster" Siciliano (19, 22).

L'aspetto interessante degli altri "clusters" è che, eccettuato uno (9,20,21) che collega un giacimento in posizione interna della Liguria (Arma dello Stefanin) con due strati di giacimento costiero del Salento (Grotta del Cavallo), tutti gli altri correlano giacimenti costieri (Arene Candide (8) in Liguria con Praia L (17) in Calabria, ad esempio), o in posizione interna (Ripoli (13) in Abruzzo e S.Basilio (25) in Sicilia, ad esempio).

L'uso di un altro coefficiente di similarità (quello cosiddetto del coseno dell'angolo) (fig.29), lascia praticamente invariata la situazione, con la sola eccezione del riparo Blanc al Monte Circeo (15), che ora non è più correlato con il livello superiore di Praia (18) (che ora si aggiunge al "cluster" (8,17)), ma resta a basso coefficiente di similarità con tutti i "clusters", come già erano il giacimento di Peschio Ranaro (14), in provincia di Frosinone, il giacimento di Capo d'Acqua (12), in provincia dell'Aquila e Grotta Erica (16), sulla costa in provincia di Salerno.

Mentre i "clusters" geografici hanno una interpretazione abbastanza ovvia come espressioni di gruppi culturali ben determinati, è ragionevole interpretare quelli non geografici come attività simili da parte di gruppi culturali anche diversi, e i giacimenti scorrelati come siti di *attività specializzate*, forse stagionali (Bietti, 1978), come Peschio Ranaro per la caccia ad animali di montagna (lo stambecco) e da pelliccia, o come riparo Blanc per la raccolta intensiva di molluschi marini di scoglio.

Questa interpretazione è in accordo con l'ipotesi delle culture mesolitiche come insieme di processi culturali che portano le comunità di cacciatori della fine del Paleolitico ad un

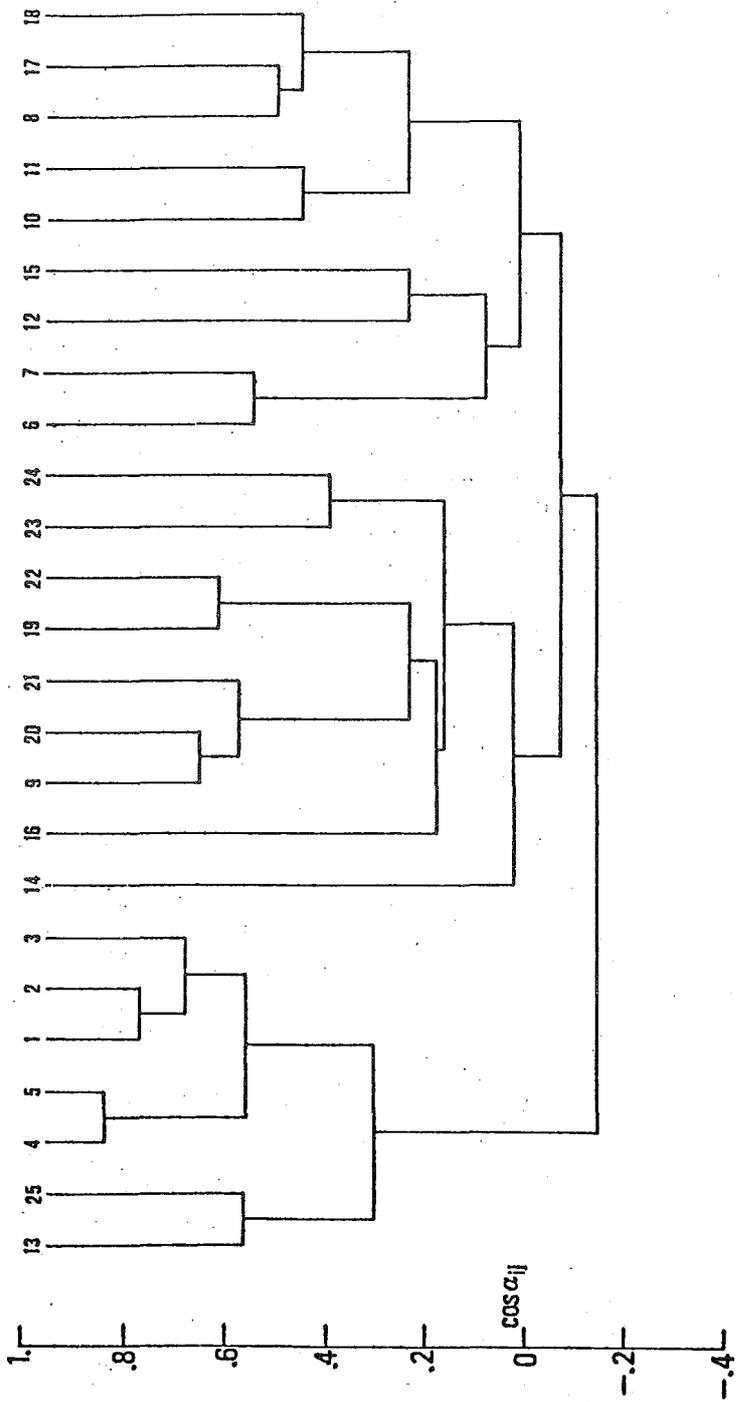


Fig.29
Come in fig.13, ma con il coefficiente del coseno dell'angolo (da Bietti, 1978).

ampliamento e ad un arricchimento, dovuto in parte ai grandi cambiamenti climatici in atto, delle loro attività di sussistenza, ipotesi avanzata da M. Taschini alcuni anni fa (Taschini, 1968).

Si è cercato di vedere meglio quest'aspetto di attività specializzate costruendo dei dendrogrammi di tipo R.

In fig.30 si vede il dendrogramma calcolato per i giacimenti costieri con il coefficiente di correlazione (eq.(2.6)), e, a parte alcuni "clusters" negativi (e cioè di oggetti che sono scarsamente rappresentati nei vari giacimenti, come ad esempio i micromammiferi (16), gli uccelli (17) ed i molluschi terrestri (19)), sono interessanti ad esempio il "cluster" tra microbulini (14), geometrici (10) e troncature (5), e soprattutto quello tra punteruoli (2), intaccature e denticolati (8) e molluschi marini di scoglio (21), in accordo con l'ipotesi di un uso specializzato di questi strumenti per il distacco di questi molluschi (specialmente *Patellae*) dalle pareti rocciose. E' infatti significativa, nel già citato riparo Blanc, la grande quantità di questi tipi di strumenti (più del 60% del totale) in concomitanza di un'attività economica quasi esclusivamente dedita alla raccolta di molluschi marini (*Trochus* e *Patella*).

Correlazioni funzionali connesse con l'attività di caccia, come quella tra grattatoi (1) e macrofauna (15), sono invece evidenti nel dendrogramma di tipo R per i giacimenti in posizione interna, sempre calcolato con il coefficiente di correlazione (eq.(2.6)) (fig.31). Anche qui abbiamo la correlazione "litotecnica", se così si può definire, tra geometrici (10) e microbulini (14), e "clusters" negativi, cioè che correlano oggetti poco rappresentati nei giacimenti, come gli strumenti "esquilles" e diversi (13) e i molluschi terrestri.

Come esempi successivi di applicazione di "cluster analysis", sempre eseguita con il programma DENDR, voglio invece mostrare il caso di complessi archeologici più recenti, e cioè

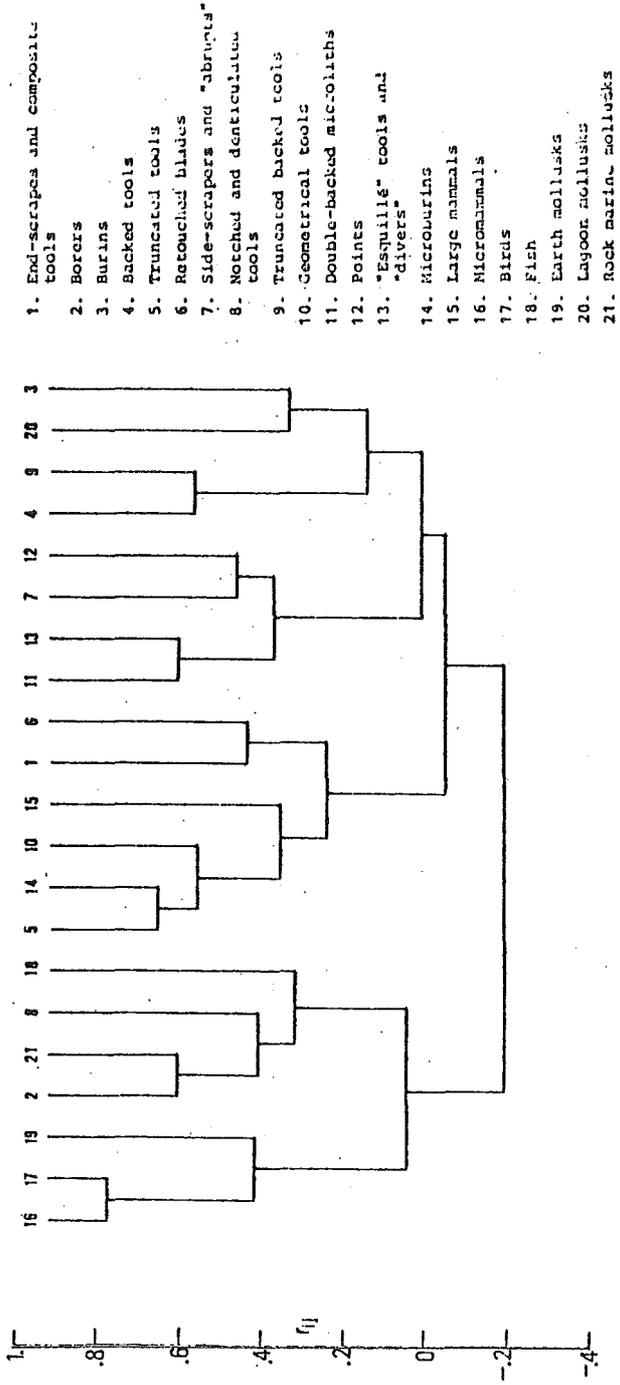


Fig. 30

Dendrogramma di tipo R per i giacimenti costieri del Mesolitico italiano, ottenuto mediante il coefficiente di correlazione lineare (da Bietti, 1978).

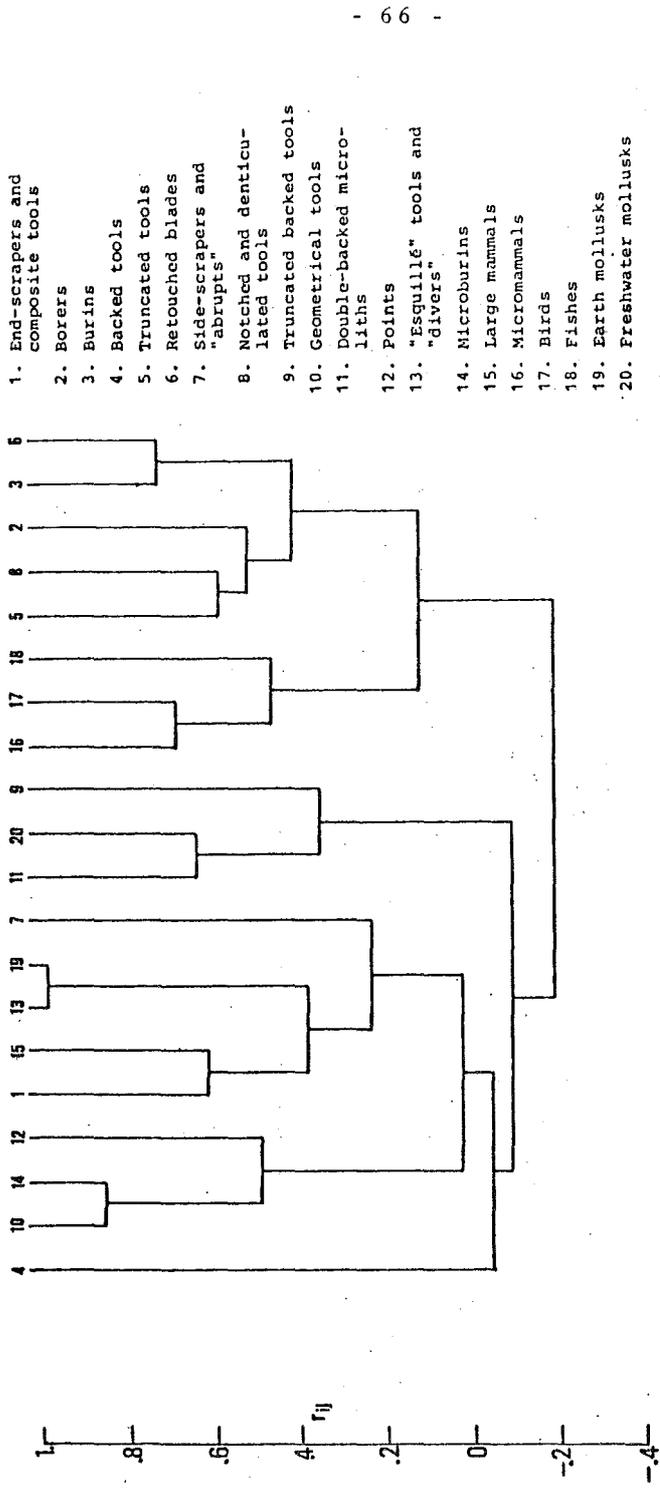


Fig.31

Dendrogramma di tipo R per i giacimenti in posizione interna del Mesolitico italiano, ottenuto mediante il coefficiente di correlazione lineare (da Bietti, 1978).

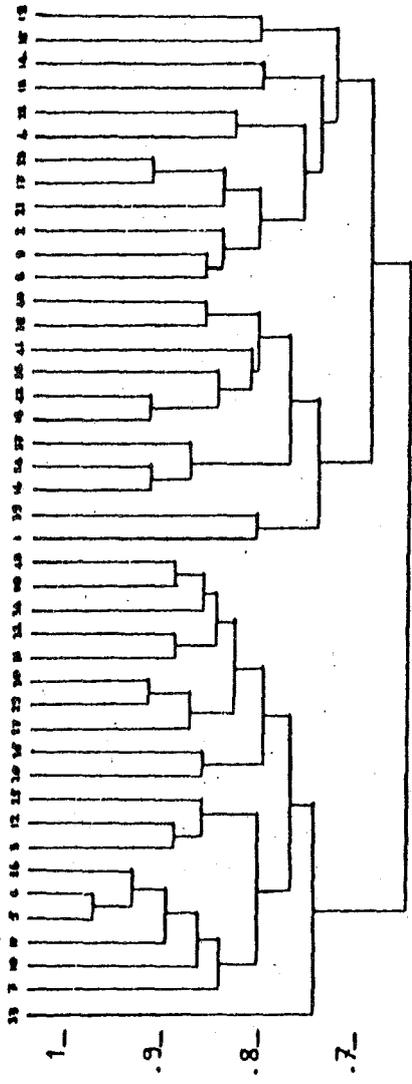
necropoli protostoriche e classiche, proprio allo scopo di mettere in evidenza le ampie possibilita' di questi metodi, per lo studio di complessi di queste epoche.

I risultati che presentero' sono in corso di pubblicazione e debbo ringraziare le Dr.sse A.M. Bietti-Sestieri, G. Bartoloni e A. Greco-Pontrandolfo per avermi fornito i dati. Naturalmente si tratta di risultati preliminari, e conclusioni piu' ampie si potranno trovare nei lavori specializzati degli Autori.

Il primo esempio e' uno studio delle tombe ad incinerazione della prima e seconda fase della cultura laziale (Bietti-Sestieri e Bergonzi, 1979; Bietti-Sestieri, 1979). La fig.32 da' il dendrogramma di tipo Q per 43 tombe ad incinerazione della I fase Laziale di varie localita'. Per ogni tomba sono stati considerati 36 attributi, a partire dalla presenza o assenza di dolio o dell'urna a capanna fino a tutti gli oggetti del corredo, elencati in fig.33, che da' invece il dendrogramma di tipo R. Entrambi i dendrogrammi sono stati calcolati mediante la semplice presenza o assenza degli attributi, usando il coefficiente di Sokal e Michener, definito dall'eq.(2.5).

In fig.32 vediamo che vi sono sostanzialmente tre grandi "clusters". Il primo a partire da sinistra e' in realta' composto (a parte la tomba di Villa Cavalletti 7 (33), sostanzialmente isolata) da due "clusters": il primo considerato piuttosto antico (salvo Villa Cavalletti 1 (26)), formato da tombe sia maschili che femminili del Foro Romano (7,10,11,12), Boschetto (5), Anzio (16) e Lavinio (3,25) con oggetti miniaturizzati nei corredi; il secondo da tombe piu' tarde cronologicamente, sempre miste come sesso, essenzialmente del Foro (27,32) Lavinio (20,35), Villa Cavalletti 4 (34) e Osteria dell'Osa T56 (43), caratterizzate sostanzialmente dall'assenza di oggetti miniaturizzati.

Il secondo grande "cluster" e' composto sostanzialmente da tombe dell'Osteria dell'Osa (v. per esempio tav.Ia) (36-42), piu' Lavinio 21, Villa Cavalletti 6, Foro Y e Palatino (1,16,



- S M_{ij}
- | | | |
|-----------------|-----------------|----------------------|
| 1 LAV. 21 | 19 FORO C | 37 OST. OSA AS 53-54 |
| 2 S. LORENZO V. | 20 LAV. 16 | 38 OST. OSA AFA |
| 3 LAV. 18 | 21 FORO U | 39 OST. OSA A 54-55 |
| 4 LAV. 24 | 22 FORO G | 40 OST. OSA T 55-56 |
| 5 BOSCHETTO | 23 FORO A | 41 OST. OSA TU 57 |
| 6 ANZIO 1A | 24 FORO V | 42 OST. OSA T 57 |
| 7 FORO 1 | 25 LAV. 15 | 43 OST. OSA T 56 |
| 8 ANDREOLI | 26 V. CAV. 1 | |
| 9 V. D'ANDREA | 27 FORO X | |
| 10 CAMPOFATTORE | 28 FORO V | |
| 11 FORO 3 | 29 FORO S | |
| 12 FORO 2 | 30 FORO T | |
| 13 FORO 4 | 31 FORO A | |
| 14 LAV. 7 | 32 FORO G-G | |
| 15 V. CAV. 2 | 33 V. CAV. 3 | |
| 16 V. CAV. 6 | 34 V. CAV. 4 | |
| 17 V. CAV. 5 | 35 LAV. 33 | |
| 18 PALATINO | 36 OST. OSA 55A | |

Fig. 32
 Dendrogramma di tipo Q per le tombe ad incinerazione della prima fase della cultura Laziale, ottenuto mediante il coefficiente di Sokal e Michener.

- 1 - dolio
- 2 - urna a cap.
- 3 - olle in coppia
- 4 - calefattoio
- 5 - askos
- 6 - vaso barchetta
- 7 - piattello trepp.
- 8 - tazza carenata m.
- 9 - tazza profonda m.
- 10 - scodella orlo rient. m.
- 11 - scodella biancata
- 12 - scodella troncoc.
- 13 - olletta orlo rientr.
- 14 - vaso a ciessidra
- 15 - orciolo globulare
- 16 - orciolo biconico
- 17 - anfora m.
- 18 - anfora
- 19 - tazza carenata
- 20 - tazza profonda
- 21 - scodella orlo rientr.
- 22 - scodella villanov.
- 23 - bocca-s
- 24 - fuseruola
- 25 - statuetta
- 26 - vaso biancato m.
- 27 - vaso biancato
- 28 - lancia
- 29 - rasoio
- 30 - fib. arco serp.
- 31 - fib. arco gomito
- 32 - fib. arco semp.
- 33 - scudi
- 34 - coltello
- 35 - perline, ornamenti
- 36 - tazza bassa e larga

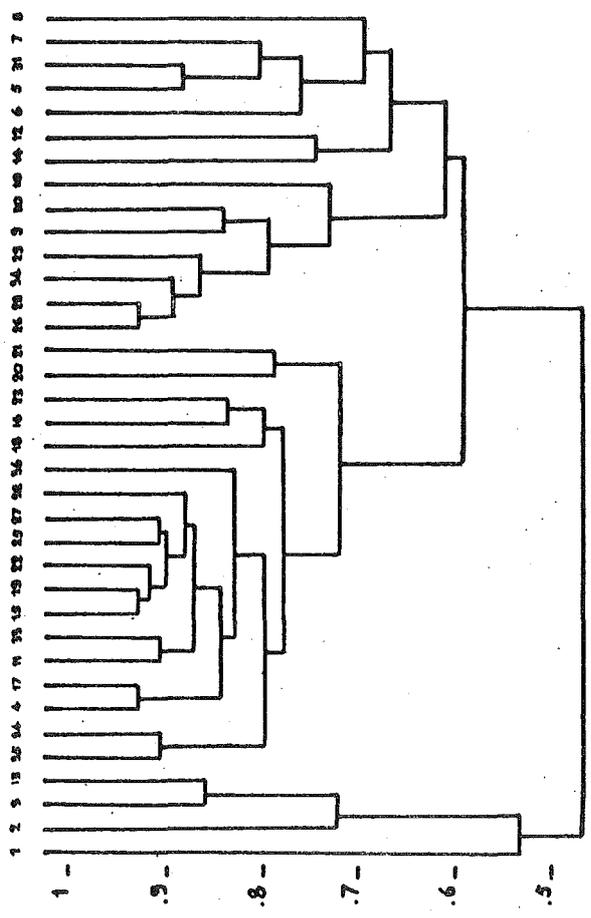


Fig.33

Dendrogramma di tipo R per le tombe ad incinerazione della I fase della cultura Laziale, ottenuto con il coefficiente di Sokal e Michener.

24,18). Si tratta, secondo l'Autore, di tombe maschili, piuttosto tarde cronologicamente, con l'eccezione di Lavinio 21 (1), che è più antica, non a caso correlata con la R54 bis dell'Osteria dell'Osa (39), piuttosto singolare nell'insieme delle tombe di questa necropoli, ed entrambe correlate a coefficiente di similarità più basso con il resto del "cluster".

Il terzo grande "cluster" è piuttosto composito, e comprende sia tombe femminili più antiche, come Andreoli, Vigna d'Andrea e S. Lorenzo Vecchio (8,9,2) e tombe miste più recenti (Foro U,A e Villa Cavalletti 5) (21,23,17), correlate a livello di similarità più basso con tombe miste ad esempio tombe del Foro (13,19,22), Lavinio (4,14) e Villa Cavalletti (15).

Secondo l'Autore, si nota una omogeneità culturale abbastanza evidente, sia in senso cronologico che spaziale, con una differenziazione nella fase più recente, legata in parte al sesso, come mostra il "cluster" con rituale maschile tipico dell'Osteria dell'Osa.

Per curiosità del lettore oltre al dendrogramma disegnato a mano nella fig.32, in fig.34 si può vedere lo stesso dendrogramma disegnato direttamente dal calcolatore mediante il già citato programma DENDR (Bietti et al., 1978).

L'analisi di tipo R (fig.33), oltre ad alcuni "clusters" tra oggetti molto rari in quasi tutte le tombe, come ad esempio l'orciolo globulare, la tazza carenata e la scodella villanoviana (15,19,22), mostra alcuni "clusters" di oggetti riferibili al sesso del defunto, come la fuseruola e le perline (24,39), l'askos e la fibula ad arco a gomito (5,31) (femminili) e il vaso biansato e la tazza profonda miniaturizzata, la lancia, il coltello ed il rasoio miniaturizzato e la fibula ad arco serpeggiante (26,9,28,34,29,30) (maschili).

È interessante, secondo l'Autore, notare alcune correlazioni che denotano associazioni funzionali: le olle in coppia con l'olletta ad orlo rientrante (3,13), correlate a più basso livello di similarità con l'urna a capanna (2), il vaso a cles-

Un secondo esempio di interesse protostorico riguarda alcune (18) tombe integre della necropoli Villanoviana di Quattro Fontanili di Veio (Roma) (ricerche in corso di G. Bartoloni), fornite dall'Autore a titolo dimostrativo (v. per esempio Tav. Ib).

Gli oggetti di corredo sono 43, che non specificheremo in questa sede, dato che il nostro esempio riguarda solo analisi di tipo Q, i cui risultati si possono vedere nel dendrogramma di fig.35, calcolato sempre con il coefficiente di Sokal e Michener (eq.2.5)).

A partire da sinistra, il "cluster" (10,1) consiste in due tombe femminili particolarmente ricche, separato sia dal "cluster" (8,9), rappresentato da tombe maschili (una di guerriero e l'altra senza armi, ma con molti elementi di ceramica locale in comune), che da quello rappresentato da tombe femminili di media ricchezza (2,3,6). Quest'ultimo "cluster" si trova correlato, a livello più basso, con un altro piuttosto composito (12,5,7,17).

Infine vi è un "cluster" di tombe maschili piuttosto ricche (11,13,14,16,18) (anche se alcune con basso coefficiente di similarità) di cui alcune (14,13,15,18) hanno anche il carro ed i morsi di cavallo in bronzo.

Questa possibile differenziazione di "status" e di ricchezza per entrambi i sessi, anche se il numero di tombe considerato è molto limitato, sembra essere messa in luce anche dal dendrogramma eseguito con il coefficiente di Jaccard (che non tiene conto delle assenze contemporanee di oggetti di corredo in due tombe, come si è già detto), rappresentato in fig. 36. Come si vede, i primi tre "clusters" della fig.35 sono ancora presenti: (10,1), (8,9) e (2,3,6). Il "cluster" di tombe maschili ricche è ora sostanzialmente diviso in due: (11,14) (già isolato nel dendrogramma precedente) e (13,15,18). Il "cluster" indistinto di fig.35 (12,5,7,17) è ora smembrato tra altri "clusters", con bassi valori del coefficiente di similarità.

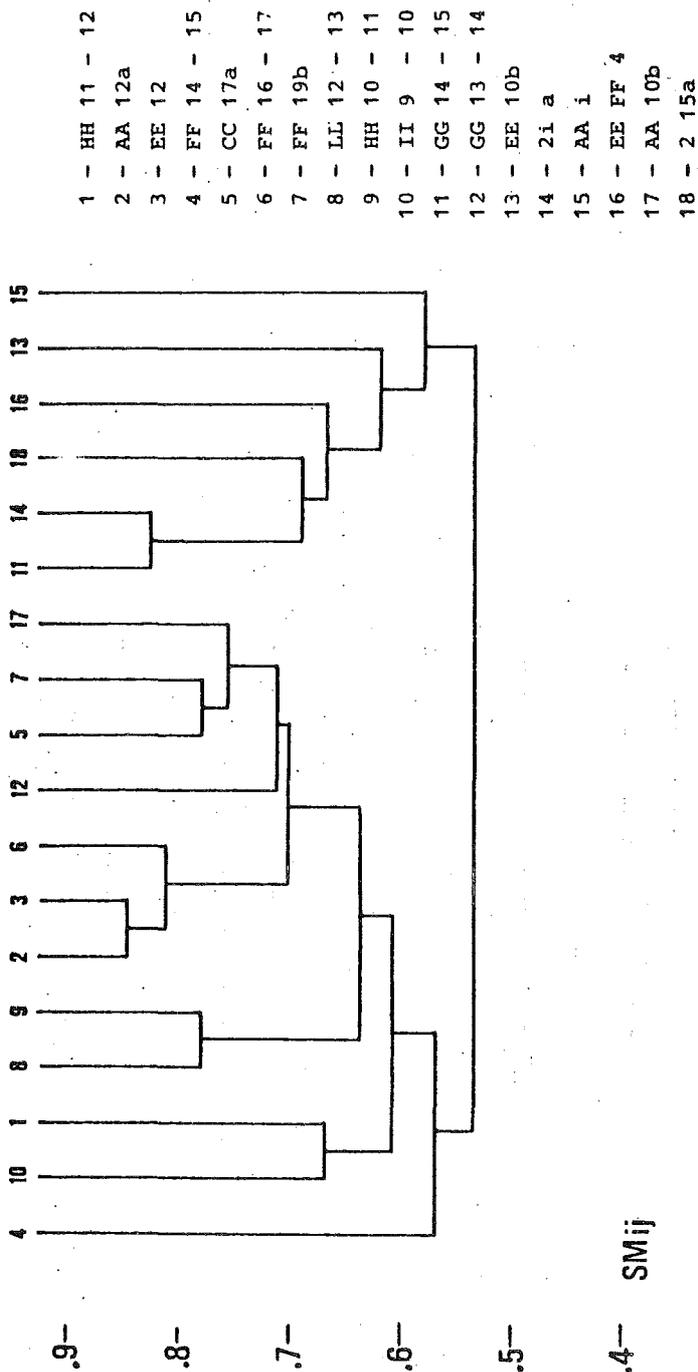


Fig.35

Dendrogramma di tipo Q per alcune tombe della necropoli di Quattro Fontanili di Veio, ottenuto mediante il coefficiente di Sokal e Michener.

ta' (da notare che, per sua stessa definizione, il coefficiente di Jaccard ha in generale valori piu' bassi rispetto a quelli di Sokal e Michener, a parita' di unita' da correlare).

Come esempio di "cluster analysis" applicata ad un problema di archeologia classica, in fig.37 si puo' vedere il dendrogramma di tipo Q per 37 tombe di sesso misto (riconoscibile solo dai corredi, perche', data la natura del terreno, gli scheletri non sono conservati) di una necropoli Lucana del IV secolo A.C. di Paestum (Salerno) (ricerche in corso di A. Greco-Pontrandolfo; v. anche Greco-Pontrandolfo, 1977) (v. per esempio Tav.II).

Il dendrogramma e' calcolato questa volta sulla base del coefficiente di correlazione (eq.(2.6)) dato che gli oggetti di corredo sono stati considerati dall'Autore come attributi a

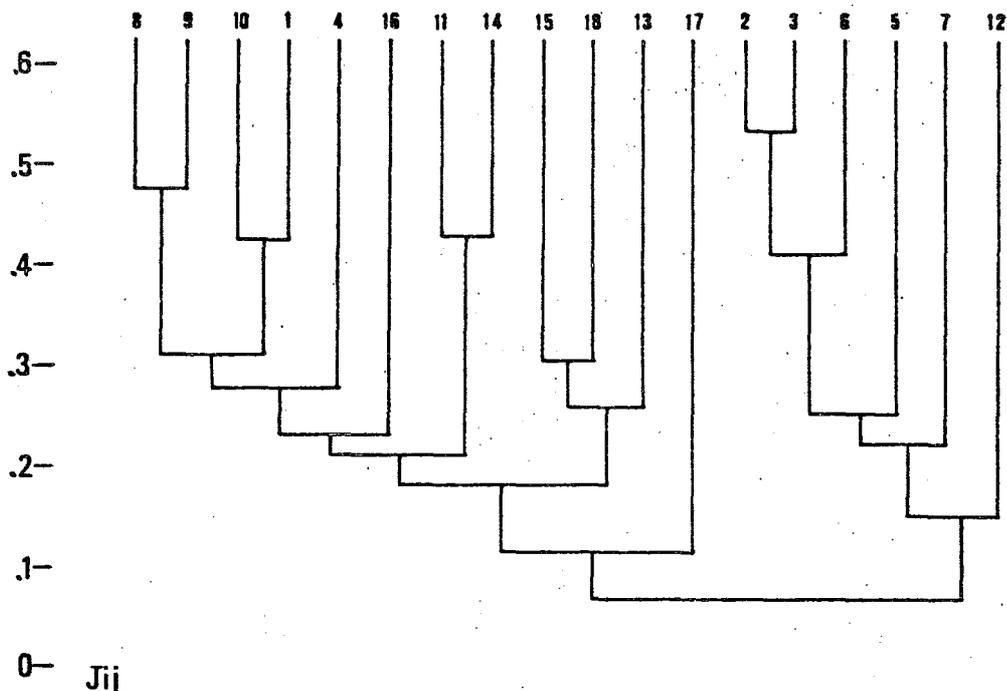


Fig. 36

Come in fig.35 ma con il coefficiente di Jaccard.

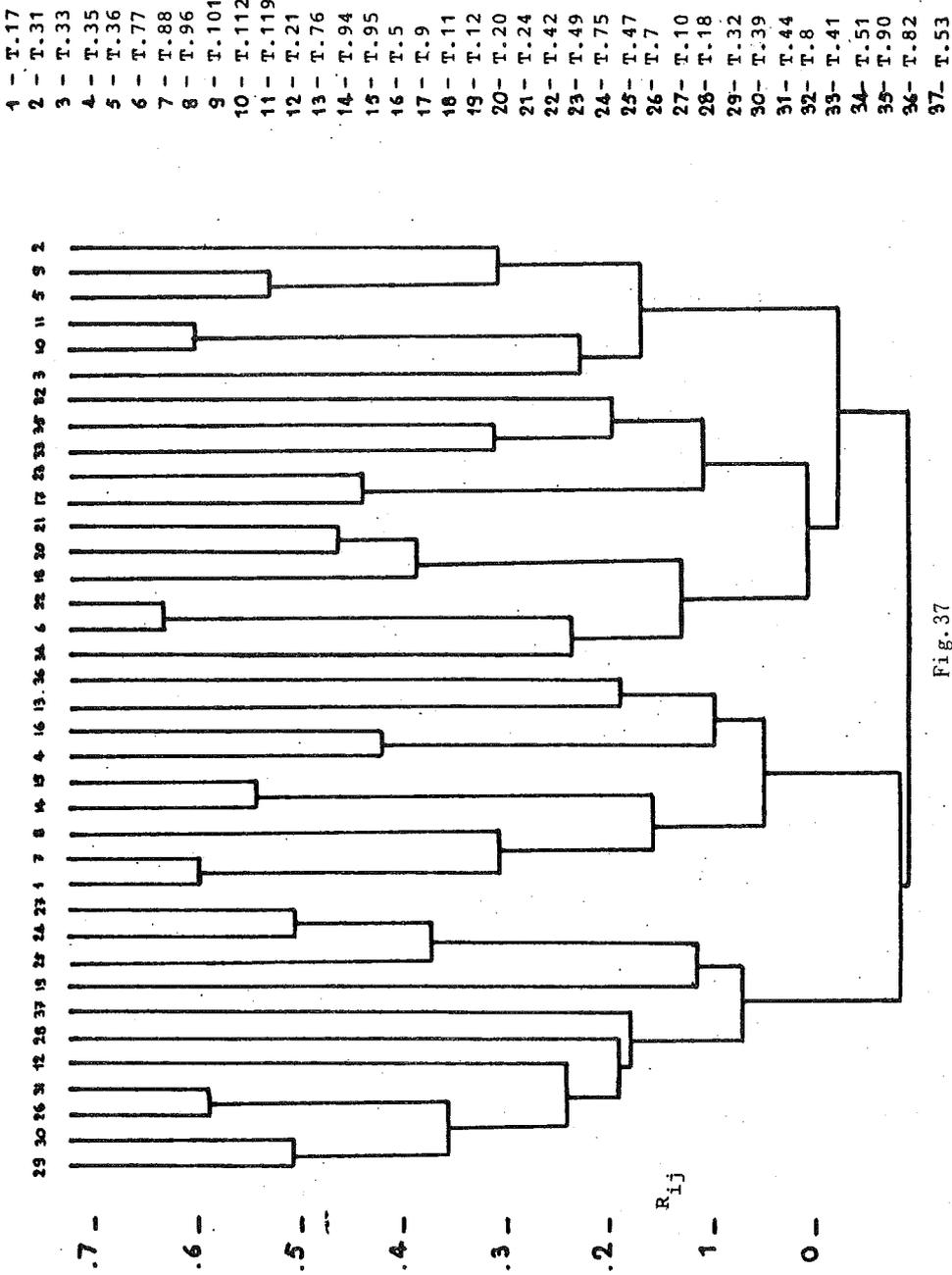


Fig.37

Dendrogramma di tipo Q per le tombe della necropoli Lucana di Paestum, ottenuto mediante il coefficiente di correlazione lineare.

quattro stati (assente, presente, frequente, molto frequente).

A partire da sinistra in fig. 37 abbiamo un primo gruppo di tombe, composto innanzitutto da un "cluster" di tombe femminili (29,30,26,27) legate però, con basso valore del coefficiente di similarità con altre tombe femminili (12,28,37). L'altra "meta" del gruppo è composta da un "cluster" sempre di tombe femminili (25,24,27) legate però, con basso valore del coefficiente di similarità, ad una tomba maschile (19). Secondo l'Autore, vi è una probabile differenza cronologica e sociale tra i "clusters" di questo primo gruppo, da controllare meglio in futuro sia con l'analisi di nuove tombe, che con dendrogrammi di tipo R.

Il secondo gruppo di tombe consiste in tombe ibride, in maggioranza maschili (per esempio i "clusters" (1,7,8) e (14,15)), con l'eccezione di due tombe femminili (13,36).

Il terzo gruppo, sulla destra in fig. 37, è composto di sole tombe maschili, e consta di tre "clusters", interessanti, secondo l'Autore, perché rappresentano distinzioni sia cronologiche che di "status" sociale.

Le tombe del primo "cluster" ((34,6,22) e (18,20,21)) sono più antiche: contengono essenzialmente armi e scarso corredo ceramico, anche se le tombe (18,20,21) hanno un corredo di armi che denota più prestigio di quello delle tombe 6 e 22, ad esempio.

Nel secondo "cluster", considerato dall'Autore più recente, formato dalle tombe (17,23) e (33,35,32), le armi sono assenti (solo il coltello è presente), mentre il corredo ceramico aumenta in quantità e tipi (sono presenti i vasi da mensa).

L'ultimo "cluster" del terzo gruppo è rappresentato dalle tombe (3,10,11) (quest'ultima è dipinta), correlate a livello più basso di similarità con le tombe (5,9,2), dove, oltre al corredo ceramico, ricompaiono le armi.

L'Autore ritiene che questi primi risultati siano molto incoraggianti, e che un'analisi di tipo R, nonché un'analisi di

tipo Q estesa anche ad altre tombe, possa contribuire a chiarire ancor meglio sia la sequenza cronologica della necropoli che la differenziazione di "status" sociale, sia per le tombe maschili che per quelle femminili.

* * *

I pochi dendrogrammi di tipo R che abbiamo visto hanno un carattere essenzialmente qualitativo: non c'è modo di sapere *quantitativamente*, ad esempio, un "cluster" tra denticolati, punteruoli e molluschi marini di scoglio del Mesolitico (fig.30) in che percentuali è composto da questi caratteri, e in che misura i vari giacimenti considerati contribuiscono al "cluster".

Il metodo delle componenti principali, con le sue varianti, come ad esempio l'analisi dei fattori ("factor analysis") è in pratica quello più usato a questo proposito, e cercherò qui di illustrarlo in modo molto semplice.

Supponiamo di avere due grandezze da misurare X_1 e X_2 (la larghezza e la lunghezza di una serie di strumenti di selce, e così via) che mi darà al variare dei campioni, il solito "scatter-plot". Calcoliamo ora le medie di ognuna delle grandezze definite al solito modo, le varianze e la *covarianza* tra le due grandezze definite, come è noto da

$$\sigma_{12}^2 = \frac{1}{N-1} \sum_i (\bar{X}_1 - X_1^{(i)}) (\bar{X}_2 - X_2^{(i)}) \quad (2.8)$$

dove ovviamente $X_1^{(i)}$ e $X_2^{(i)}$ rappresentano l'*i*-esima osservazione per la grandezza 1 e quella 2.

Se le due grandezze fossero completamente scorrelate è ovvio che la covarianza sarebbe zero. Supponiamo di avere allora ad esempio $\sigma_1^2 = 6$, $\sigma_2^2 = 2$ e $\sigma_{12}^2 = 3$. Graficamente si può rappresentare la situazione in un piano in cui i due assi siano le varianze (fig.38).

I "vettori" indicati con 1 e 2 nella figura sono rispettivamente quello che ha come ascissa σ_1^2 e ordinata σ_{12}^2 , e quel-

lo che ha come ascissa σ_{12}^2 e ordinata σ_2^2 .

Si puo' costruire una *matrice* varianze-covarianze che e' cosi' costituita

$$A = \begin{pmatrix} 6 & 3 \\ 3 & 2 \end{pmatrix}$$

Lo scopo dell'analisi in componenti principali e' quella di "disaccoppiare" le due grandezze e cioe' di arrivare a una

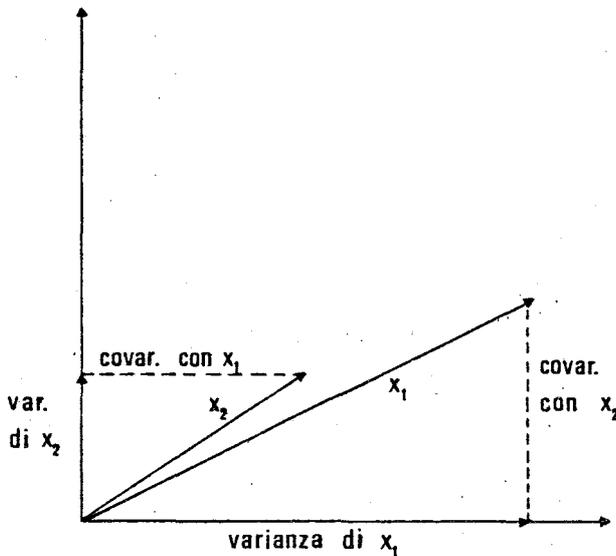


Fig. 38

Rappresentazione grafica delle varianze e covarianze tra due grandezze X_1 ed X_2 .

combinazione di esse, o meglio a due di esse tali che la rispettiva covarianza sia zero. In altre parole si tratta di effettuare una rotazione di assi, e cioe' cercare un nuovo sistema di riferimento, dove la matrice diventi *diagonale* e cioe' gli elementi al di fuori della diagonale principale (3 e 3 nella matrice A, siano nulli).

Questo e' un classico problema di *autovalori* e *autovettori*

di una matrice (simmetrica nel nostro caso, come in tutti i casi di varianza-covarianza) che si risolve tramite l'equazione secolare

$$(6-\lambda) \cdot (2-\lambda) - 9 = 0 \quad (2.9)$$

che dà i due valori $\lambda_1 = 4 + \sqrt{13} = 7,6$ e $\lambda_2 = 4 - \sqrt{13} = 0,4$, così che la matrice diagonalizzata è ora

$$D = \begin{pmatrix} 7,6 & 0 \\ 0 & 0,4 \end{pmatrix}$$

Notare che ovviamente (proprietà degli autovalori) la varianza totale è la stessa (8) sia per la matrice A che per la matrice D.

Gli autovettori, che sono in pratica i nuovi assi coordinati hanno componenti, nel vecchio sistema di riferimento, rispettivamente 0,88 e +0,46 e 0,46 e -0,88, che sono appunto l'inclinazione rispetto ai vecchi assi X_1 e X_2 , come si vede nella fig.39 che mostra i sistemi di riferimento e l'ellisse definita dalle varianze e covarianze: le varianze delle nuove grandezze, date dalla matrice D, sono proprio gli assi principali dell'ellisse.

Il risultato è che abbiamo ora due nuove variabili Y_1 e Y_2 legate alle precedenti da $Y_1 = 0,88 X_1 + 0,46 X_2$ e $Y_2 = 0,46 X_1 - 0,88 X_2$, e che hanno le varianze date dagli autovalori e covarianze nulle.

In sostanza queste relazioni dicono solo in che percentuale le nuove variabili sono costituite dalle vecchie, e possono essere invertite, e cioè trasformate in relazioni che danno le vecchie variabili in termini delle nuove il che ovviamente ci interessa molto di più, visto che le nuove variabili sono scorrelate.

Nel nostro caso abbiamo $X_1 = 0,88 Y_1 + 0,46 Y_2$ e $X_2 = 0,46 Y_1 - 0,88 Y_2$ (che sono le stesse relazioni di prima per-

che' la "matrice" degli autovettori e' simmetrica).

E' importante notare che gli assi principali non hanno niente a che vedere con i coefficienti di regressione tra due variabili, ed in particolare con il coefficiente b , che rappresenta la pendenza della retta. In fig.40 si vede appunto, in due dimensioni, l'ellisse che rappresenta l'aspetto generale

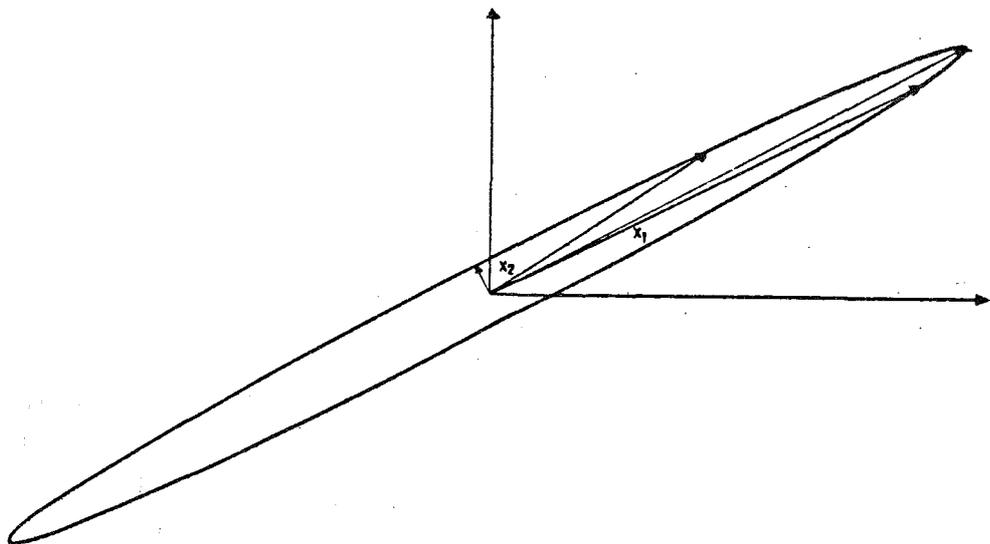


Fig.39

L'ellisse definita dalle varianze e covarianze delle grandezze X_1 e X_2 della fig.38 e i suoi assi principali.

di un insieme di osservazioni positivamente correlate; le due rette di regressione (x su y e y su x , indicate dai coefficienti b_{xy} e b_{yx} , visto che nell'esempio i coefficienti a sono zero) anche se vicine, sono diverse dalla retta che rappresenta il 1° asse principale.

L'esempio ora fornito e' banale perche' e' in due dimensioni anche se abbiamo visto un risultato importante e cioe' che nelle nuove variabili la varianza della prima variabile (che si chiama anche "fattore") e' considerevolmente aumentata rispetto al-

la seconda.

L'utilità del metodo delle componenti principali si ha però quando il numero di variabili è abbastanza grande, al solito, maggiore di 4 o 5. Se le variabili di partenza sono N avremo quindi un problema di diagonalizzazione di una matrice $N \times N$, che non si può risolvere analiticamente, ma con metodi di approssimazione numerica, che qui non voglio descrivere, e che sono largamente usati in innumerevoli programmi di calcolatore elettronico.

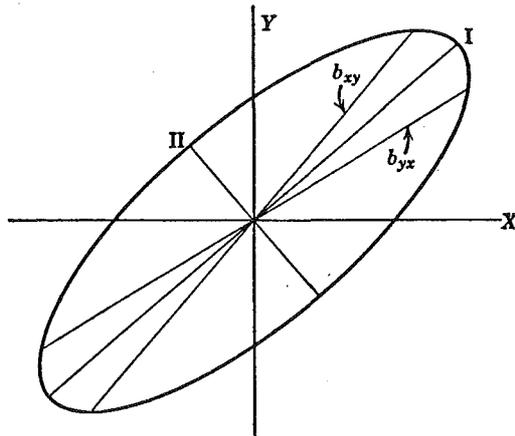


Fig. 40

Differenza tra assi principali e rette di regressione nel caso bidimensionale (da Cooley e Lohnes, 1971).

Se le vecchie variabili erano X_i ($i=1, \dots, N$) allora i "fattori" saranno legati ad esse dalla relazione

$$X_i = \sum_1^N a_{ik} Y_k \quad (2.10)$$

dove gli a_{ik} , che si chiamano comunemente "factor loadings" sono collegati appunto agli autovettori della matrice varianza-covarianza (costituiscono la matrice trasposta di quella degli autovettori).

Inutile dire che l'eq. (2.10) si può applicare ad ogni sin-

gola misura delle variabili X_i . Se le misure sono n avremo quindi

$$X_i^{(h)} = \sum_1^N a_{ik} Y_k^{(h)} \quad h = 1, \dots, n \quad (2.11)$$

dove appunto $X_i^{(h)}$ rappresenta la h -esima misura sulla variabilita' X_i e $Y_k^{(h)}$ il corrispondente effetto sull' k -esimo fattore.

In generale i fattori sono N ma la proprieta' piu' importante dell'analisi in componenti principali e' che spesso pochi fattori sono necessari per rappresentare la quasi totalita' della varianza (come abbiamo visto nell'esempio banale in due dimensioni, il I° fattore rappresenta il 95% della varianza totale, mentre la variabile X_1 dava solo il 75% della varianza totale).

Il risultato pratico e' quindi quello di una notevole *riduzione del numero delle variabili*, che rende chiaramente conto dell'utilita' e dell'importanza del metodo.

Come ultima aggiunta di questa breve introduzione all'analisi dei componenti principali va detto che spesso, nella matrice di correlazione si usano variabili *standardizzate*, definite cioe' da

$$Z_i = \frac{x_i - \bar{x}}{\sigma}$$

In termini di queste nuove variabili le varianze sono tutte uguali ad 1 (le medie sono zero), e le covarianze diventano ora uguali ai coefficienti di correlazione (gia' definiti nell'eq.(1.6)), che variano tra -1 e $+1$.

* * *

L'analisi dei fattori (o "factor analysis") differisce dall'analisi in componenti principali in primo luogo perche' in par-

te è già "a priori" orientata verso una riduzione del numero di variabili, e soprattutto perché una volta ottenuti gli assi principali importanti viene effettuata una "rotazione" di assi che dovrebbe avere lo scopo di fornire una migliore comprensione dei risultati ottenuti.

È inutile sottolineare che le più comuni applicazioni dell'analisi in componenti principali e della "factor analysis" sono state effettuate essenzialmente nel campo dei "tests" di psicologia, già da diversi decenni.

Ovviamente i metodi di rotazione sono svariati, ed esiste tutta una letteratura su di essa, con i loro vantaggi e svantaggi (v. ad esempio il libro di Harman, 1960).

Cercherò ora tramite un semplice esempio tratto dal classico libro di W.W. Cooley e P.R. Lohnes (1971), di spiegare brevemente la differenza tra analisi dei fattori e analisi in componenti principali.

Supponiamo di avere un certo numero (più precisamente 100) di rettangoli di vari tipi e dimensioni. Per ogni rettangolo si scelgono otto misure più o meno "random" e cioè, oltre ad una lunghezza ed una larghezza arbitrarie per ogni rettangolo L_i e W_i , opportune combinazioni di lunghezza e larghezza più un numero intero di una cifra e_{ji} scelto a caso, considerato come "errore" sulla variabile j per l' i -esimo rettangolo.

In tab. 6 vengono mostrati gli otto tipi di "tests" (e cioè di variabili) misurati sull' i -esimo rettangolo, e le misure sono indicate con X_{ji} con $j = 1, \dots, 8$.

La scelta di questo tipo di misura, che potrebbe sorprendere a prima vista, è invece effettuata appositamente in modo "random" per simulare dei problemi (come quelli dei "tests" psicologici) in cui appunto bisogna scoprire quali sono i raggruppi interessanti di variabili da considerare.

Il prossimo passo è naturalmente quello di calcolare la matrice varianza-covarianza tra le otto variabili (standardizzate) e di diagonalizzarla.

Tabella 6

Variabile	Misura
1	$X_{1i} = L_i$
2	$X_{2i} = W_i$
3	$X_{3i} = 10 L_i + e_{3i}$
4	$X_{4i} = 10 W_i + e_{4i}$
5	$X_{5i} = 20 L_i + 10 W_i + e_{5i}$
6	$X_{6i} = 20 L_i + 20 W_i + e_{6i}$
7	$X_{7i} = 10 L_i + 20 W_i + e_{7i}$
8	$X_{8i} = 40 L_i + 10 W_i + e_{8i}$

Si trova che le due prime componenti principali da sole esauriscono il 98,5% della varianza totale (rispettivamente 5.96 e 1.93 su 8 (Cooley e Lohnes, 1971, p.134) e quindi abbiamo una sostanziale riduzione del numero di variabili, da otto a due.

In fig.41 sono mostrati chiaramente i "factor loadings" delle variabili di partenza rispetto agli assi delle componenti principali (I e II). Le vecchie variabili sono indicate con circoletti e si vede chiaramente che hanno tutte un "loading" positivo rispetto al I asse (che viene chiamato il "fattore generale") e "loadings" sia positivi che negativi rispetto al II asse (che, come si è visto, è il più grande degli altri sette fattori, chiamati "fattori bipolari").

Si può effettuare invece una rotazione, ed è questo il risultato della "factor analysis", in questo caso con il metodo VARIMAX, che è uno dei più usati, e si ottengono due nuovi assi, chiamati fattori "di gruppo" o "comuni", indicati in fig.41 con L e W.

Come si vede, ora i "factor loadings" delle otto variabili rispetto ai nuovi assi sono tutti positivi. Si può osserva-

re che le variabili 1,3,5,6,8 hanno ora un "loading" piu' alto rispetto all'asse L, mentre le altre (2,4,7) hanno un "loading" piu' alto rispetto all'asse W.

Riferendoci ora alla tab.6 si vede molto bene che le variabili 1,3,5,6,8 sono quelle che hanno la lunghezza del rettangolo come parametro dominante, a parte gli "errori" e_{ji} , e d'altra parte le variabili 2,4,7 sono quelle che hanno la lar-

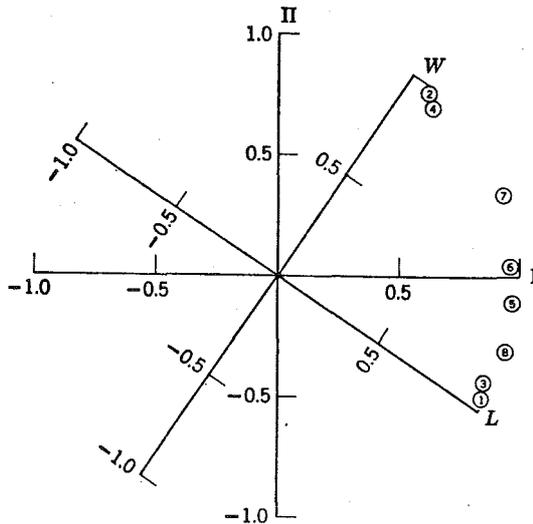


Fig. 41

Rappresentazione grafica dei "factor loadings" delle variabili date in tabella 6 rispetto agli assi principali (I,II) e rispetto agli assi fattoriali (L,W) ruotati con il metodo VARIMAX (da Cooley e Lohnes, 1971).

ghezza come parametro dominante. Non e' un caso che W.W. Cooley e P.R. Lohnes abbiano chiamato i due nuovi assi fattoriali L ("length") e W ("width") in fig.41.

L'interpretazione di questo semplice esercizio "didascalico", mostra che nell'analisi dei 100 rettangoli il metodo dei fattori mette in evidenza, a partire da dati piu' o meno "random", le due soluzioni piu' semplici: una classificazione secondo la lunghezza, ed un'altra secondo la larghezza; e' ovvio che la variabile 1 di partenza (solo lunghezza) e' praticamente incollata

all'asse fattoriale L con il massimo coefficiente, così come accade per la variabile 2 per l'asse fattoriale W, come si può vedere dalla fig.41.

E' comunque importante osservare che anche i fattori derivati dall'analisi dei componenti principali possono avere una loro interpretazione. Il primo fattore (generale), con tutti i "factor loadings" positivi, e' una misura generale di "dimensione" dei rettangoli, mentre il secondo (bipolare) dà la correlazione delle lunghezze con le larghezze, che e' negativa, dato che le variabili 1,3,5,8 hanno "loadings" negativi con l'asse II, come si vede in fig.41.

L'interpretazione piu' semplice e' che il fattore II dipende solo dalle lunghezze e larghezze relative, ed e' indipendente dalle dimensioni generali (fattore I) dei rettangoli.

Ho voluto citare per esteso questo esempio di W.W. Cooley e P.R. Lohnes (1971) perché penso che anche se sostanzialmente banale, serva a dimostrare sia l'indubbio interesse, che i limiti forniti dall'analisi dei fattori.

Per tornare ora ad esempi di carattere archeologico, dove in gran parte e' stata usata la "factor analysis", vorrei citare il lavoro di L.R. Binford e S.R. Binford (1966) su di alcuni complessi Musteriani di "facies levallois", sia francesi (Houpeville, Rouen) che medio-orientali (Jabrud (Siria) e Mugharet es-Shubbabiq (Israele)).

Si tratta appunto di un'analisi di tipo R in cui il punto di partenza e' la lista tipologica di F. Bordes (1961) per il Musteriano, e al variare delle percentuali di questi tipi nei giacimenti (o strati dello stesso giacimento) si cercano correlazioni significative tra strumenti. Il contenuto "culturale" di queste correlazioni dovrebbe essere, secondo gli Autori, isolare gruppi di strumenti che rappresentano "tool kits" con preciso significato di attivita' "funzionale" specializzata: insieme di strumenti per macellazione, per attivita' di sussistenza, ecc.

E' inutile dire che questo tipo di prospettiva é esattamente "ortogonale" se cosi' si puo' dire, alla prospettiva cronostratigrafica classica del Musteriano, anche se F. Bordes (1953, 1970), pur dividendo il Musteriano francese di tipo Levallois in quattro facies (Musteriano tipico, Musteriano di tradizione Acheulana, Musteriano tipo Ferrassie, Musteriano denticolato), sostiene l'ipotesi di una *interstratificazione* tra di esse.

L.R. Binford e S.R. Binford, sulla base delle percentuali dei tipi della lista Bordes, usando un'analisi dei fattori con il metodo di rotazione VARIMAX, hanno estratto dai 62 tipi della lista Bordes 5 fattori e sulla base dei "factor loadings" piu' importanti quantitativamente, hanno cercato di dare appunto un'interpretazione "funzionale" di questi fattori. Siccome si tratta di 5 fattori, non é ovviamente possibile dare una rappresentazione grafica di tutti e cinque i fattori con i loro "factor loadings", cosi' che i risultati vengono proiettati su due assi fattoriali volta a volta.

In fig.42 si vede, per esempio, la distribuzione delle vecchie variabili (i tipi della lista Bordes per le variabili "essenziali" e cioe' con tutte le schegge e punte Levallois non ritoccate) rispetto agli assi fattoriali II e V.

Per l'interpretazione funzionale di questo grafico é bene guardare la fig.43, dove gli Autori danno delle interpretazioni precise, a seconda dei valori dei "factor loadings" (e cioe' delle posizioni dei tipi di Bordes nel piano dei due fattori) sul genere di schema ("pattern") da attribuire ai fattori. Nel caso della fig.42 lo schema sarebbe quello di tipo C della fig.43 (Binford e Binford, 1966, p.256) e quindi rappresenterebbe fattori che hanno attivita' simili da un punto di vista funzionale ma con sole differenze stilistiche come quelli del tipo B (perfettamente simmetrico sulla diagonale, oppure con completa differenza funzionale come quelli del tipo D.

In ogni caso in fig.44 viene data la tabella completa di tutti i fattori, il loro contenuto in termini di variabili di partenza (i tipi della lista Bordes), il loro significato funzionale, ed infine la relazione possibile tra questi fattori, nuovi "clusters" di oggetti funzionalmente interpretati, e la classificazione tradizionale del Musteriano di F. Bordes.

Non è in questa sede il caso di discutere in modo dettagliato tutte le obiezioni sollevate a proposito di questo la-

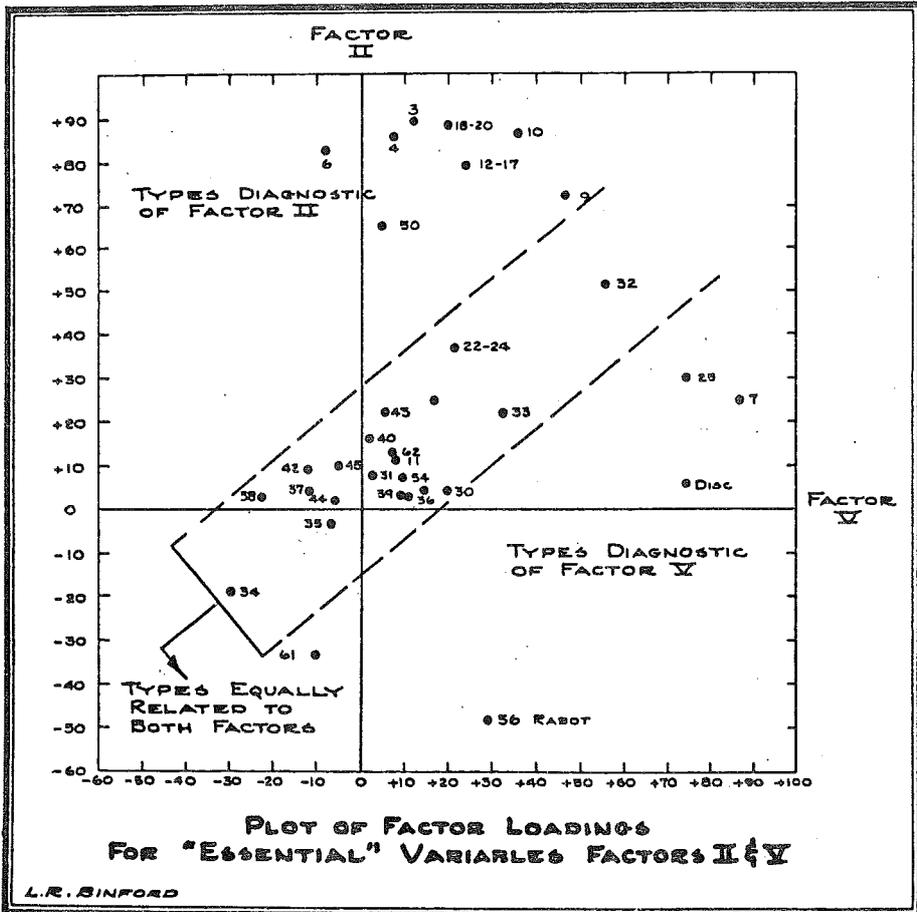


Fig.42

"Factor loadings" dei tipi della lista Bordes per alcuni giacimenti musteriani di facies Levallois rispetto a due assi fattoriali (da Binford e Binford, 1966).

voro dei Binford da vari Autori (Bordes e da Sonnevile Bordes, 1970; Mellars, 1970), nonché le controargomentazioni (Binford, 1973). Basti dire forse che il punto piu' debole di quest'analisi, a parte l'uso della "factor analysis" con la rotazione

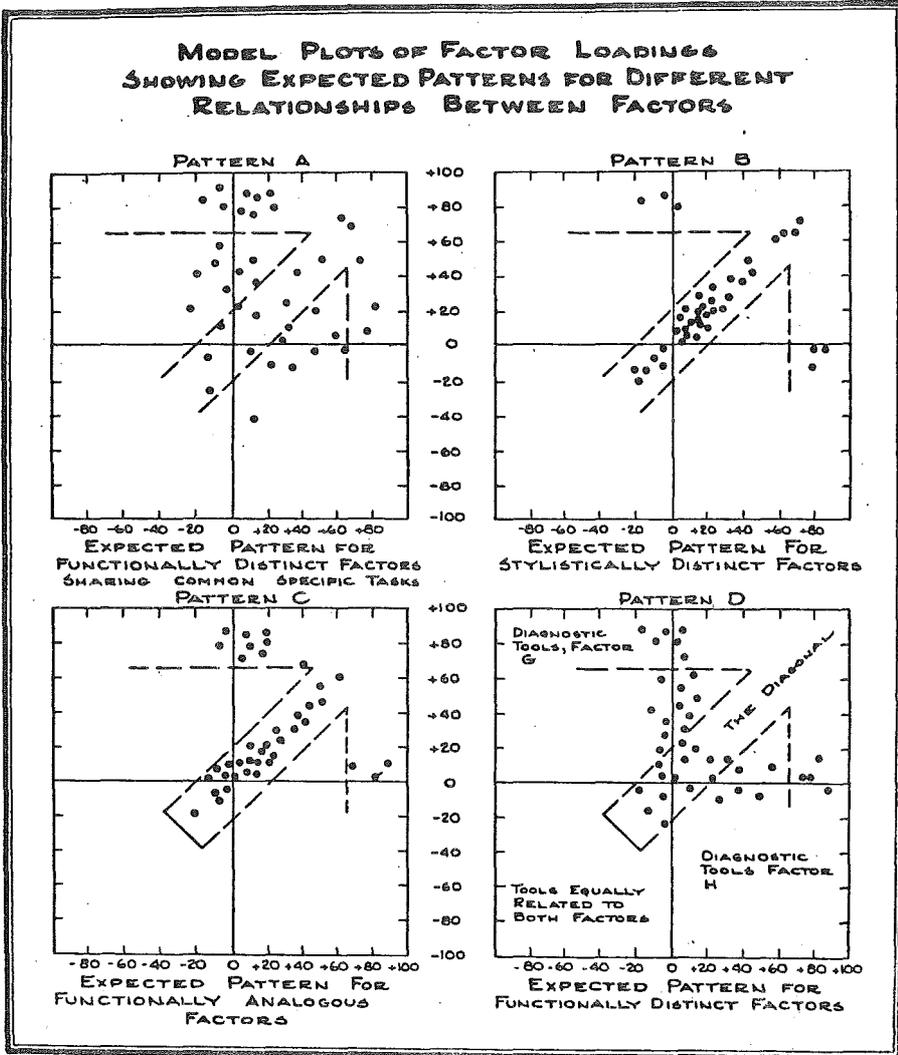


Fig. 43

"Patterns" previsti per i "factor loadings" a seconda di diverse interpretazioni funzionali dei fattori (da Binford e Binford, 1966).

Factor Number	Diagnostic Variables	Suggested Activity	Type of Activity	Analogy to Bordes' Types
I	Typical borer Atypical borer Bec Atypical burin Typical end-scraper Truncated flake Notches Miscellaneous tools Simple concave scraper Ventrally retouched piece Naturally backed knife	Manufacture of tools from non-flint materials	Maintenance tasks	Typical Mousterian (concave graph)
II	Levallois point Retouched Levallois point Mousterian point Convergent scrapers Double scrapers Simple convex scrapers Simple straight scrapers Bifacially retouched piece Typical Levallois flake Unretouched blade	Killing and Butchering	Extractive tasks	Ferrassie (convex graph)
III	Typical backed knife Atypical backed knife Naturally backed knife End-notched piece Typical Levallois flake Atypical Levallois flake Unretouched flake	Cutting and incising (food processing)	Maintenance tasks	Mousterian of Acheulian Tradition (concave graph)
IV	Utilized flakes Scrapers with abrupt retouch Raclettes Denticulates	Shredding and cutting (of plant materials?)	Extractive tasks	Denticulate (concave graph)
V	Elongated Mousterian point Simple straight scrapers Unretouched blade Scraper with retouch on ventral surface Typical burin Disc	Killing and butchering	Extractive tasks	Ferrassie (convex graph)

Fig. 44

Elenco dei fattori riconosciuti dall'analisi di alcuni giacimenti musteriani di *facies* Levallois, e loro relazione con i tipi della lista Bordes (da Binford e Binford, 1966).

VARIMAX, è di carattere *paleontologico*: non è possibile, in vista di una interpretazione che si presume *globale* dei fenomeni culturali, considerare soltanto le industrie. In questa prospettiva le correlazioni delle industrie con i parametri ecologico-ambientali sono essenziali, come abbiamo già visto discutendo i dendrogrammi dei giacimenti del Mesolitico italiano.

Resta comunque il fatto che questo lavoro di L.R. Binford e S.R. Binford ha un valore "storico" fondamentale, non solo per il fatto di essere uno dei primi ad usare tecniche matematiche multivariate, ma soprattutto perché esercita un minimo di critica sulla "intoccabilità" del metodo "cronostratigrafico" per la ricostruzione delle culture preistoriche del Paleolitico.

* * *

In linea di principio, sia l'analisi in componenti principali che l'analisi dei fattori "comuni" hanno come variabili di partenza delle *misure continue*, come percentuali dei tipi, ad esempio. In caso di attributi nominali, non metrici, come presenza o assenza di certi corredi nelle tombe, come si è già visto precedentemente per l'analisi di tipo Q, oltre alla normale creazione di varianze e covarianze tra elementi *standardizzati* (anche se a partire da quantificazione discreta come 0 = assente, 1 = presente, 2 = abbondante, ecc.), che però spesso può restare più o meno semiquantitativa, si possono usare altri metodi che tengono conto di queste circostanze.

Citerò qui, solo come esempio, *l'analisi delle corrispondenze*, introdotta da C.P. Benzecri (1973), che è in sostanza un'analisi di fattori, dove però il punto di partenza, dato che si tratta di variabili nominali, e non continue, è dato da una *distanza di contingenza* derivata appunto da una tabella di contingenza ($n \times p$) dei siti (o strati del medesimo sito) considerati, definita, per quanto riguarda le righe della tabella da

$$d^2(i, i') = \sum_1^p j \frac{1}{N_j^C} \left[\frac{k_{ij}}{N_i^R} - \frac{k_{i'j}}{N_{i'}^R} \right]^2 \quad (2.12)$$

dove se il numero totale degli oggetti è N e se n_{ij} è la frequenza di oggetti della cella definita dalla i -esima riga e j -esima colonna si ha

$$k_{ij} = \frac{n_{ij}}{N} \quad N_j^C = \sum_1^n i k_{ij} \quad N_i^R = \sum_1^p j k_{ij}.$$

Questi ultimi sono ovviamente i *totali marginali* delle tabelle di contingenza. Analogamente si può definire una stessa distanza di contingenza per quanto riguarda le colonne

$$d^2(j, j') = \sum_1^n i \frac{1}{N_i^R} \left[\frac{k_{ij}}{N_j^C} - \frac{k_{ij'}}{N_{j'}^C} \right]^2 \quad (2.13)$$

Da queste distanze si costruisce una matrice di varianza-covarianza che, come al solito, viene diagonalizzata trovando gli assi principali, ecc.

Come esempio volevo solo citare lo studio preliminare degli strati Aurignaziani della Grotta delle Ferrassie in Dordogna (Delporte et al., 1977) che sono diciotto, con varie sigle a partire dal basso: K6-K2, J, I, H, G, F ed E, e che dovrebbero quindi essere in ordine cronologico.

L'analisi quantitativa del lavoro è dovuta a F. Djindjian e la fig. 45 mostra un primo livello di analisi, dove sulla base di due assi fattoriali si vedono le posizioni dei vari tipi (numeri interi) della lista del Paleolitico superiore francese già citata (Sonneville-Bordès e Perrot, 1954-1956), nonché le posizioni dei vari strati (non ci dimentichiamo che i "factor loadings" sono gli stessi anche per le singole osservazioni, eq.(2.11)) e sembra che vi siano tre gruppi abbastanza ben distinti, e più o meno cronologicamente ben delimitati.

Togliendo alcuni tipi ridondanti e rari, nonché alcuni strati poco rappresentativi quantitativamente, con una nuova

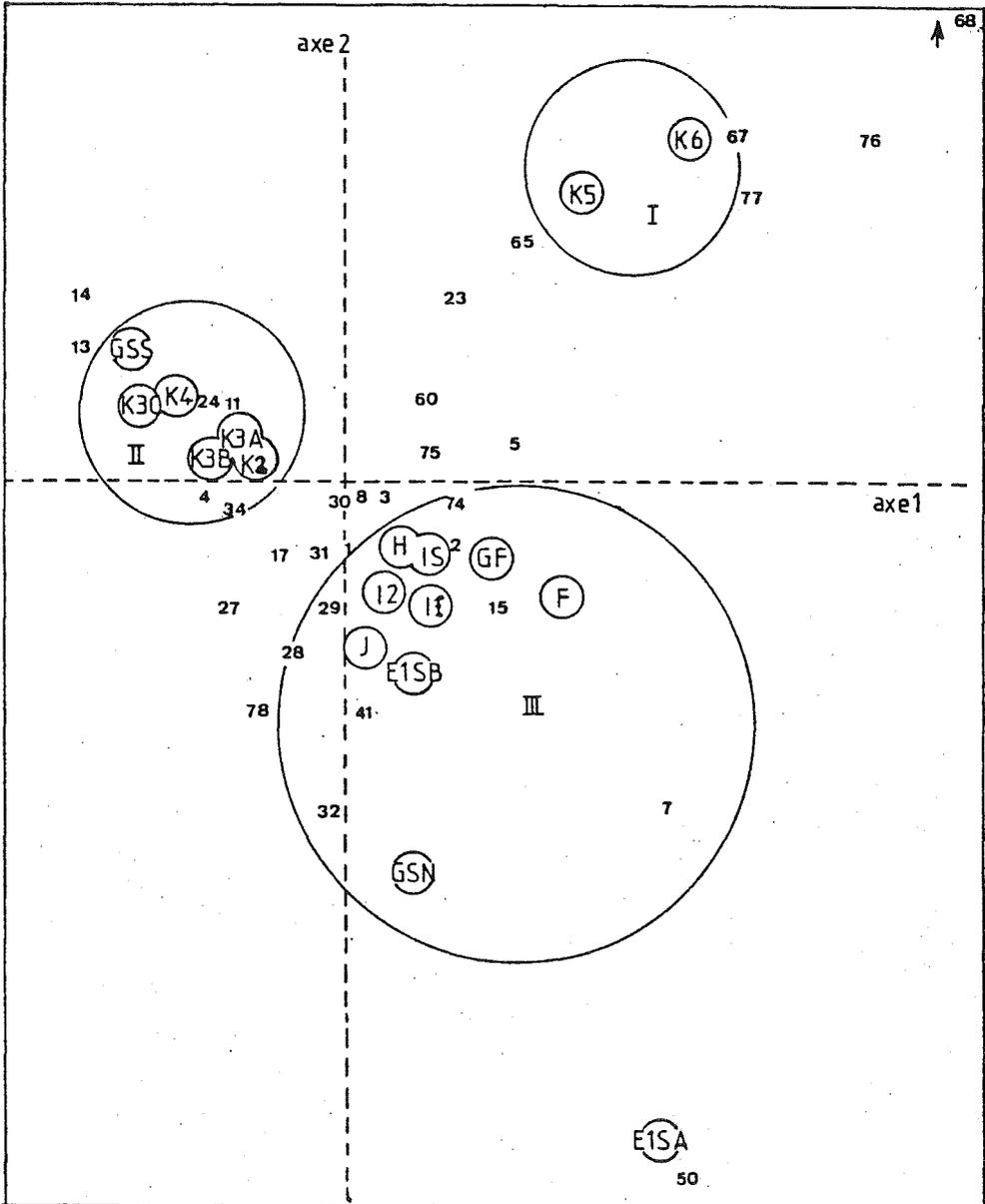


Fig. 45

Analisi delle corrispondenze per vari strati del giacimento di La Ferrassie (Dordogna): rappresentazione grafica rispetto a due assi fattoriali (da Delporte et al., 1977).

analisi che da' quattro fattori importanti (circa il 75% della varianza totale) l'evoluzione cronologico-tipologica dei livelli Aurignaziani della Grotta delle Ferrassie sembra essere meglio delineata. (Delporte et al., 1977, pp. 359-360).

* * *

Come avevo già annunciato nel capitolo precedente, le varie tecniche di analisi multivariata hanno anche un interesse per problemi di tipologia. Uno degli esempi più classici è quello dello studio di 28 fibule dell'Eta' del Ferro (La Tene, V° secolo A.C.) della necropoli di Munsingen in Svizzera, più una dell'Inghilterra e una di Hallstatt (Eta' del Ferro in Austria), largamente descritto da J.E. Doran e F.R. Hodson (1975, pp.218-237). In fig.46 si vedono infatti le 28 fibule di Munsingen (il n.7 e il n.29, che mancano sono appunto quelle del Tamigi, presso Reading, e quella della necropoli di Hallstatt, in Austria).

In fig.47 si vedono invece gli attributi considerati, e cioè le variabili misurate per ogni singola fibula e la loro definizione.

Gli Autori hanno effettuato diversi tipi di analisi: la classica analisi di componenti principali, quella dei fattori con la rotazione VARIMAX, ed infine un altro tipo di analisi statistica, di cui non ho il tempo di parlarne in questa lezione, che è quella *discriminante*.

L'analisi in componenti principali da' ovviamente 13 fattori e gli Autori considerano soltanto i primi quattro, che da soli rendono conto dell'80% della varianza totale (in realtà il quinto fattore, non considerato, rende conto del 9,6% e il sesto, non considerato del 7,2%). In fig.48 si vedono tutti i risultati quantitativi, e in particolare i "factor loadings" sia per l'analisi in componenti principali che per l'analisi dei fattori con la rotazione VARIMAX.

In fig.49a si vedono le rappresentazioni grafiche dei

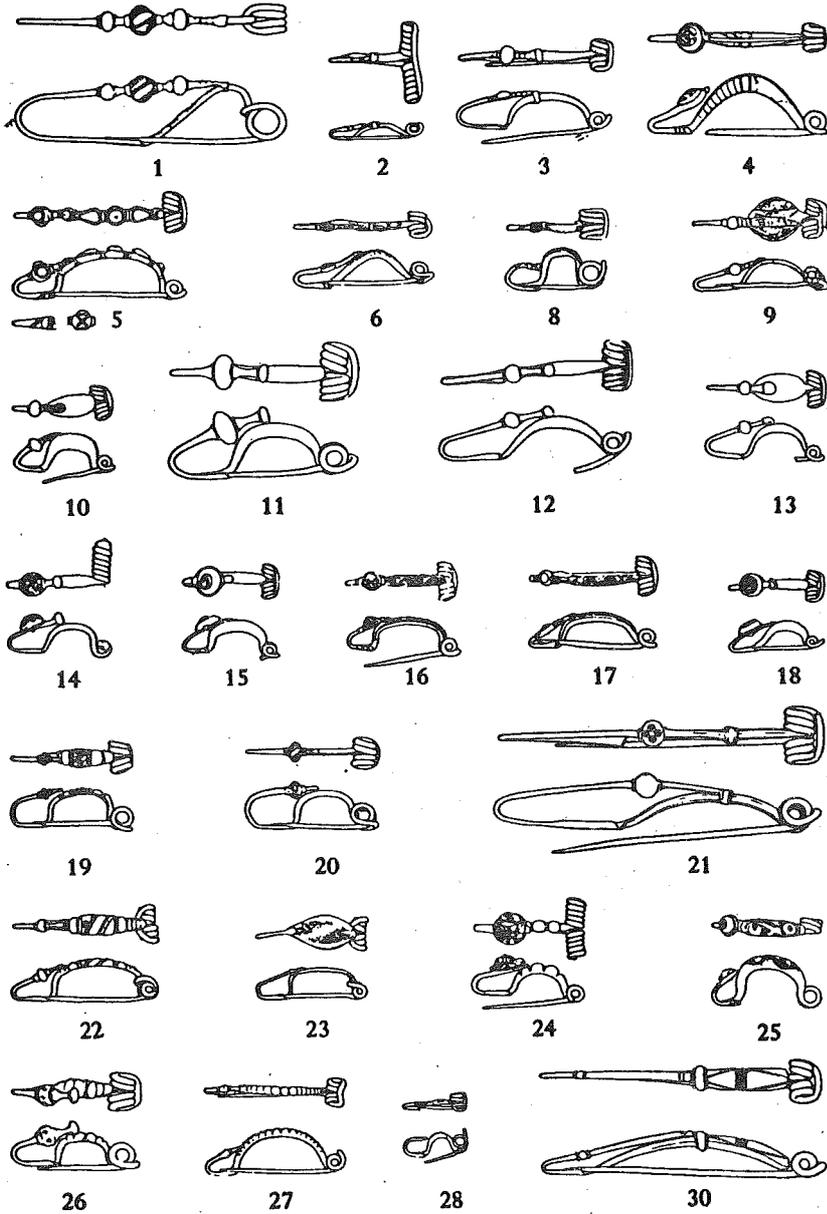


Fig.46

Le Fibule dell'Eta' del Ferro di Munsingen (Svizzera) (da Doran e Hodson, 1975).

"factor loadings" delle fibule rispetto ai primi due assi fattoriali, dove si vede che le fibule "diverse", sia da un punto di vista spaziale che cronologico, sono vicine tra loro ma non particolarmente distinte dal gruppo principale di quelle di Munsingen. In fig.49b che mostra i "factor loadings" rispetto

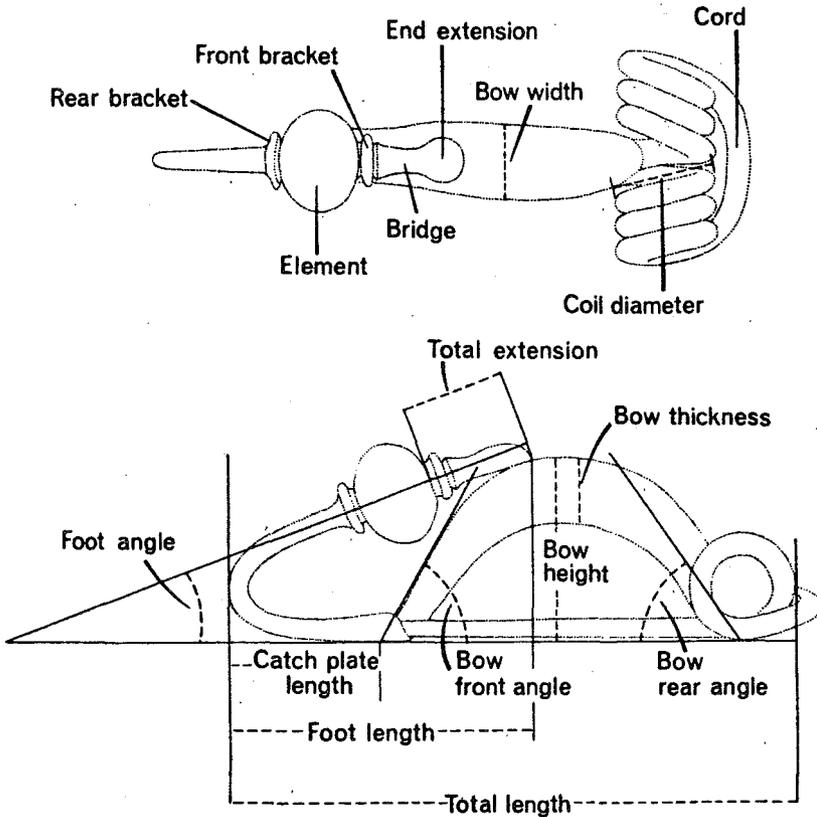


Fig. 47

Definizione degli attributi considerati per le fibule della fig.46 (da Doran e Hodson, 1975).

al primo e al quarto fattore VARIMAX, le due sono un po' piu' separate, anche se insieme ad altre fibule di Munsingen (le n.8 e n.25), per esempio, che gia' erano peraltro vicine nell'analisi in componenti principali (primi due assi, fig.49a)).

E' chiaro che tra i tipi (le fibule) si possono effettuare

anche analisi di tipo Q, e cioè cercare correlazioni tra le fibule al variare del loro contenuto di attributi. In fig.50 si vede infatti un dendrogramma, effettuato con il metodo dell'"average linkage", come quelli mostrati nelle figg.27-37 per il Mesolitico italiano, le tombe ad incinerazione della cultura laziale, Veio, le tombe lucane.

attributes	(a) Principal Components				(b) Varimax Rotated Components			
	1	2	3	4	1	2	3	4
1 foot length	76	-51	25	07	90	-27	06	-12
2 bow height	-94	-02	09	08	-60	31	35	57
3 front bow angle	-80	-16	-11	34	-40	31	07	73
4 foot angle	-32	34	53	61	01	91	18	08
5 coil diameter	-61	-48	-34	-06	-33	-29	07	73
6 rear bow angle	-63	-46	-21	40	-14	12	-02	88
7 element diameter	-57	-44	47	-10	-07	06	72	48
8 foot extension length	78	-12	31	44	88	20	-16	-25
9 catchplate length	06	-95	-04	-08	47	-55	23	58
10 bow width	-48	02	52	01	-18	36	56	15
11 bow thickness	-69	14	37	-12	-48	31	53	17
12 foot extension width	78	-29	27	28	89	-01	-08	-20
13 number of coils	05	14	-73	52	-08	10	-86	26
Relative importance of components indicated by								
(a) eigenvalue	5.2	2.1	1.9	1.2	80.2% of total			
(b) per cent variance	40.2	16.1	14.3	9.6	variance preserved			
(c) cumulative per cent variance	40.2	56.3	70.6	80.2*				
* for remaining components 5-13: 87.4, 91.9, 93.7, 95.5, 97.2, 98.3, 99.1, 99.6, 100								

Fig. 48

Contributi delle variabili definite in fig.47 rispetto ad a) componenti principali, b) fattori VARIMAX (da Doran e Hodson, 1975).

Si vede che la fibula di Hallstatt (H) è praticamente scorrelata da tutto il complesso, mentre quella del Tamigi (T) è correlata solo con le fibule n.8, 28 e 25 a più basso livello, in accordo più o meno qualitativo con le fig.49 a) e b).

Bisogna comunque ricordare che in queste due figure si

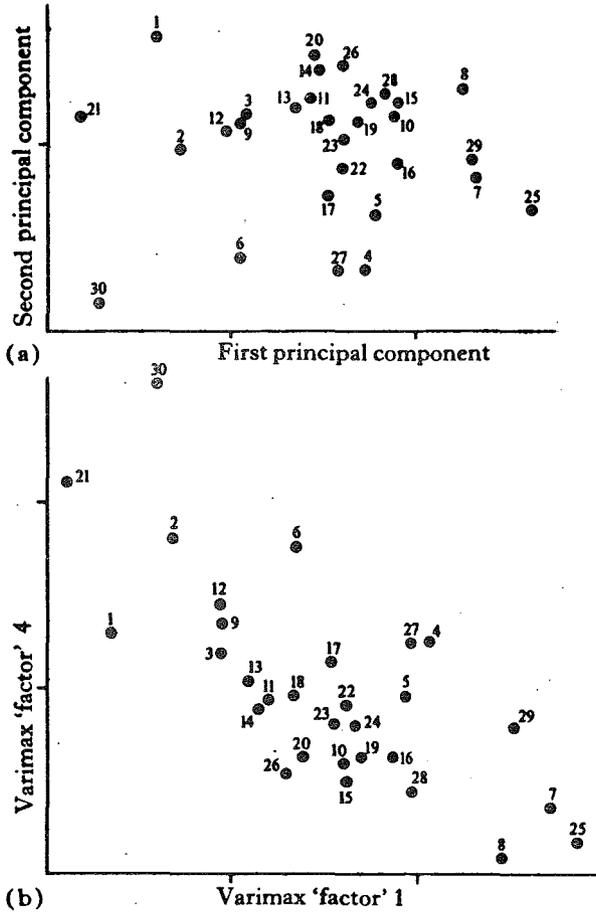


Fig.49

Rappresentazione grafica dei "factor loadings" delle fibule rispetto: a) i primi due fattori del metodo in componenti principali, b) il primo e il quarto fattore VARIMAX (da Doran e Hodson, 1975).

considerano i "factor loadings" solo per due fattori, mentre abbiamo visto che ne occorrono quattro per arrivare all'80% della varianza totale e almeno sei per superare il 90%.

Per questa ragione, secondo me, quando il numero dei fattori importanti e' maggiore di due, e' difficile effettuare analisi di tipo Q e cioe' "clustering" tra tipi, con metodi tipo

rarchia di attributi (Bietti, 1978a), connessi sia con caratteristiche funzionali che stilistiche.

Come esempio di metodi di classificazione gerarchica o *monotetica* dei tipi vorrei illustrare brevemente lo studio effettuato da R. Whallon (1972) sulle ceramiche paleoindiane di tipo Owasco dello Stato di New York.

La tipologia tradizionale per questi tipi di ceramiche è quella di W.A. Ritchie ed R.S. MacNeish (1949) e lo schema generale è mostrato in fig. 51. Come si vede è una *struttura gerarchica ad albero* ("tree structure") dove ovviamente alcuni attributi (tutti "nominali") sono considerati più importanti di altri e pertanto sono più "vicini" alle radici dell'albero di altri attributi che vengono considerati secondari. In questo modo "intuitivo" gli Autori hanno comunque distinto almeno dieci tipi principali più alcune "varianti".

Il problema che si pone R. Whallon è che genere di tipologia si può costruire da questi materiali (vengono considerati cinque complessi diversi, anche cronologicamente) su basi statistiche.

L'analisi del χ^2 e delle tabelle di contingenza, che abbiamo spiegato nel capitolo precedente, anche a più dimensioni, viene ampiamente criticata da R. Whallon, proprio perché è *politetica*, non dà indicazioni gerarchiche, e non riproduce, applicato a questi 5 complessi nessuno dei tipi originali di W.A. Ritchie e R.S. MacNeish, che avevano peraltro una buona definizione cronologica.

R. Whallon prova invece ad applicare un metodo statistico di classificazione *monotetica*, chiamato "association analysis", che è gerarchico e "subdivisivo", usato dagli ecologi delle piante.

È un metodo che si basa su attributi nominali e sulla loro presenza o assenza, usando ovviamente il test del χ^2 come misura di associazione tra essi. Inoltre il procedimento suddivide gli attributi in ordine di importanza e ogni ramo, in

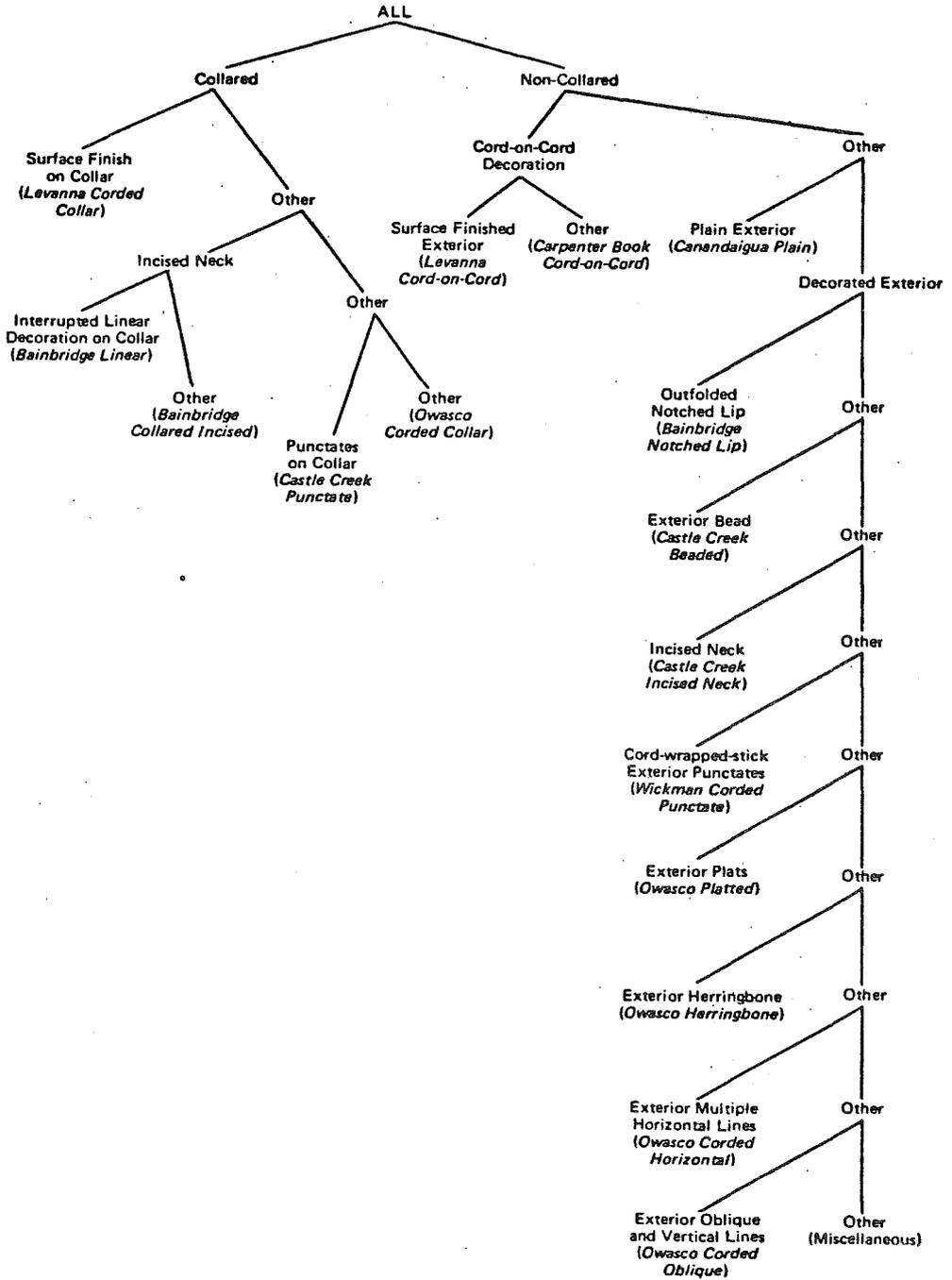


Fig. 51

Tipologia delle ceramiche paleoindiane di Owasco (New York) secondo W.A. Ritchie e R.S. MacNeish (da Whallen, 1972).

questa struttura ad albero, si divide in altri due rami.

La scelta di un nuovo ramo dipende da quale attributo è importante per permettere la suddivisione, e il criterio della selezione è dato dal χ^2 , o meglio dalla somma dei χ^2 tra i vari attributi, la loro media, o il valore più alto del χ^2 (Whallon, 1972, p.19).

Si può così stabilire una gerarchia, di importanza tra gli attributi, a partire dall'attributo che possiede la più alta di queste misure fino a quello con la più bassa.

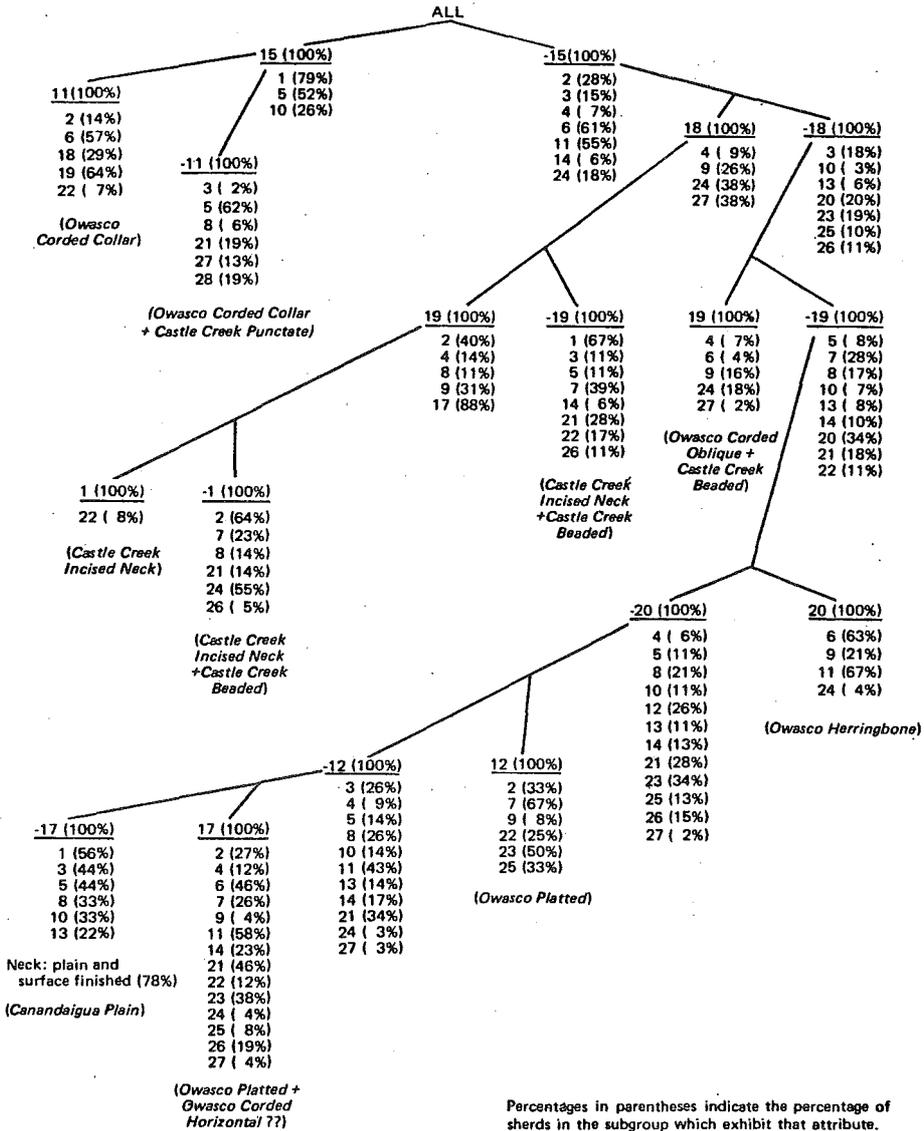
C'è ovviamente una questione che riguarda la minima frequenza accettabile per ogni cella dei valori aspettati in modo tale che il test del χ^2 sia significativo. Naturalmente questa scelta va fatta in modo abbastanza elastico e tenendo conto anche della consistenza numerica del materiale nonché dei problemi di carattere statistico che possono sorgere.

Se si sceglie zero, ad esempio, non ha neanche senso fare il test e l'associazione viene subito scartata. In fig.52 si vede la struttura ad albero (per il giacimento di Castle-Creek) ottenuta da R. Whallon (ovviamente con un programma di calcolatore) per un numero minimo di tre oggetti aspettati per cella, usando come criterio di discriminazione la somma dei χ^2 . La lista degli attributi è data in fondo alla figura, e ogni ramo rappresenta un tipo. In parentesi ad ogni ramo viene data la percentuale degli oggetti di quel ramo che possiede l'attributo, indicato corrispondentemente fuori parentesi.

Le conclusioni di Whallon (1972, p.27) che almeno 7 od 8 su 10 dei tipi definiti dalla classificazione intuitiva di Ritchie e MacNeish si ritrovano tramite l'"association analysis", il che è certamente un bel risultato, e mostra l'importanza di questi metodi di classificazione gerarchica per la costruzione di tipologie.

L'esempio qui presentato riguarda solo attributi nominali, mentre è ovvio che considerando anche attributi metrici le cose si complicano un po'.

SUM OF CHI-SQUARES
MINIMUM EXPECTED CELL FREQUENCY = 3



ATTRIBUTE LIST

- | | |
|---------------------------------|--|
| 1 - Flat lip | 15 - Collar |
| 2 - Beveled lip | 16 - Collar: multiple horizontals |
| 3 - Flat lip with protrusion | 17 - Neck: cord wrapped stick decoration |
| 4 - Beveled lip with protrusion | 18 - Neck: incising |
| 5 - Interior: plain | 19 - Neck: obliques left |
| 6 - Interior: obliques left | 20 - Neck: herringbone |
| 7 - Interior: verticals | 21 - Neck: obliques right |
| 8 - Interior: obliques right | 22 - Neck: verticals |
| 9 - Castellations | 23 - Neck: vertical plats |
| 10 - Lip: plain | 24 - Neck: bead |
| 11 - Lip: obliques left | 25 - Neck: plats right |
| 12 - Lip: verticals | 26 - Neck: multiple horizontals |
| 13 - Lip: obliques right | 27 - Neck: vertical chevrons |
| 14 - Lip: double horizontals | 28 - Collar: punctates |

Fig. 52

Struttura ad albero per i tipi ceramici della fig.51, ottenuta tramite l'"association analysis" (da Whallon, 1972).

In ogni caso, io penso che la costruzione di tipologie mediante strutture ad albero abbia bisogno di sistemi interattivi, e cioè dove si possa modificare la struttura in modo continuo, fino ad ottenere un albero "ottimale" mediante opportune tecniche statistiche.

Come esempio voglio descrivere un progetto di lavoro in corso su un insieme di grattatoi del Paleolitico superiore (Bietti e Zanella, 1980).

In questo caso si hanno sia attributi metrici che nominali. Se consideriamo solo grattatoi interi gli attributi nominali sono essenzialmente quattro: due per il "pezzo supporto": lama o scheggia (con più possibilità di suddivisioni, scheggia di I° distacco, schegge successive, ecc.) e due per la presenza o assenza di ritocco (anche qui vi potrebbero essere diverse suddivisioni: tipo di ritocco, presenza su un solo bordo o su entrambi i bordi, ecc.).

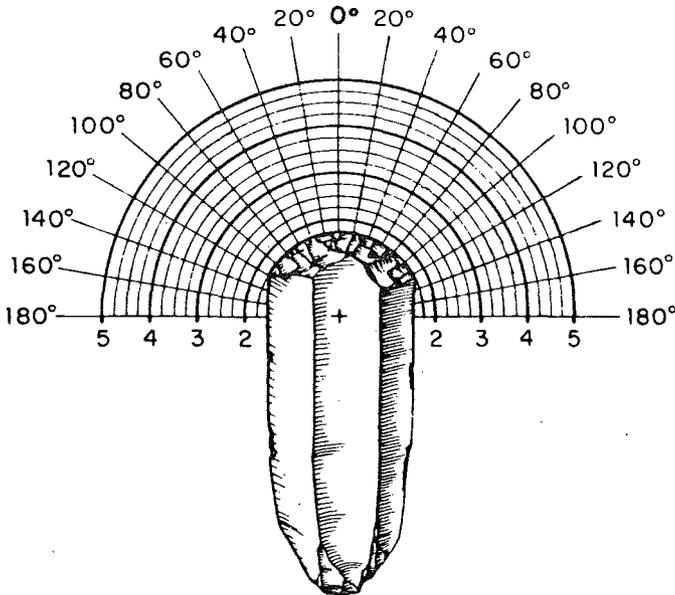


Fig. 53

Misura del raggio di curvatura del fronte di un grattatoio (da Movius et al., 1968).

rare altre variabili quantitative; nelle figg. 53 e 54 si vedono, per esempio, i parametri proposti da Movius et al. (1968) per il raggio di curvatura del fronte del grattatoio e per l'angolo di inclinazione del ritocco.

A nostro avviso, l'angolo di inclinazione del ritocco può avere un'importanza notevole per i complessi aurignaziani, mentre per i complessi apigravettiani, che sono quelli che ci interessano di più, spesso gli angoli ripidi risultano da iperravviamiento del fronte del grattatoio.

Invece del raggio di curvatura preferiamo definire un parametro "curvatura" C, indipendente dalle dimensioni, un parametro di asimmetria A, e un parametro di "triangolarità" T, che ha un particolare interesse per i giacimenti della costa tirrenica, dove si hanno diversi grattatoi triangolari.

In fig. 55 diamo le definizioni di questi parametri ed in fig. 56 uno schema pratico di misura di essi.

Lo schema più logico è quello di provare, come prima ipotesi, una struttura ad albero, come quella rappresentata in fig. 57, dove i vari nodi corrispondono tutti a variabili nominali. In base alle variabili metriche che abbiamo appena definito, si possono effettuare i tests di significanza statistica per distinguere un nodo dall'altro; i tests possono essere i più vari: dalla semplice analisi della varianza su ogni variabile, fino ad analisi discriminanti per decidere quale delle variabili metriche considerate sia "diagnostica" nella decisione tra due rami.

Dato che abbiamo più variabili metriche, si può effettuare, alla radice, o a ciascuno dei rami, un'analisi in componenti principali tra le variabili metriche e trovare opportune correlazioni che possono essere "diagnostiche" per i vari rami considerati.

E' chiaro che una procedura di questo genere, proprio perché parte da una struttura ad albero "a priori", ha bisogno di una possibilità continua di cambiamenti, sia nella scelta e nel

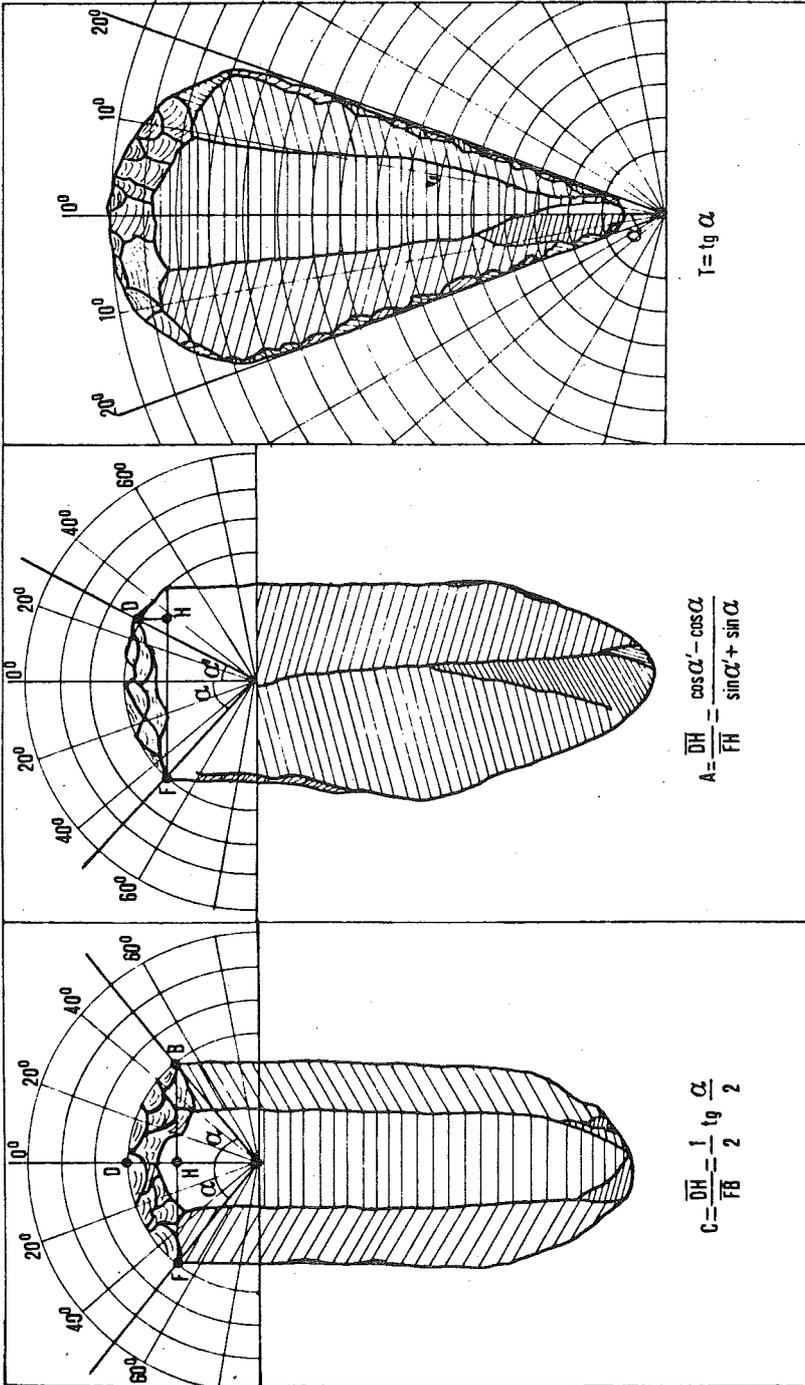


Fig.55

Definizioni della curvatura C, asimmetria A e trisegolarita' T per un grattatoio (da Bietti e Zanollo, 1980).

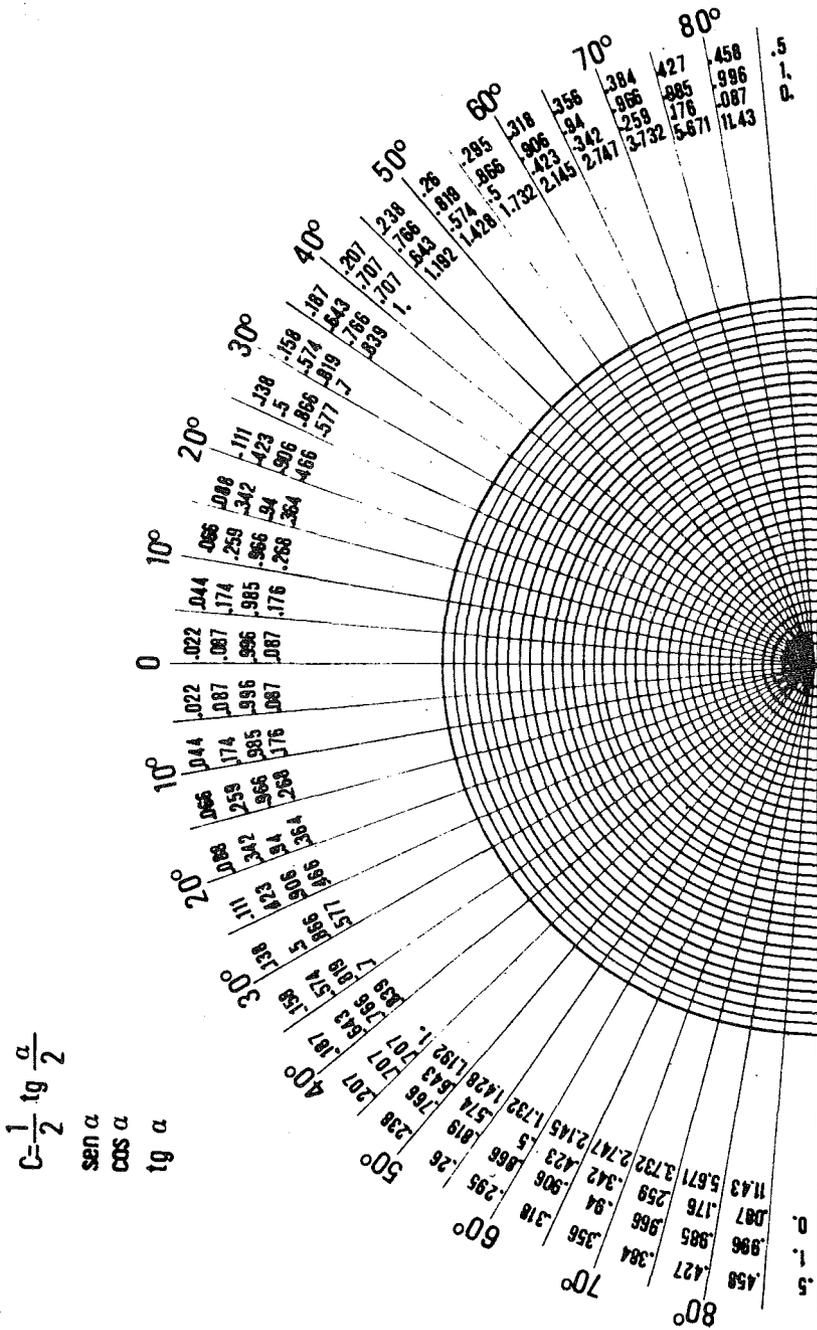
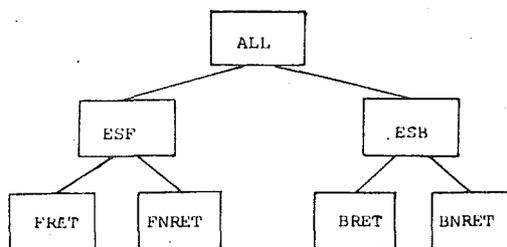


Fig. 56
 Schema pratico di misura dei parametri definiti in fig. 55 (da Bietti e Zanella, 1980).

numero dei parametri, che nello schema dei nodi e dei rami, e quindi di un sistema "interattivo", tramite un "display" video grafico, in modo tale che si possano controllare passo per passo le scelte affettuate e la significanza statistica ai vari



Where:

ESF = end-scrapers on flake
FRET = end-scrapers on flake retouched
FNRET = end scrapers on flake not retouched
ESB = end-scrapers on blade
BRET = end-scrapers on blade retouched
BNRET = end-scrapers on blade not retouched

Fig. 57

Albero "tentativo" per la classificazione dei grattatoi (da Bietti e Zanella, 1980).

nodi, nonché la possibilità di cambiare volta per volta la struttura stessa dei nodi.

A questo proposito, nel lavoro in corso, ci è parso particolarmente efficiente l'uso di un sistema interattivo (ISPAHAN) creato in Olanda per ricerche biomediche (Gelsema, 1980), adattabile su minicalcolatori del tipo DEC/PDP/11, e speriamo nel prossimo futuro di fornire qualche risultato concreto.

III - Problemi di interpretazione globale: ricostruzione di "patterns" nello spazio e nel tempo. Cenni a modelli matematici di culture

Come primo esempio di problemi "interpretativi" voglio citare quello della ricognizione di "patterns" spaziali, e cioè di raggruppamenti di associazioni spaziali significative di diversi materiali provenienti da uno o più strati di un giaci-

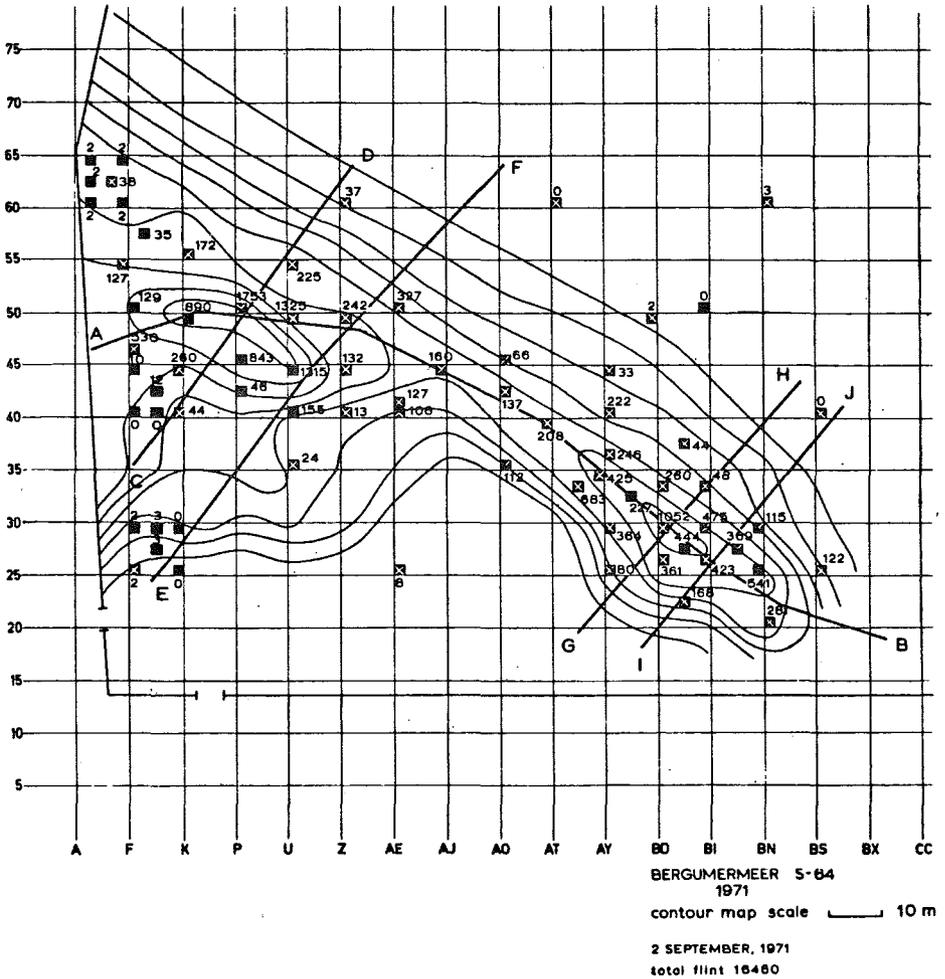


Fig. 58

Distribuzione degli strumenti del giacimento mesolitico di Bergumermeer (Olanda) (da Newell e Vroomans, 1972).

mento archeologico. Anche qui si può procedere intuitivamente, o si possono usare tecniche matematiche semplici o complesse a seconda dei casi.

Un esempio classico è quello, per giacimenti di superficie, della separazione di insediamenti differenti da un insieme che può sembrare in prima analisi omogeneo o appartenente ad uno stesso insediamento.

In fig. 58 si vede la distribuzione degli oggetti litici del giacimento mesolitico di Bergumermeer, nell'Olanda settentrionale (Newell e Vroomans, 1972). Come si vede dalle curve di livello (equidistanza 10 cm, i quadrati sono di 10 m), sembra che vi siano due insediamenti separati. Questa ipotesi è confermata, sulla base dell'analisi della varianza, sia dalla fig. 59 dove si vede la concentrazione di oggetti secondo il profilo A-B della fig. 58 e che rivela una distribuzione tipicamente bimodale, sia dalla fig. 60, che mostra chiaramente come, in termini di densità di strumenti, i due insediamenti sono separati

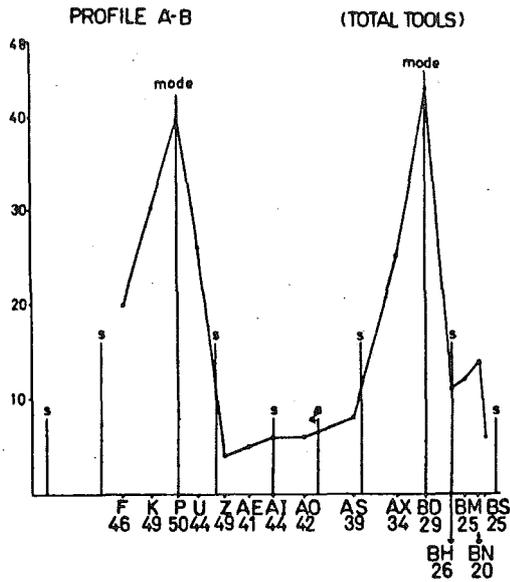


Fig. 59

Frequenza degli strumenti secondo il profilo A-B delle fig. 58 (da Newell e Vroomans, 1972).

per piu' di due deviazioni standard. Un'analisi tipologica piu' fine, in particolare dei microliti, (Newell e Vroomans, 1972, pp.91-92), sembra indicare inoltre una data piu' recente (circa due o trecento anni) per l'insediamento B rispetto a quello A (fig.60).

Per riconoscere distribuzioni spaziali intenzionali e quindi differenti da ipotesi "random" si puo' fare uso di tecniche matematiche piu' complicate, come l'analisi dimensionale della varianza o la "nearest neighbour analysis" largamente usata dagli ecologi delle piante (v.per esempio Greig-Smith, 1964; Ker-

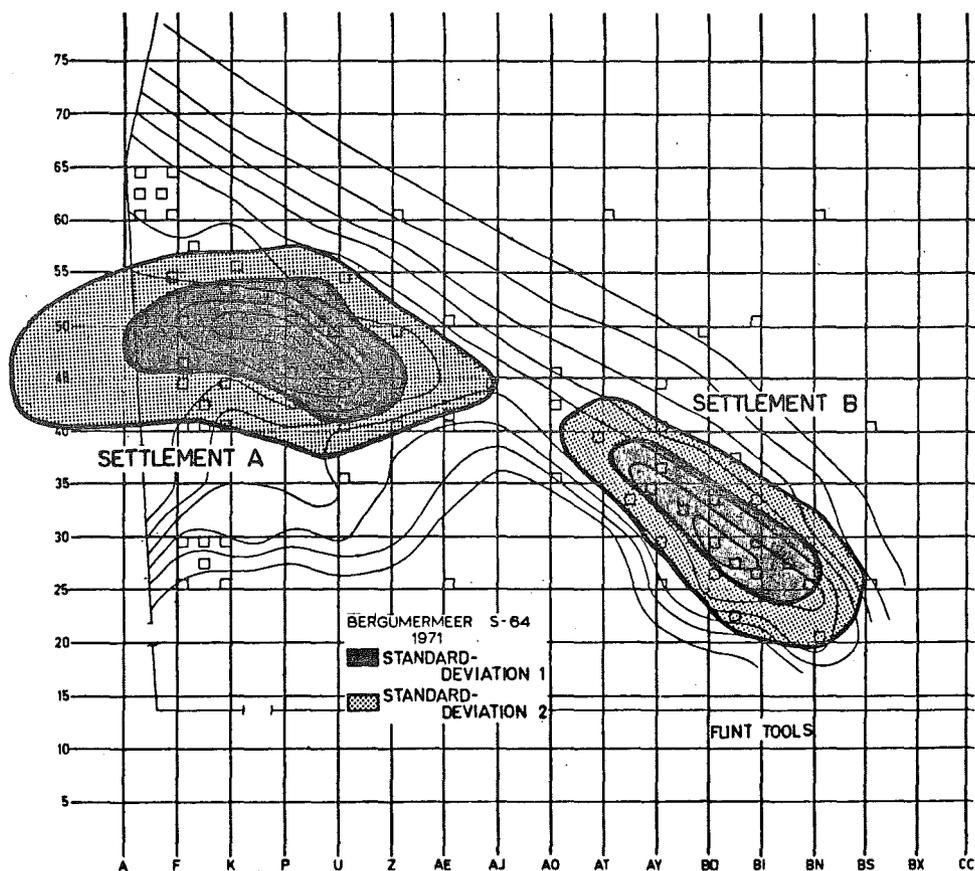


Fig.60

Separazione degli insediamenti A e B di Bergumermeer sulla base dell'analisi della varianza (da Newell e Vroomans, 1972).

shaw, 1961; Clark ed Evans, 1954; Pielou, 1969) e applicata agli scavi archeologici da R. Whallon (1973; 1974).

L'analisi dimensionale della varianza si basa sul concetto di "misura di blocco": il giacimento archeologico viene ridotto a "misura quadrata" e opportunamente sezionato in quadrati più piccoli o rettangoli (doppi o multipli di quadrati). In ogni caso la ripartizione per aree da considerare varia come le potenze di 2 (fig.61, dove si vede una "griglia" di 8×8 m). La quantità da considerare è la somma quadrata delle "misure

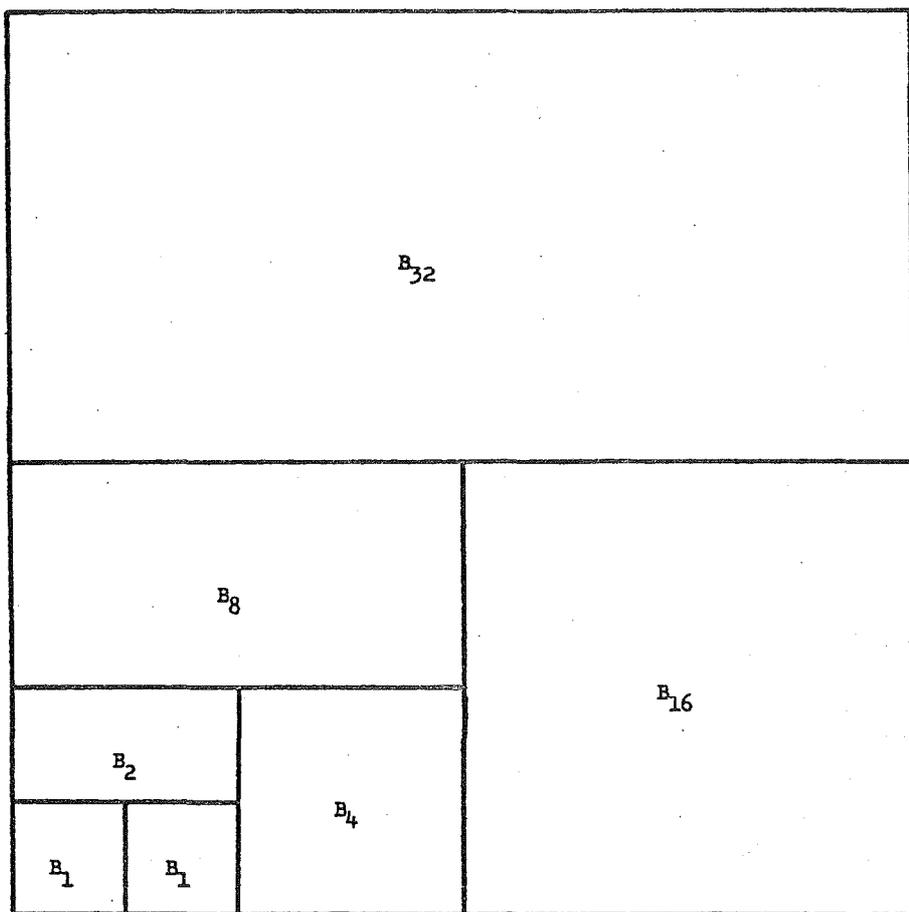


Fig.61

Diverse misure di blocco per un quadrato di 8×8 (da Whallon; 1973).

di griglia" S_j definita da

$$S_j = \frac{1}{j} \sum_{i=1}^{T/j} N_i^2 \quad j = 1, 2, 4, 8, \dots, T \quad (3.1)$$

dove T e' il numero totale di "griglie" di blocco (nel caso della fig. 61, $T = 32$) e N_i e' il numero totale di oggetti nell' i -esimo blocco di misura di griglia j .

Per vedere un "trend" di concentrazione spaziale ad un certo livello di misura di blocco si calcola una "varianza" definita da

$$M_j = (S_j - S_{j+1})/D_j \quad (3.2)$$

dove

$$D_j = \frac{T}{2_j}$$

Se c'è una concentrazione particolare di oggetti ad un blocco j' , in un diagramma che ha in ascisse la misura del blocco e in ordinata M_j (oppure M_j divisa per la media) si dovrebbe vedere un massimo per la misura di blocco j' e un calo significativo per misure di blocco superiori (nonché inferiori). La fig. 62 dà i risultati per le concentrazioni spaziali della *Jatropha* per il giacimento Mesoamericano di Guila Naquitz (Oaxaca, Messico) (Whallon, 1973, p. 273). Come si vede, c'è un picco di M_j (e del rapporto con la media) corrispondente alla misura di blocco 8, ben più significativo del livello di confidenza del 95% (curva inferiore), per distribuzione isotropa.

La limitazione dell'analisi dimensionale della varianza consiste essenzialmente nel fatto che si determina solo una misura di blocco significativa; e non in quale parte del giacimento esiste la concentrazione significativa.

In ogni caso è utile per situazioni, come spesso accade negli scavi archeologici di qualche decennio fa, in cui non c'è

una ricognizione tramite coordinate cartesiane degli oggetti recuperati, ma al massimo si conosce da quale quadrato (le dimensioni dipendono dal tipo di scavo, ovviamente) provengono gli oggetti in esame.

La parte piu' interessante, secondo me, dell'analisi, e' quella che riguarda la *correlazione spaziale* tra diversi oggetti ("items") di cui si e' riconosciuta una concentrazione preferenziale ad una certa misura di blocco.

Ottenute le diverse M_j (eq.(3.2)) per i diversi oggetti

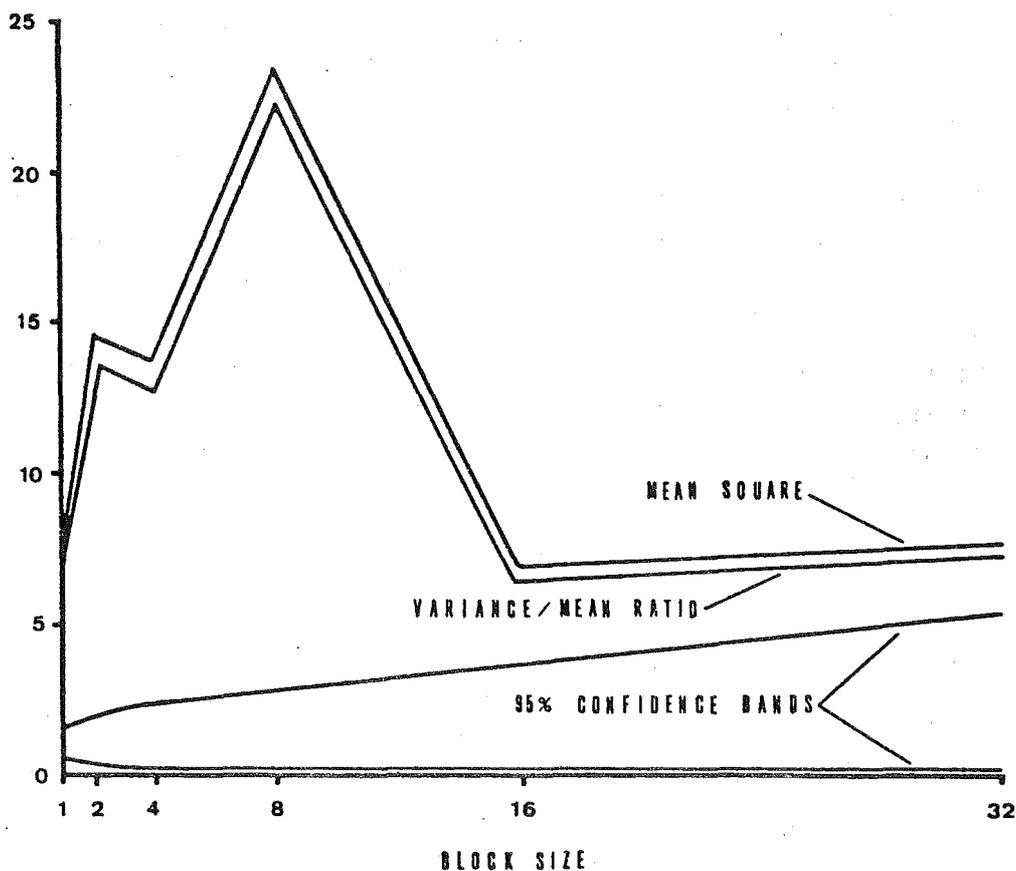


Fig.62

Concentrazione a misura di blocco 8 per la specie *Jatropha* nel giacimento di Guila Naquitz (Oaxaca, Messico) (da Whallon, 1973).

si può tramite metodi di "cluster analysis" come quelli descritti nel capitolo precedente, cercare di rilevare queste correlazioni. In **fig.63**, ad esempio si vede la matrice "sorted" e cioè riordinata secondo il massimo coefficiente di correlazione tra i vari valori di M_j tra resti vegetali, faunistici ed archeologici ("chipped stone") del già citato giacimento di Guila Naquitz (Whallon, 1973, p.276). Come si vede, ci sono quattro o cinque gruppi di oggetti, sia di soli resti vegetali, che di resti animali o misti, che presentano correlazioni spaziali significative che possono pertanto essere interpretate come attività specializzate nelle varie zone del giacimento.

Un altro limite dell'analisi dimensionale della varianza consiste nella necessità di ridurre l'area di scavo ad un "quadrato", e quindi evitare "effetti di bordo". Se si dispone delle coordinate degli oggetti (x,y) per ogni superficie, si può invece usare la "nearest neighbour analysis" e cioè calcolare la distanza media di ogni oggetto con il proprio "vicino", in un'area spazialmente identificata, e poi confrontarla con la distanza ottenuta da una distribuzione "random", oppure concentrata, sulla base della stessa densità di oggetti (Whallon, 1974).

Data un'area A con un numero di oggetti N la densità è ovviamente data

$$d = \frac{N}{A} \quad (3.3)$$

se r_i è la distanza dell'i-esimo oggetto rispetto al i+1-esimo, la distanza media sarà

$$\bar{r}_0 = \frac{1}{N} \sum_{i=1}^N r_i \quad (3.4)$$

Se la distribuzione fosse "random" e cioè governata dalla formula di Poisson, già citata in precedenza, la distanza

	Acacia pods	Opuntia stems	Opuntia leaves	Opuntia stems	Agave leaves	Agave quids	Cucurbita	Prosopis pods	Malpighia	Lucaena/Lysitoma	Jatropha	Chipped stone	Pinyon nuts	Acorns	Celtis	Phaseolus	Rabbit	Deer	Turtle	Opuntia fruit
Acacia pods	1.00	.99	.99	.99	.99	.99	.88	.71	.30	.34	.11	.00	-.22	-.34	-.25	-.32	-.43	-.23	-.41	-.33
Opuntia stems	.99	1.00	.99	.99	.99	.99	.91	.76	.34	.39	.17	.05	-.16	-.28	-.19	-.26	-.39	-.20	-.38	-.35
Opuntia leaves	.99	.99	1.00	.99	.99	.99	.92	.77	.36	.41	.19	.08	-.14	-.25	-.16	-.23	-.36	-.18	-.35	-.34
Agave quids	.99	.99	.99	1.00	.99	.99	.93	.78	.40	.43	.22	.07	-.12	-.23	-.16	-.22	-.38	-.22	-.39	-.31
Cucurbita	.88	.91	.92	.93	1.00	.95	.62	.70	.53	.38	.21	.12	.18	.13	.13	.12	.07	-.07	-.21	-.31
Prosopis pods	.71	.76	.77	.78	.95	1.00	.69	.84	.73	.63	.47	.39	.47	.41	.16	.14	.14	.03	.03	-.37
Malpighia	.30	.34	.36	.40	.62	.69	1.00	.68	.71	.36	.46	.54	.49	.61	.00	-.25	-.12	.38		
Lucaena/Lysitoma	.34	.39	.41	.43	.70	.84	.68	1.00	.96	.80	.82	.74	.66	.65	.32	.14	.04	.30		
Jatropha	.11	.17	.19	.22	.53	.73	.71	.96	1.00	.84	.92	.88	.80	.82	.47	.21	.18	-.17		
Chipped stone	.00	.05	.08	.07	.38	.63	.36	.80	.84	1.00	.90	.84	.90	.81	.81	.69	.61	-.52		
Pinyon nuts	-.22	-.16	-.14	-.12	.21	.47	.46	.82	.92	.90	1.00	.97	.88	.88	.70	.42	.41	-.23		
Acorns	-.34	-.28	-.25	-.23	.12	.39	.54	.74	.88	.84	.97	1.00	.92	.95	.70	.38	.45	-.03		
Celtis	-.25	-.19	-.16	-.16	.18	.47	.49	.66	.80	.90	.88	.92	1.00	.96	.86	.63	.70	-.16		
Phaseolus	-.32	-.26	-.23	-.22	.13	.41	.61	.65	.82	.81	.88	.95	.96	1.00	.75	.45	.57	.05		
Rabbit	-.43	-.39	-.36	-.38	-.12	.16	.00	.32	.47	.81	.70	.70	.86	.75	1.00	.91	.93	-.40		
Deer	-.23	-.20	-.18	-.22	-.07	.14	-.25	.14	.21	.69	.42	.38	.63	.45	.91	1.00	.94	-.64		
Turtle	-.41	-.38	-.35	-.39	-.21	.03	-.12	.04	.18	.61	.41	.45	.70	.57	.93	.94	1.00	-.34		
Opuntia fruit	-.33	-.35	-.34	-.31	-.31	-.37	.38	-.30	-.17	-.52	-.23	-.03	-.16	.05	-.40	-.64	-.34	1.00		

Group 1—Acacia, Opuntia stems, Agave, Cucurbita, Prosopis, Group 2—Malpighia, Group 3—(a) Lucaena/Lysitoma, Jatropha; (b) Pinyon nuts, Acorns, Celtis, Phaseolus; (c) Chipped stone. Group 4—Rabbit, Deer, Turtle. Group 5—Opuntia fruit.

Fig. 63.

Matrice di similarita' riordinata per i diversi oggetti del giacimento di Guila Naquitz (Oaxaca, Messico) (da Whallon, 1973).

media sarebbe (Clark ed Evans, 1954)

$$\bar{r}_e = \frac{1}{2\sqrt{d}} \quad (3.5)$$

dove d è la densità data nell'eq.(3.3). È chiaro che questa distanza media è zero per un insieme fortemente concentrato, e in ogni modo il rapporto

$$R = \frac{\bar{r}_0}{\bar{r}_e}$$

può dare una misura di "clustering" spaziale nell'area A .

Se vi fosse una "spaziatura" ben ordinata (struttura ad esagoni) allora il valore teorico di \bar{r}_0 sarebbe

$$\bar{r}_0 = \frac{1.0746}{\sqrt{d}}$$

ben diverso dal valore "random" dato dall'eq. (3.5) e si avrebbe quindi un valore $R = 2.1491$ ovviamente diverso non solo dal valore 0 di massima concentrazione ma anche dal valore 1 di ipotesi "random" già citato.

Naturalmente per il valore \bar{r}_0 effettivamente osservato si rende necessario un test di significanza statistica rispetto alle diverse ipotesi, introducendo l'opportuna variabile standardizzata

$$C = \frac{\bar{r}_0 - \bar{r}_e}{\sigma_{\bar{r}_e}} \quad (3.6)$$

con

$$\sigma_{\bar{r}_e} = \frac{0.261}{\sqrt{Nd}}$$

dalla solita distribuzione di Poisson. Anche qui su C si applica la solita analisi della varianza con i vari livelli di confidenza definiti sulla base di 2 o 3 deviazioni standard, ecc.

Le poche formule che abbiamo fornito si riferiscono ad analisi di livelli di confidenza basate su distribuzioni *normali*. Non e' il caso qui di dilungarsi sulle possibili correzioni e sui differenti tipi di test per altri tipi di distribuzioni teoriche e rimando per questo all'analisi di R. Whallon (1974, pp.19-22) nonche' alla letteratura gia' citata.

La "nearest neighbor analysis" ha ovviamente il vantaggio di dare una "localizzazione spaziale" di un eventuale "clustering", che varia pero' per quanto riguarda il livello di significanza statistica, dalla dimensione dell'area in esame.

Se, per esempio, siamo in una situazione in cui c'e' un'attivita' specializzata di produzione di schegge da nuclei di selce, con concentrazione di una zona particolare, dell'ordine di mezzo metro quadrato, e cioe' il luogo stesso di attivita' dell'operatore preistorico, l'analisi su un'area di queste dimensioni dara' probabilmente un valore di \bar{r}_0 (eq.(3.4)) dell'ordine di zero, mentre un'analisi su uno o due metri quadrati dara' valori diversi e il "test" sulla variabile standardizzata C (eq. (3.6)) puo' divenire piu' delicato.

E' naturale quindi che per poter variare rapidamente l'area in esame, e cioe' le varie misure di "griglia" occorre l'uso dell'elaboratore elettronico, per ragioni di risparmio di tempo. A questo scopo, nel 1978, e' stata data una tesi di laurea presso l'Istituto di Fisica dell'Universita' di Roma (D. Bulgarelli, 1978) che per il momento affronta solo l'uso della statistica normale di Clark ed Evans (1954), nonche', ovviamente, il disegno automatico tramite calcolatore, della distribuzione degli oggetti nell'area considerata, e della distribuzione in *stratigrafia* (piani xz e yz) dei singoli oggetti. Per il momento si dispone praticamente di dati "simulati" ma si spera

nel prossimo futuro di applicare il programma a dati reali.

Una volta eseguita la "nearest neighbor analysis" per ogni oggetto, il problema è naturalmente quello di stabilire la relazione spaziale tra i vari oggetti. R. Whallon suggerisce alcuni sistemi, tra i quali il più semplice è quello di costruire un "indice di similarità percentuale" (analogo a quello di Jaccard definito nel paragrafo precedente), a seconda della percentuale comune di area ricoperta dagli esemplari dei due oggetti A e B sul totale dell'area in cui si è effettuata la "nearest neighbor analysis" (Whallon, 1974, p.24).

$$I = \frac{100 a}{a + b + c} \quad (3.7)$$

dove a è la percentuale di area comune agli oggetti A e B, b e c sono le percentuali di area dove non si trovano esemplari rispettivamente dell'oggetto A e dell'oggetto B.

A mio giudizio, a parte i risultati ottenuti da R. Whallon per lo studio di un celebre giacimento del Paleolitico superiore francese (Abri Pataud), il calcolo dell'indice I è piuttosto macchinoso e poco pratico per le applicazioni con l'elaboratore elettronico.

È forse preferibile, una volta individuata una "griglia" di aree (le cui dimensioni possono variare ovviamente al variare degli elementi), determinare una densità per unità di area di ogni tipo di oggetto e poi, coi metodi di "cluster analysis" come quelli esposti nel paragrafo precedente, sia di riordinamento di matrice che dendrogrammi, usando, per esempio, un coefficiente di distanza percentuale o un coefficiente di correlazione, determinare l'esistenza o meno di associazione spaziale tra gli oggetti stessi (D. Bulgarelli, 1978, pp.75-79).

È naturale che anche qui, questo genere di analisi va eseguita più volte, variando l'unità di area che si prende in considerazione, dato che non è detto che, per i diversi oggetti considerati, l'eventuale concentrazione spaziale sia simile

sia in estensione (sostanzialmente la "misura in blocco" dell'analisi dimensionale della varianza) che in localizzazione.

Mi sembra importante sottolineare, ancora una volta, l'importanza della scelta della "griglia ottimale" che definisce le aree di misura. A questo proposito voglio soltanto citare l'attenzione che a questo problema è stata data in un recente lavoro di R. Newell e A. Dekin (1978) dove sulla base di tecniche matematiche essenzialmente basate sul metodo del χ^2 non parametrico, di cui abbiamo già parlato nel primo paragrafo, si cercano di definire "patterns" spaziali che possono dare informazioni di carattere "behavioural" per giacimenti di interesse "etno-archeologico" delle culture esquimesi dell'Alaska.

* * *

Abbiamo parlato fino ad ora della ricerca di "patterns" spaziali nell'organizzazione interna di un singolo giacimento.

Si può anche parlare di ricostruzione di "patterns" spaziali anche per quanto riguarda la distribuzione di insediamenti in un certo territorio. Questo problema introduce la nozione di *modello* di "settlement", e cioè un'ipotesi teorica sulla distribuzione degli insediamenti, e soprattutto il concetto di "simulazione" e cioè di una *ricostruzione completa* (almeno questo è lo scopo che si vorrebbe raggiungere!) della situazione archeologica da parte del modello, che va poi modificata sui dati effettivi (matematicamente è un problema molto simile a quello appena visto della ricognizione preferenziale di oggetti su di una superficie di scavo, confrontata con distribuzioni teoriche ben definite, come ad esempio, quella "random").

Come esempio, possiamo citare un modello di distribuzione (e di diffusione spaziale) per gli insediamenti del Neolitico della Polonia meridionale, citato da I. Hodder e C. Orton nel libro "Spatial analysis in Archaeology" (1976, pp.85-97).

Il problema è quello di decidere, per tre tipi di culture, la "Bandkeramik", "Lengyel", e la "TRB" (cultura dei va-

si a bocca ad imbuto) che tipo di diffusione si ha degli insediamenti spaziali nel territorio: a) di "contagio" effettivo ("true contagion") in cui la diffusione è sostanzialmente omogenea e l'esistenza di un "cluster" spaziale significa che un insediamento ha più probabilità di avere come *prossimi vicini* insediamenti simili; b) di "contagio" apparente o "spurio" in cui c'è una inomogeneità nella distribuzione dei "cluster" di insediamento, dovuta forse a caratteristiche regionali che non permettono la diffusione delle "informazioni" oppure a inomogeneità della distribuzione spaziale delle caratteristiche culturali delle popolazioni che ricevono l'informazione.

La zona in esame viene "quadrettata" e cioè si costruisce una griglia di base, ad esempio di unità $1 \text{ km} \times 1 \text{ km}$ e si aumenta successivamente la superficie dell'unità di base (quadraticamente: 2×2 , 4×4 , ecc. un po' come per l'analisi dimensionale della varianza che abbiamo citato prima).

Il modello teorico che viene usato dagli Autori è la *distribuzione binomiale negativa* che dà, come probabilità di trovare x insediamenti (su di un numero totale n), per una certa "misura di blocco" S

$$P(x) = \frac{(x+k-1)!}{(k-1)! x!} p^k (1-p)^x \quad (x=0, 1, \dots, n) \quad (3.8)$$

dove

$$p = \bar{x}/S^2 \quad \text{e} \quad k = \bar{x}^2/(S^2 - \bar{x}) \quad (3.9)$$

\bar{x} è naturalmente la media dei valori di x e $x!$ si chiama fattoriale di x ed è definito, come è noto, da $x! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot x$.

(Vale forse la pena di ricordare che la più familiare *distribuzione binomiale*, per intenderci quella che si applica al numero di lanci x di una moneta o di un dado, è data da

$$P(x) = \frac{n!}{(n-x)! x!} p^x (1-p)^{n-x} \quad (x=0, 1, \dots, n)$$

dove ad esempio, $p = \frac{1}{2}$ per la moneta (testa o croce) e $p = \frac{1}{6}$ per il dado).

Si puo' dimostrare che nel caso di "contagio" effettivo, indicando con p_1 e k_1 i valori di p e k dati dall'eq. (3.9) con $S = 1$ si ha, per un valore generico di S

$$k_S = Sk_1 \quad \text{e} \quad p_S = p_1 \quad (3.10)$$

mentre nell'ipotesi di "contagio" apparente si ha

$$k_S = k_1 \quad \text{e} \quad p_S = \frac{1}{1 + S \frac{1-p_1}{p_1}} \quad (3.11)$$

Si tratta allora di fare un "best fit" tra la distribuzione osservata e quella aspettata a partire dall'eq.(3.8) con le due diverse ipotesi eq.(3.10) e (3.11) e scegliere quindi, tramite il metodo del χ^2 (*parametrico*, questa volta) quale delle due ipotesi e' verosimile.

Il risultato (Hodder e Orton, 1976, pp.91-97) e' che mentre per le fasi della "Bandkeramik" e Lengyel l'ipotesi di un "pattern" di insediamenti piu' probabile e' quella del tipo di "contagio" effettivo, per la cultura dei vasi a bocca ad imbuto (TRB), nella zona in esame il modello di "contagio" apparente spiega ugualmente bene i dati, e si avrebbe quindi la possibilita' di una differenziazione cronologica: inizialmente si ha una diffusione ben spiegata dal modello di distribuzione binomiale negativa con i parametri dati dall'eq.(3.10) e successivamente una situazione spiegabile con il modello di "contagio" spurio con i parametri dati dall'eq.(3.11).

Un altro esempio di studio di distribuzione di insediamenti, pero' piu' orientato sull'approccio di *simulazione* tramite l'elaboratore elettronico e' quello dato da A.J. Chadwick (1978) per la Messenia in Grecia, durante il periodo Miceneo tenendo conto anche di un sviluppo temporale, e cioe' a partire dall'Elladico medio (circa 1600 A.C.) fino all'Elladico tardo

IIIb (circa 1250 A.C.).

L'area di studio, i cui dati sono stati ricavati da una spedizione dell'Universita' di Minnesota (McDonald e Rapp, 1972) si puo' vedere in fig.64.

I dati della spedizione suddetta sono poi stati usati per dare una rappresentazione schematica ottenuta tramite una "qua-

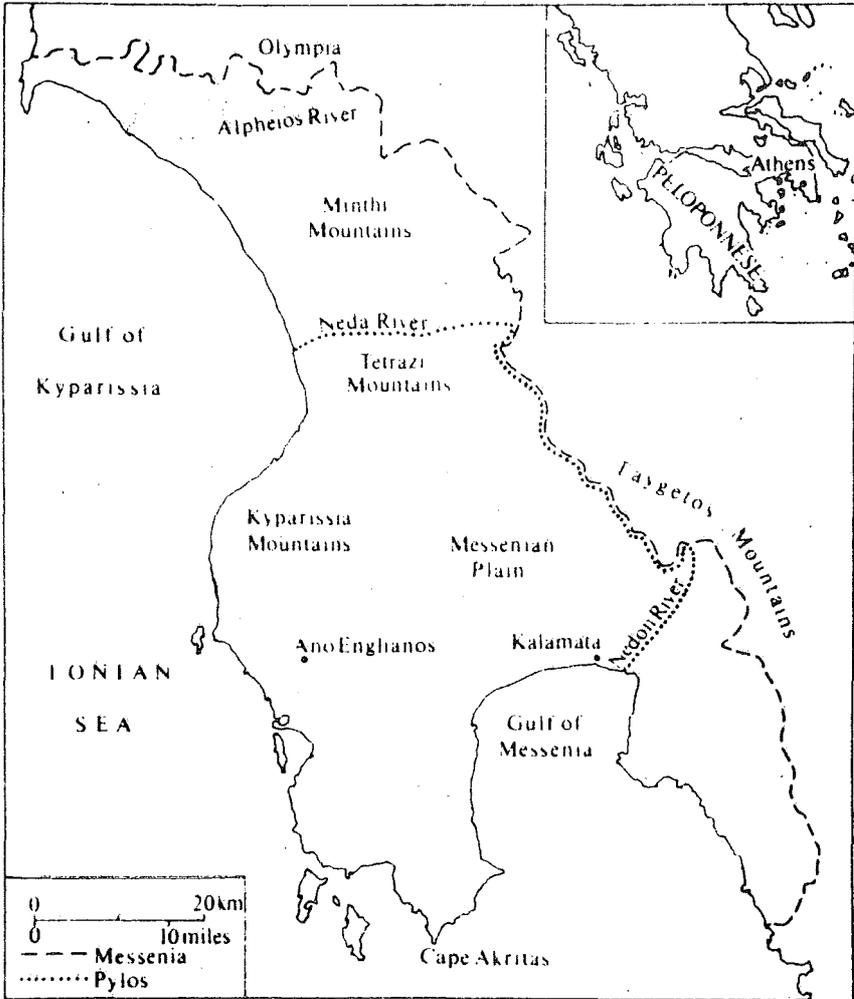


Fig.64

Pianta dell'area di studio (Messenia) della spedizione dell'Universita' di Minnesota (da Chadwick, 1978).

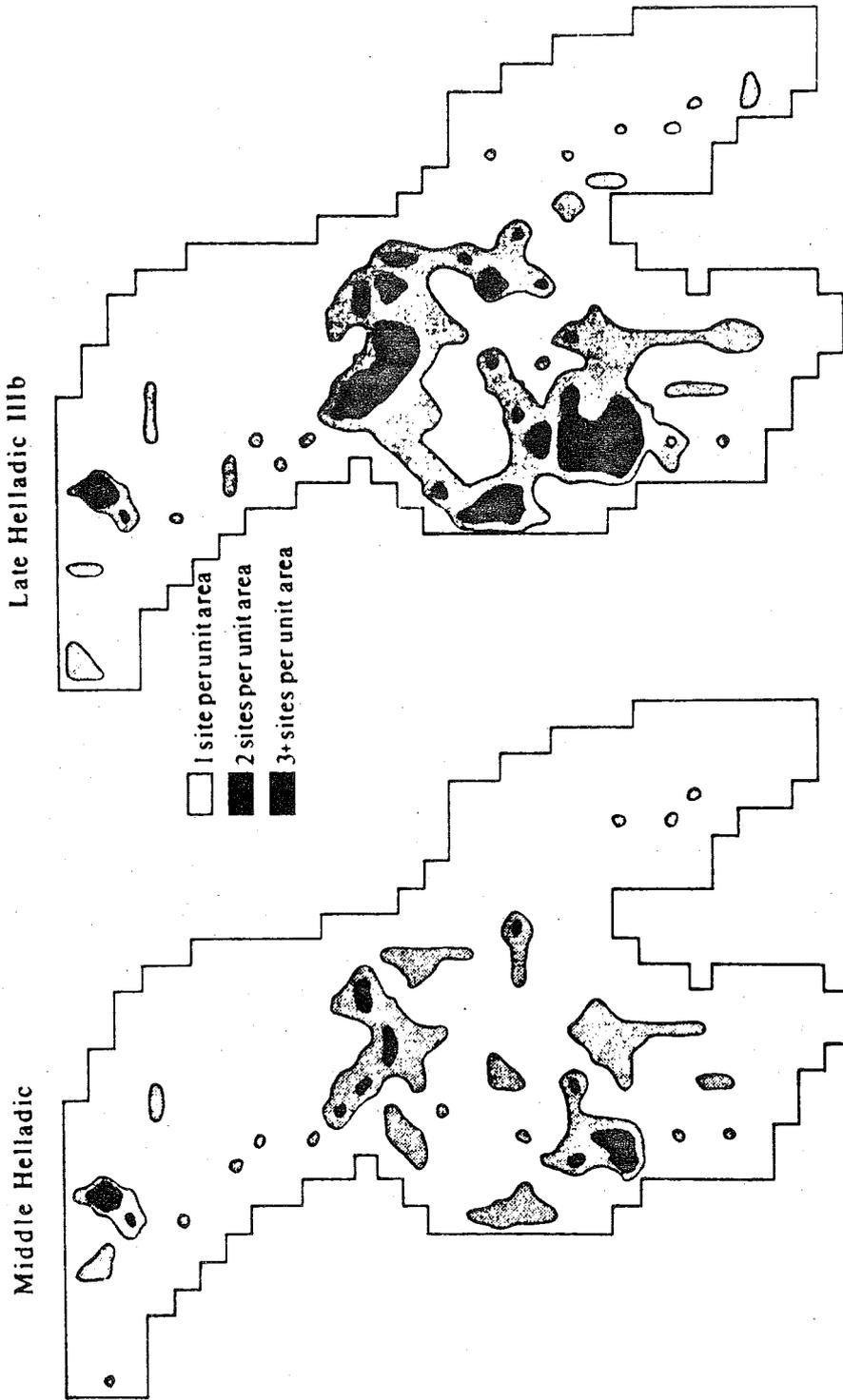


Fig.65

Rappresentazione schematica degli insediamenti in Messenia per l'Elladico medio e l'Elladico tar-
do III b (da Chadwick, 1978).

drettatura" dell'area presa in esame della densità degli insediamenti sia dell'Elladico medio (94) che del tardo Elladico IIIb (169) (fig.65).

Come modello di simulazione il punto di partenza è naturalmente la solita distribuzione binomiale negativa già descritta (eq.(3.8)) con una griglia che ha come unità di base 2×2 km. L'Autore non si aspetta ovviamente di ottenere buoni risultati con l'ipotesi di "contagio" effettivo, visto che non c'è alcuna ragione di considerare la Messenia come un territorio isotropo in due dimensioni: vi sono diverse montagne, fiumi che scorrono in valli con rive piuttosto scoscese, ecc.

L'Autore propone allora un modello che tiene conto di due variabili principali: 1) accessibilità a risorse d'acqua; 2) caratteristiche del territorio, per quanto riguarda sia le possibilità reali di insediamento, che quelle di sfruttamento agricolo.

A ogni quadrato 2×2 viene attribuito un indice opportuno che "quantizza" le due variabili precedenti, definendo così un "indice di superficie" (Chadwick, 1978, p.52) che dovrebbe dare una misura di "preferenza" per certe zone e quindi di concentrazioni di insediamenti.

Questo modello si basa essenzialmente su due ipotesi: 1) la scelta della posizione dell'insediamento è casuale ma la probabilità di avere una certa densità è proporzionale al valore dell'"indice di superficie"; 2) la posizione di ogni insediamento è indipendente da tutte quelle degli altri. Si tratta, come si vede, di un modello che tiene conto soltanto di variabili ambientali ("environmental") e non c'è da stupirsi pertanto (Chadwick, 1978, p.53) che non dia assolutamente la distribuzione spaziale degli insediamenti osservata dalla spedizione dell'Università di Minnesota.

L'Autore tenta allora di eseguire la simulazione con un modello più complesso le cui regole essenziali sono: la posizione dell'insediamento è casuale ma la probabilità di "acces-

so" a ogni unita' di "griglia" (2×2 km) dipende da tre fattori: a) il valore dell'"indice di superficie" trasformato pero' in modo da tener conto di quello delle celle vicine; b) la sua posizione rispetto a tutti i siti esistenti, e c) il numero totale di insediamenti.

In fig.66 si vede un risultato di simulazione, basato su queste premesse, per il tardo Elladico che, almeno qualitativamente, rappresenta piuttosto bene il "pattern" dell'insediamento, come si puo' confrontare con la fig.65.

Sempre in tema di modelli di simulazione, volevo qui citarne uno di tipo un po' diverso, ed e' quello proposto da H. Wobst (1974), con l'intento di stabilire l'estensione territoriale e l'entita' numerica (nonche' la sua *stabilita'* nel tempo) di una cultura di cacciatori e raccoglitori del Paleolitico, mediante un modello stocastico di simulazione che usa il metodo di Montecarlo.

Il modello ha alcune ipotesi di base, che vengono mantenute fisse, e cioe': 1) il gruppo culturale e' costituito da una "macrobanda" articolata in "microbande", che sono a loro volta costituite da una o piu' unita' familiari; 2) il gruppo e' "territoriale", nel senso che si muove in un'area delimitata sia da ostacoli naturali che da confini "culturali", ed in cui si trova a suo agio, anche per la familiarita' con l'ambiente circostante. Si assume inoltre un modello geometrico ad esagoni (la struttura piu' regolare), come organizzazione geografica del gruppo; 3) il sistema culturale articola le sue componenti in risposta agli stimoli dell'ambiente esterno, e si suppone che le dimensioni numeriche della "macrobanda" siano mantenute a livello di equilibrio, bilanciando le strategie di rifornimento del cibo, il sistema di organizzazione dei matrimoni (endogamia od esogamia, ecc.), il tasso di natalita', ecc.

Tutti questi parametri vengono ovviamente variati durante la simulazione tramite calcolatore elettronico.

I risultati dipendono chiaramente dal valore di questi pa-

rametri (oltre a quelli già citati vengono considerati ad esempio la monogamia e la poligamia, la media della vita degli adulti, la presenza o assenza del tabu' dell'incesto, ecc.). Si

Late Helladic

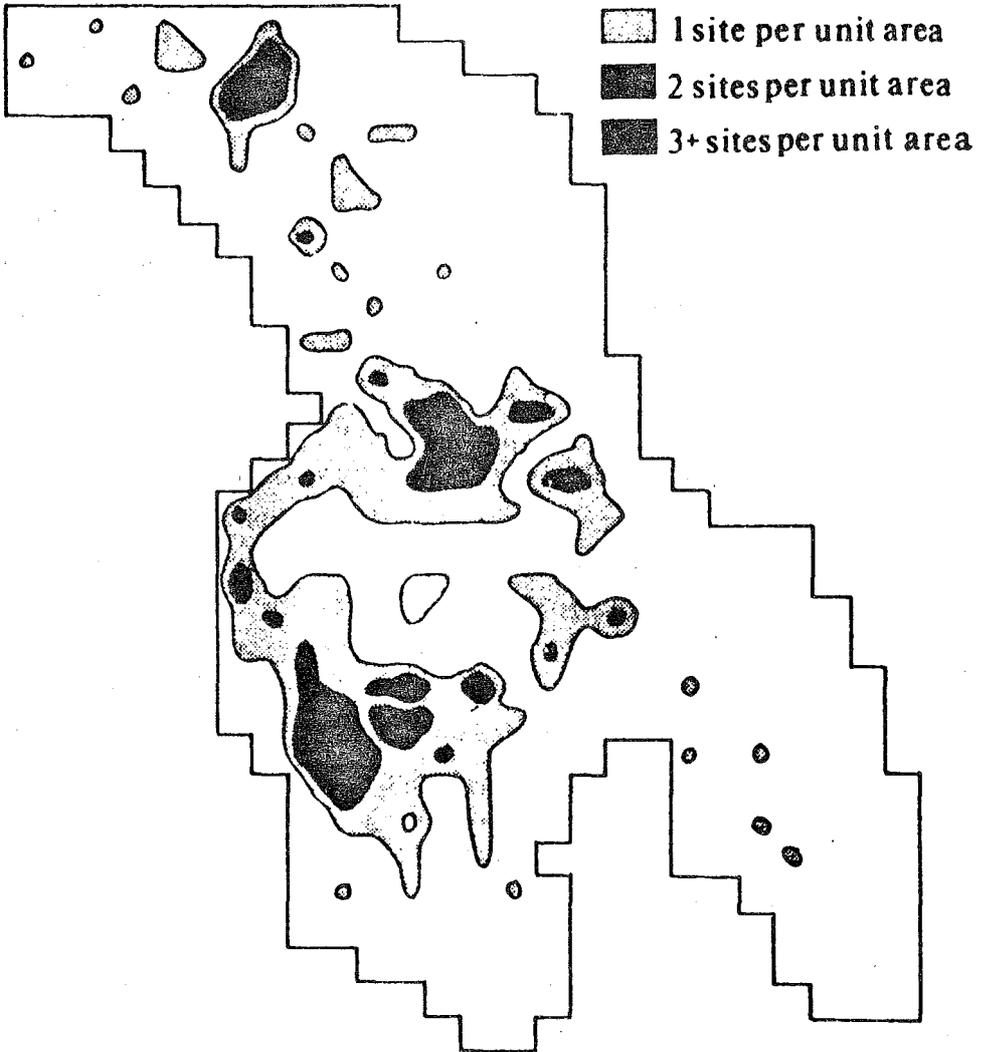


Fig.66

Modello di insediamento simulato per il tardo Elladico (da Chadwick, 1978).

trova comunque una dimensione minima di equilibrio della "macrobanda" che può variare da circa 200 a 500 membri; e pertanto, con una dimensione media di 25 persone per la "microbanda" (composta da due a quattro nuclei familiari, circa), come è ricavato anche da osservazioni etnografiche recenti, si ricava che un gruppo culturale, rappresentato dalla "macrobanda", è costituito da un insieme variabile da circa 8 a circa 20 "microbande".

H. Wobst definisce anche una "vita media" del gruppo culturale: il numero di anni richiesto perché la metà del nucleo originale di cacciatori e raccoglitori abbia cessato di esistere. Mediante il programma di simulazione trova un valore dell'ordine di 180 anni, per "macrobande" di dimensioni minime di equilibrio. Questo valore, anche se si tratta di una stima grossolana, permette, per una "microbanda" di circa 25 persone, come si è già detto, una sopravvivenza nel territorio per un gran numero di generazioni.

Questo modello prevede anche delle dimensioni medie del territorio occupato dal gruppo culturale, ed è stato applicato per le culture mesolitiche dell'Europa settentrionale da T.D. Price (1978), che ha mostrato come, in assenza di ostacoli naturali, si possa stimare un'occupazione di un territorio con un raggio medio di circa 150 km, in accordo sostanziale con la distribuzione dei giacimenti del Mesolitico osservato nella Germania settentrionale.

* * *

Fino a questo momento abbiamo dato una base panoramica dei metodi che si possono usare per la interpretazione di correlazioni *spaziali* sia di materiali su di una superficie, che di insediamenti in una regione tramite opportuni modelli matematici.

Voglio invece ora occuparmi dei problemi di correlazione *temporale* e cioè di cronologia relativa tra diversi giacimen-

ti, tombe di una necropoli, ecc., nonché di modelli *dinamici* di processi culturali e cioè dell'evoluzione e/o della diffusione temporale delle culture.

Il problema della cronologia relativa tra entità archeologiche, potrebbe sembrare oggi di facile soluzione, visto che si hanno ormai a disposizione numerose tecniche "sussidiarie" di datazione: analisi sedimentologiche, paleobotaniche, ecc., per non parlare dei vari metodi, ormai sempre più diffusi, di datazione assoluta.

Basti pensare, ad esempio, che l'uso del metodo del C^{14} con l'opportuna calibrazione tramite la dendrocronologia, ci permette di avere ormai date sicure, con un errore statistico che è al massimo dell'1%, fino a circa 7000 anni da oggi!

Si può però giustamente obiettare che non sempre si hanno a disposizione materiali databili (basti pensare ad esempio a complessi di superficie dove non sono conservati né carboni né ossa), e in questo spirito parlerò brevemente dei metodi di *seriazione* cronologica, che si basano essenzialmente sull'uso della tipologia per riconoscere sequenze temporali tra complessi.

Si possono distinguere grosso modo due diversi tipi di approccio: il primo si basa su un'"affinità globale" tra complessi, e quindi su una base *politetica* di confronto tra contenuti dei vari complessi, mentre il secondo approccio tende ad identificare, quantitativamente, alcuni *tipi* (litici, ceramici, ecc.) che possono caratterizzare *cronologicamente* diversi complessi, in qualche modo *diagnostici* o "fossili direttori" come si è già accennato nel capitolo precedente.

Debbo subito dire che varie critiche (che in gran parte io personalmente condivido) sono state sollevate sulla validità di entrambi i metodi: il capitolo "Automatic Seriation" del già citato libro di J.E. Doran e F.R. Hodson (1975) mi sembra che renda conto abbastanza bene dei problemi di interpretazione che sorgono, in generale, dalle tecniche di seriazione.

Il primo metodo, "politetico" consiste in sostanza in un problema di "cluster analysis".

Come esempio riprendiamo le matrici dei coefficienti di similarita' tra sei complessi (tombe, giacimenti, ecc.) gia' viste nelle tabelle n.4 e n.5 del capitolo precedente, quando si parlava del "riordinamento" degli elementi della matrice. Come si e' gia' detto, il metodo e' stato introdotto da G.W. Brainerd (1951) e W.S. Robinson (1951) proprio allo scopo di raggruppare cronologicamente complessi a seconda del valore del coefficiente mutuo di similarita'.

In quest'ottica, i complessi (tombe, giacimenti, ecc.) cronologicamente piu' vicini avremmo coefficienti di similarita' comparabili e piu' alti. La matrice riordinata ("sorted", come quella rappresentata in tab.5) avra' pertanto, come si e' gia' detto, i valori di coefficiente di similarita' piu' alto vicini alla diagonale principale. Per questa ragione, nella matrice riordinata, dovrebbe esserci una "contiguita'" tra i valori dei coefficienti di similarita' di giacimenti cronologicamente vicini, e pertanto, la somma di questi coefficienti per ogni colonna della matrice (o riga, visto che la matrice e' simmetrica) dovrebbe mostrare un valore crescente fin verso la meta' delle colonne (nella tab.5 il massimo valore (4.0) della somma si ha appunto per la 4^a colonna) per poi decrescere successivamente.

E' proprio questo comportamento della somma dei coefficienti di correlazione delle varie colonne (o righe) la base del metodo di approssimazioni successive, a partire dalla matrice "non ordinata" dei coefficienti di similarita', introdotto da G.W. Brainerd e W.S. Robinson (1951), e che ha dato luogo a numerosi programmi di calcolatore abbondantemente usati da vari autori, anche se con risultati alterni.

Da un punto di vista matematico, come e' chiaro, il problema e' semplicemente di "cluster analysis" di tipo Q: invece del riordinamento della matrice dei coefficienti di similarita' si potrebbe benissimo costruire un dendrogramma, secondo lo schema

del capitolo precedente.

Le difficoltà interpretative nascono invece, come è ovvio, da ragioni paramente paleontologiche e archeologiche: non è detto che complessi culturali *tipologicamente vicini*, sulla base di un opportuno coefficiente di similarità (e qui bisognerebbe anche discutere le *variabili* considerate nel calcolo del coefficiente!) debbono essere necessariamente *cronologicamente vicini*.

Il secondo metodo, come si è detto, cerca invece di individuare elementi "cronologicamente diagnostici" e quindi, in qualche modo è più un metodo "gerarchico" perché a certi tipi (che sono quelli "diagnostici") si dà una preponderanza su certi altri.

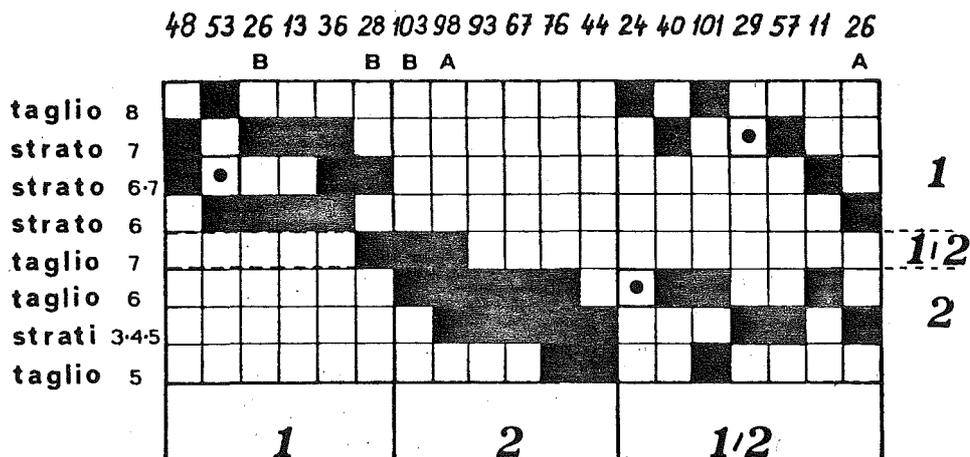
I principi di questo metodo sono stati probabilmente posti per la prima volta da W.F. Petrie nel 1899 per riordinare un insieme di tombe egiziane (Petrie, 1899) e, molto più tardi, sistematizzati da D.G. Kendall (1963).

In Italia, la più semplice e nota espressione di **seriazioni** di questo genere sono le così dette *tabelle di associazione* (ottenute generalmente a mano, con metodi empirici) che hanno come variabili una sequenza di tipi (righe, ad esempio) e una sequenza di complessi (colonne quindi) ordinati mediante presenza o assenza dei tipi in questione, che appunto sono considerati "diagnostici" (per le necropoli della cultura laziale v. ad esempio Peroni, 1959; Müller-Karpe, 1962).

In fig.67, vediamo, ad esempio, una tabella di associazione per diversi strati o tagli dell'Età del Bronzo (appenninico) di Palidoro, presso Roma (Fugazzola - Delpino, 1976: si tratta degli strati a ceramica soprastanti il giacimento del Paleolitico superiore già citato). In "ascisse" (colonne) vengono dati i tipi ceramici (ad esempio il n.48 è un'olla carenata e il n.28B è una scodella ad orlo rientrante e larga (Fugazzola-Delpino, 1976, p.223)), mentre in "ordinate" (righe) vi sono gli strati o tagli del giacimento: vi sono alcuni tipi

quindi che sono diagnostici per una distinzione cronologica degli strati in due periodi, mentre altri (n.24, 40, 101, 29, 11 e 26A) non danno luogo a possibili seriazioni cronologiche.

In tav. III vediamo invece una tabella d'associazione per le necropoli dell'antica cultura laziale (Müller-Karpe, 1962, tav. 44) usando come tipi diagnostici sia le ceramiche che i bronzi rinvenuti nelle tombe.



■ CERTI
 □ ● VARIANTI

PALIDORO

Fig.67

Tabella d'associazione per gli strati appenninici di Palidoro (da Fugazzola-Delpino, 1976).

Gli esempi che abbiamo citato si basano solo sulla presenza o assenza dei tipi, perché come abbiamo già detto nel capitolo precedente, per la preistoria recente non si tiene conto, in generale, delle loro percentuali.

Se si introduce una variabile quantitativa, come frequenze o percentuali, è chiaro che le colonne delle tabelle di associazione si "allargano", e cioè al posto di una crocetta o

di un quadratino nero, come in fig.67, abbiamo una striscia, la cui larghezza rappresenta la consistenza quantitativa (per esempio la percentuale) in quel complesso.

Se si crede allora che i tipi "diagnostici" hanno una concentrazione temporale abbastanza ben definita, e se i complessi (giacimenti, strati, necropoli, ecc.) sono cronologicamente ordinati, si dovrebbe osservare una massima "larghezza" per ogni tipo, in corrispondenza del suo massimo "cronologico" e una conseguente diminuzione per i complessi sia precedenti che posteriori temporalmente.

Si ottiene così una struttura, per ogni tipo considerato, di tipo ovale, o, come viene generalmente definita "battleship" (letteralmente: nave da battaglia) come si può vedere dalla

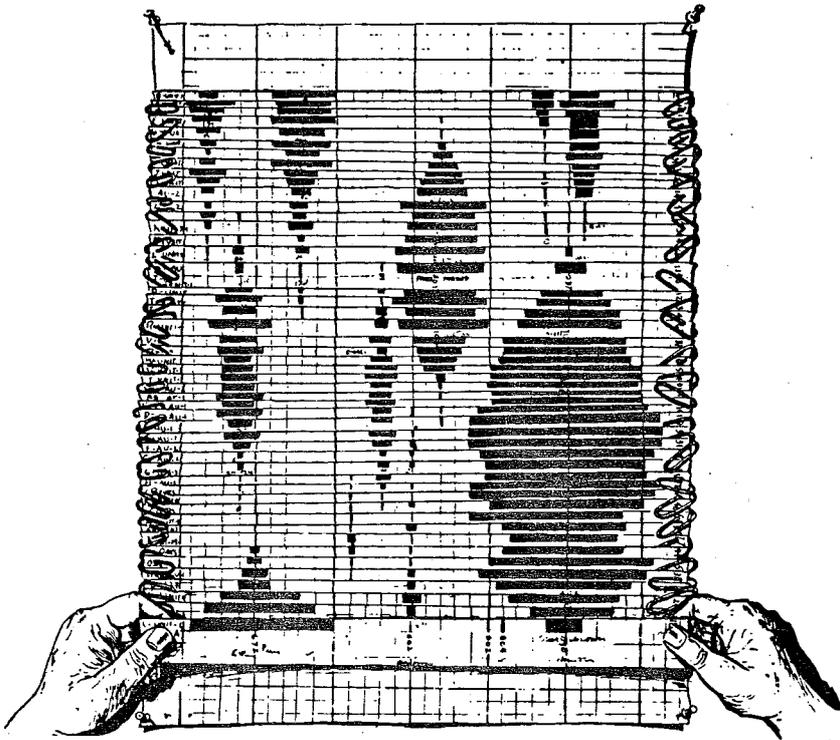


Fig.68

Schema di seriazione con il metodo delle "battleships" (da Ford, 1962).

fig.68 (Ford, 1962) che mostra appunto un tentativo di serie-
ne "a mano" su questa base.

Per citare un esempio concreto, in fig.69 abbiamo l'illu-
strazione della variazione in percentuale di alcuni tipi liti-
ci, considerati "diagnostici", per diversi strati del giacimen-
to mesolitico di Romagnano III (Broglia, 1975). I tipi sono es-
senzialmente dei microliti geometrici, e, come si vede, una grossa

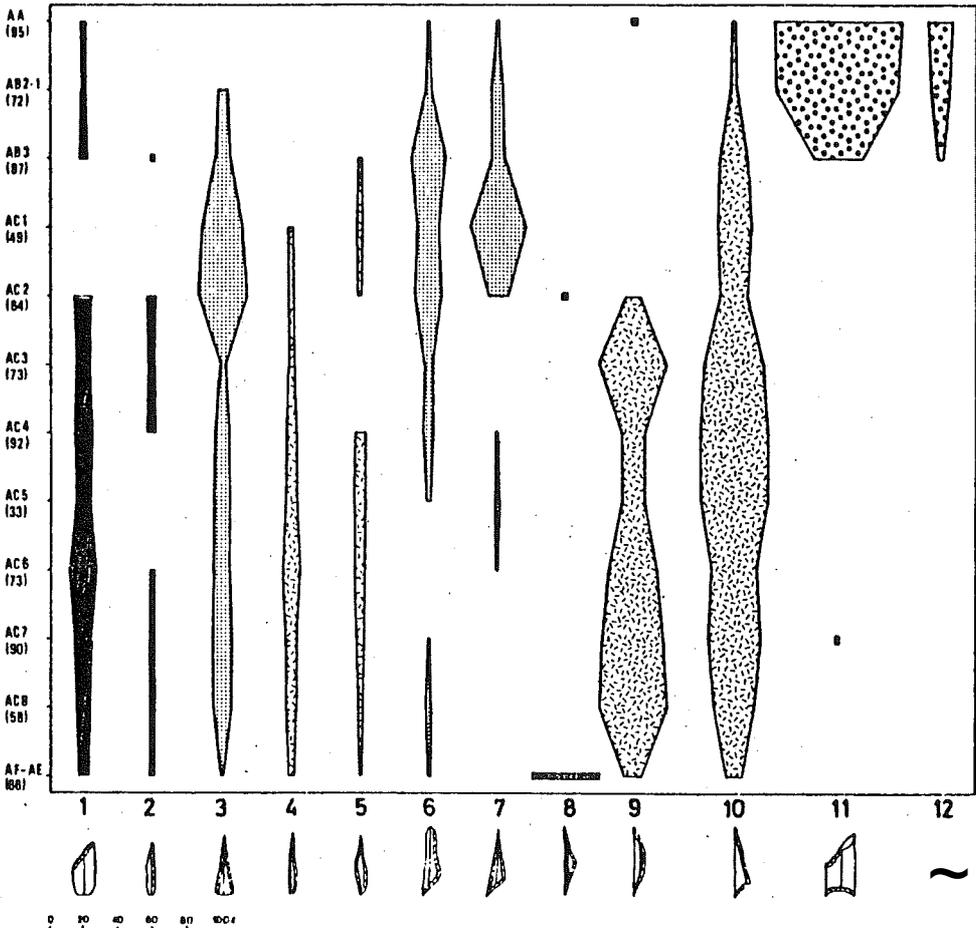


Fig.69

Tabella d'associazione per alcuni strati del giacimento mesoli-
tico di Romagnano (da Broglia, 1975).

differenziazione cronologica si ha per i trapezi e romboidi (n.11 e 12 in fig.69) caratteristici dei complessi "tardenoidi" AB2-1 e AA, mentre altri microliti sono presenti (anche in percentuale significativa) in piu' strati, con l'eccezione del triangolo ritoccato su tre lati (n.8) presente praticamente solo nei livelli inferiori (AF-AE) del complesso "Sauveterroide".

Da un punto di vista piu' strettamente matematico il metodo di seriazione mediante le "battleships" e' stato oggetto di varie ricerche, e voglio qui citare solo quella di P. Ihm, che si serve sostanzialmente di un modello *variazionale* di "maximum likelihood" (massima verosimiglianza) per la stima dei parametri (Ihm, 1976). Inoltre, si puo' dimostrare che il metodo delle "battleships" (e cioe' le tabelle di associazione "quantitative", come si e' descritto) e' sostanzialmente equivalente, matematicamente e come contenuto di informazione, al primo metodo di seriazione, e cioe' quello di Brainerd-Robinson del riordinamento delle righe e colonne della matrice di similarita' (v., ad esempio, Doran e Hodson, 1975, pp.279-280).

In ogni modo, per concludere, penso che, a parte qualsiasi critica sull'impostazione teorica, di cui abbiamo gia' discusso, l'interesse per i metodi di seriazione cronologica stia diminuendo in modo sostanziale con il passare del tempo, proprio perche' i metodi di datazione assoluta forniscono risultati sempre piu' abbondanti e piu' attendibili, cosi' che forse si assiste ad un recupero dei manufatti (tipi litici, ceramici, ecc.) come *indicatori culturali* piuttosto che *cronologici*.

* * *

La ricostruzione delle culture del passato, e', come e' noto, una operazione al tempo stesso *sincronica* e *diacronica*, e, nella prima parte di questo capitolo, ci siamo occupati essenzialmente di procedimenti e modelli in una prospettiva piu' o

meno sincronica.

Prendo invece l'occasione dall'ultimo paragrafo, che trattava di problemi di cronologia mediante tecniche di seriazioni, per accennare a *modelli dinamici* di culture, e cioè modelli che si propongono lo studio dei *processi culturali* e quindi lo sviluppo nello spazio e nel tempo delle culture.

I due modelli che voglio citare riguardano essenzialmente il problema della diffusione spazio-temporale di elementi culturali, anche se, in linea di principio, lo sviluppo spazio-temporale di una o più culture, a seconda dei casi, potrebbe avere forse più semplici spiegazioni di carattere evoluzionistico-policentrico.

Non è questa la sede per discutere più in dettaglio la validità di queste due impostazioni, basterà forse solo ricordare che in generale l'ipotesi evoluzionistica è stata maggiormente impiegata per i processi culturali del Paleolitico, mentre per età più recenti è quella diffusionistica ad avere il maggior campo di applicazione.

A mio avviso, come insegna anche il dibattito antropologico della prima metà di questo secolo, entrambe le ipotesi possono essere valide, caso per caso, purché ci si renda conto precisamente del *contesto* in cui vengono espresse. In altre parole, la formulazione di ipotesi di questo genere deve essere successiva ad una rigorosa analisi sincronica di *tutti* gli elementi che concorrono alla definizione di una cultura.

Il primo esempio che voglio citare è un modello di diffusione dell'agricoltura, a partire dal Medio Oriente, descritto da A.J.A. Ammermann e L.L. Cavalli Sforza (1973).

Il punto di partenza è un grafico, decisamente suggestivo, rappresentato in fig.70, della relazione tra tempi e distanze di insediamenti del Neolitico antico in Europa, a partire da un "centro diffusore" (Gerico in Palestina, ad esempio, datato a circa 9000 anni da oggi). Come si vede dalla figura la relazione è lineare e il coefficiente di correlazione,

definito dall'eq. (1.6) e' particolarmente alto (0.89).

Il modello si basa sul concetto di diffusione "demica" (Ammermann e Cavalli-Sforza, 1973, p.344) e cioe' sostanzialmente dovuta ad un incremento demografico, causato, secondo gli Autori, da una migliore distribuzione delle risorse motivata dal-

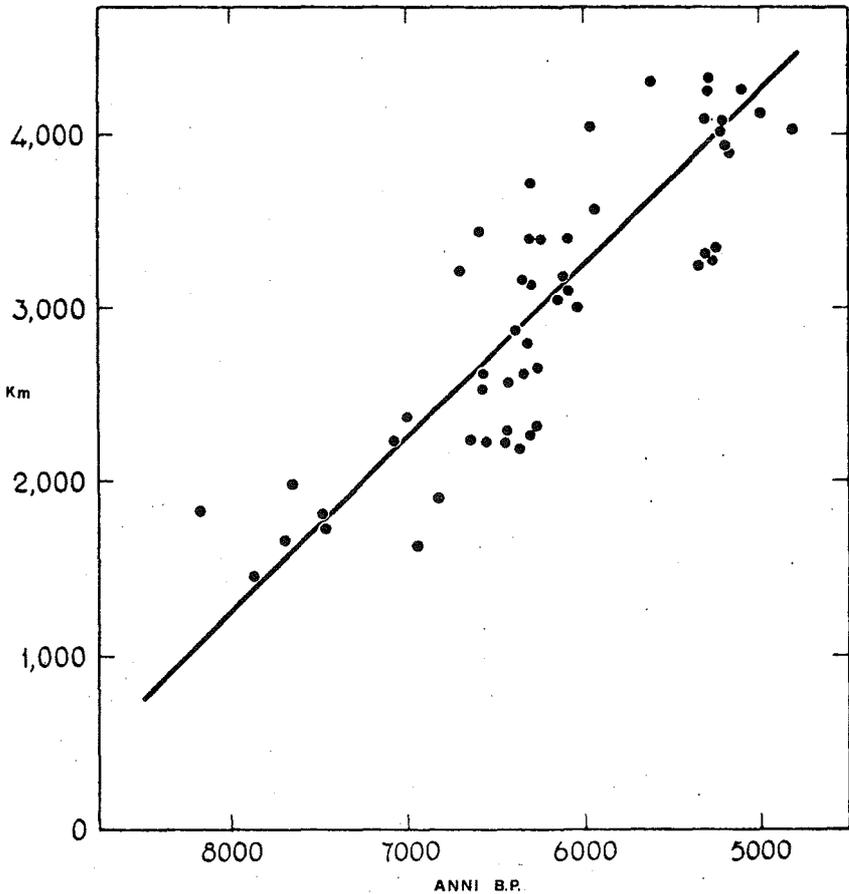


Fig.70

"Scatter diagram" e retta di regressione per i giacimenti del Neolitico antico in Europa. In ascisse l'eta' da oggi (in migliaia di anni); in ordinate la distanza del centro diffusore (Gerico, in questo caso) in migliaia di km (da Ammermann e Cavalli-Sforza, 1973).

lo sfruttamento produttivo, e quindi pianificato, dell'ambiente circostante tramite la coltivazione delle piante (in particolare cereali) e l'allevamento di animali.

Nel modello si postula un'origine di diffusione a partire da un unico centro situato essenzialmente nel Medio Oriente, anche perché non vi è prova di esistenza di specie di cereali coltivabili (come il *Triticum monococcum* o *dicoccum*) allo stato selvatico in Europa. Il tipo di diffusione è sostanzialmente considerato come una estensione lenta di nuove aree da coltivare, quasi come una successione di stati di equilibrio, secondo un modello, già applicato in problemi di genetica, di *onda di avanzamento* ("wave of advance"), basato su di un'equazione differenziale del tipo

$$\frac{\partial \psi}{\partial t} = \alpha f(\psi) + \beta \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right)$$

dove ψ rappresenta la densità di popolazione, t è il tempo e β è il coefficiente di diffusione, collegato con il parametro α alla velocità di avanzamento del fronte d'onda, e cioè alla velocità di espansione geografica dell'agricoltura a partire dal centro diffusore. La funzione $f(\psi)$ può essere scelta uguale a ψ per una crescita esponenziale della densità di popolazione (molto improbabile), oppure uguale a $1-\psi$, per una crescita che va in saturazione dopo un certo tempo. Scegliendo ad esempio il valore di saturazione di 5 abitanti per km^2 , con $\alpha = 0.039$, partendo da una densità iniziale di 0.1 abitanti per km^2 occorrono circa 200 anni per ottenere il valore di saturazione per ψ .

Il problema è naturalmente quello di verificare sui dati archeologici la validità di questo modello, e soprattutto la sua consistenza con certe ipotesi sui parametri, come ad esempio il già citato tasso ("rate") di avanzamento del fronte d'onda $\rho = 2\sqrt{\beta\alpha}$, che è una costante, secondo il modello. In linea

di principio i dati archeologici dovrebbero dare delle indicazioni sul valore di questi parametri, ma non bisogna dimenticare che il modello parte da ipotesi abbastanza semplificate, tra cui principale mi sembra quella di una continua espansione demografica (anche se, al limite, in saturazione).

E' chiaro che gli Autori stessi riconoscono comunque che la linearita' spazio-temporale per la diffusione dell'agricol-

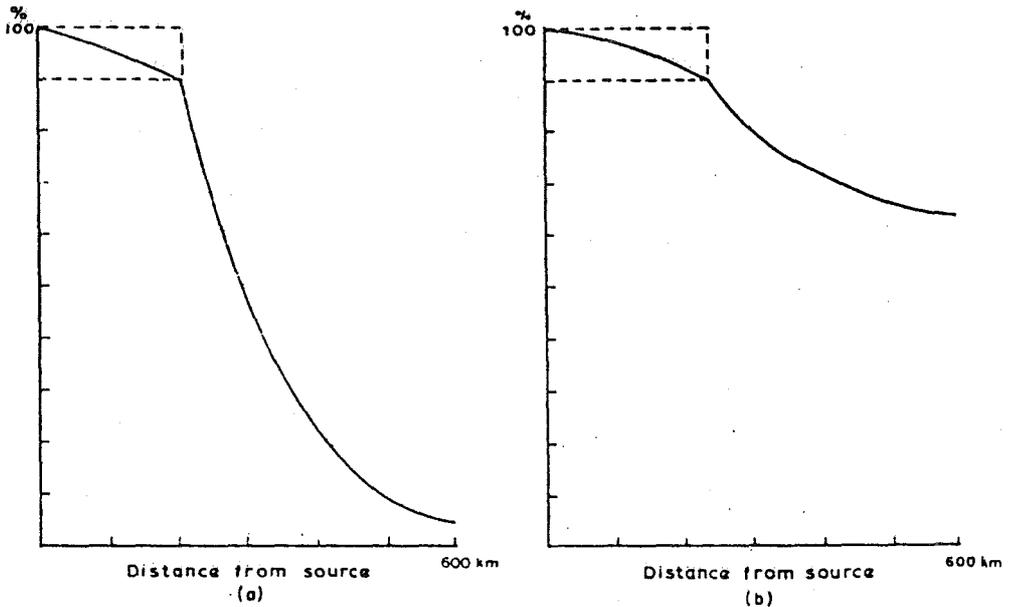


Fig.71

Curve di "fall-off": a) per i beni di consumo normale, b) per beni di prestigio. La zona tratteggiata rappresenta l'area di produzione (da Renfrew, 1972).

tura in Europa a partire dal Medio Oriente, illustrata nella fig.70, puo' essere spiegata non solo tramite la loro ipotesi di incremento demografico, ma anche per semplice scambio culturale nell'area di contatto tra culture diverse.

Un altro esempio di impiego di modelli matematici e' lo studio di possibili alternative di scambio, con conseguente ef-

fetto sulla distribuzione spaziale, per diverse classi di oggetti (ceramici, litici, ecc.) nell'area mediterranea a partire dal Neolitico (Renfrew, 1977). In un precedente lavoro l'Autore (Renfrew, 1972) aveva ipotizzato che a partire dal centro di produzione, la quantità dei beni scambiati subisce una *attenuazione* e che la rapidità dell'attenuazione (il "fall-off") dipende dal tipo di beni: per beni di consumo normale, si può ipotizzare un "fall-off" piuttosto rapido (v. fig. 71a)), ed è il cosiddetto modello di scambio "down-the-line", mentre per beni di prestigio il "fall-off" è molto più lento (v. fig. 71b)), e si avrebbe uno scambio di tipo "prestige-chain".

Per quanto riguarda i beni di consumo normale, e cioè con uno scambio del tipo "down-the-line" si possono creare modelli matematici, e quindi tipi di curve di "fall-off" diversi, a seconda delle ipotesi sulla densità di popolazione e sul tipo di decremento del bene scambiato, la cui quantità (funzione della distanza) possiamo indicare con I .

Se la variazione di questa quantità dI è indipendente dal suo valore $I(x)$ ad una data distanza x dal centro di diffusione, avremo per densità di popolazione costante, in una sola dimensione

$$dI = -kdx \quad (3.12)$$

dove dx è appunto la variazione di distanza che corrisponde alla variazione di quantità dI , k è una costante connessa alla densità di popolazione, e il segno $-$ indica appunto che al variare della distanza la quantità I *diminuisce*.

Questa semplicissima equazione differenziale ha come soluzione

$$I = I_0 - kx$$

dove I_0 è la quantità iniziale di oggetti scambiati: il "fall-off" è quindi lineare.

L'ipotesi di indipendenza della variazione dI dalla quantità stessa non è molto realistica. Se si assume, come è ragionevole, una relazione lineare avremo, invece della (3.12) l'equazione differenziale

$$dI = -k I dx \quad (3.13)$$

che integrata, dà

$$I = I_0 e^{-kx} \quad (3.14)$$

cioè un "fall-off" esponenziale.

Questi risultati valgono per una dimensione, come se lo scambio, in altre parole, avvenisse lungo una linea. Se invece si considera una situazione bidimensionale, come può essere quella reale (sempre che il processo di scambio abbia un supporto geografico omogeneo, e cioè non vi sono ostacoli come montagne alte o grandi fiumi), le equazioni possono essere diverse.

Infatti, ad esempio, la quantità I può ora essere funzione delle due coordinate cartesiane x, y , $I(x, y)$, e quindi al posto della eq. (3.13) si potrebbe supporre un'espressione del tipo

$$dI = -k I dx dy \quad (3.15)$$

dove appunto dx e dy rappresentano la variazione della distanza nelle due dimensioni x e y .

Se ora si suppone un processo di scambio *isotropo* e cioè che dipende solo dalla distanza totale $r = \sqrt{x^2 + y^2}$, e non dalla direzione, passando in coordinate polari l'eq. (3.15) diventa

$$dI = -h I r dr \quad (3.16)$$

dove $h = 2\pi k$. La soluzione di quest'equazione dà ora un "fall-off" *gaussiano* e più precisamente

$$I = I_0 e^{-\frac{hr^2}{2}} \quad (3.17)$$

che è tra l'altro in accordo anche con ipotesi di diffusione di tipo stocastico ("random walk" o "flight", v. Renfrew, 1977, pp.79-80).

Nel caso di variazione dI, indipendente dal valore di I (l'eq. (3.12) nell'approssimazione unidimensionale) invece della (3.17) avremo una espressione *parabolica* e cioè

$$I = I_0 - hx^2$$

E' ovvio che se non si suppone piu' una densità di popolazione costante ma che varia in qualche modo con la distanza, allora si avranno modelli matematici e "fall-off" diversi a seconda del tipo di funzione che descrive la densità di popolazione. Se, ad esempio, si avesse una relazione con la distanza di proporzionalità inversa, come ad esempio

$$h = \frac{a}{r}$$

è chiaro che l'eq. (3.16) diventerebbe

$$dI = -a I dr$$

e pertanto il "fall-off" sarebbe di tipo esponenziale (v. eq. (3.14)).

Il problema di modelli di questo tipo è, in generale, la scarsità di dati archeologici che possono permettere di "accettarlo" o di "rifiutarlo" (è il procedimento dell'"hypothesis testing" molto caro ai "new archaeologists" anglosassoni).

In ogni modo, sono essenzialmente applicabili a problemi di diffusione di elementi di tipi ceramici, o di materie prime, come l'ossidiana nell'area mediterranea a partire dal Neolitico, come è documentato dall'intensa attività di ricerca svolta in questo campo da C. Renfrew.

* * *

Non mi dilungherò ulteriormente nell'illustrare altri modelli matematici di culture, essenzialmente per ragioni di spa-

zio. Vorrei però, su quest'argomento, svolgere alcune considerazioni conclusive.

E' probabilmente chiaro a chiunque che la maggior parte di questi modelli hanno un valore interpretativo ("explanatory" nel senso dei "new archaeologists") abbastanza limitato e, di conseguenza, anche possibilità *predittive* piuttosto modeste; trattandosi di scienze umane, ed in particolare di discipline come quelle archeologiche, dove la raccolta di informazioni è naturalmente incompleta, questo genere di modelli è ovviamente di carattere troppo schematico per tenere conto della variabilità e delle correlazioni tra innumerosi elementi che concorrono a definire una cultura e il suo sviluppo.

Penso però che l'interesse scientifica della formulazione di modelli, risieda in ogni caso nel fatto che costituiscono quanto meno un tentativo di costruzione di *teorie archeologiche*, di cui si avverte, a mio avviso, sempre più la necessità nel momento presente, sia per la quantità di dati da elaborare, che è in continua espansione, e sia perché la parte "sperimentale" di raccolta dei dati suddetti, insieme con le informazioni complementari (geologiche, paleontologiche, paleobotaniche, datazioni assolute, ecc.) è ormai giunta ad un notevole livello tecnico di accuratezza ed affidabilità.

E' quindi il momento di pensare a schemi concettuali, ad angolazioni e punti di vista teorici, in una parola, per usare il famoso termine di T.S. Kuhn (1970) a *paradigmi*, tali da coinvolgere la comunità scientifica degli archeologi (ovviamente con le loro diverse specializzazioni) in un'attività di *scienza normale* (sempre per usare le parole di T.S. Kuhn) con tutto ciò che essa comporta: l'attività sperimentale di scavo e raccolta dei materiali, la "fenomenologia" e cioè l'analisi dei dati (l'"explication" degli anglosassoni) e la costruzione di *teorie o modelli interpretativi* che si propongono appunto, nel quadro generale del paradigma dominante, di "spiegare" i dati con un valore anche, se è possibile, predittivo.

La prima obiezione che puo' essere posta ad uno schema di questo genere e' ovviamente che si tratta del risultato di una analisi storica, come quella di T.S. Kuhn, essenzialmente centrata su comunita' scientifiche che oggi sono altamente sviluppate, come la fisica o la chimica, ad esempio.

Non ho alcuna difficolta' a condividere questa obiezione, e del resto la mia sensazione e' che nel campo dell'archeologia, almeno in quella preistorica, ci troviamo in una situazione che si potrebbe definire "preparadigmatica" o "protoparadigmatica" (v. ad esempio Bietti, 1976-1977; 1978a) che pero' rivela, a mio giudizio, l'embrione di una *comunita' scientifica* (almeno dei preistorici, che si riconoscono appunto nella Paleontologia o nell'Ecologia preistorica, come accennero' piu' oltre).

Bisogna inoltre osservare come gia' da qualche anno, specialmente nei paesi anglosassoni, il dibattito "epistemologico" e anche per un'analisi storica delle scienze archeologiche (v. ad esempio Sterud, 1973) sia molto vivo, sia per quanto riguarda una loro struttura paradigmatica, sempre nel senso di T.S. Kuhn (anche se con qualche fraintendimento concettuale, a mio giudizio: v. ad esempio Clarke, 1972; Binford, 1972) che per impostazioni di carattere piu' astratto o logico-deduttivo (v. ad esempio Fritz e Plog, 1970).

Un'altra critica alla introduzione di *teorie e modelli*, viene probabilmente da coloro che vedono con preoccupazione in questa circostanza, un processo progressivo di *divisione del lavoro*: si potrebbe formare una categoria di "archeologi teorici" rispetto agli "archeologi sperimentali" che sono sostanzialmente coloro che effettuano gli scavi e classificano i dati.

L'esempio piu' vistoso di una divisione del lavoro di questo genere e' chiaramente quello della fisica, dove la "professione" di fisico teorico, per particolari ragioni storiche e sociali si e' andata affermando dalla fine del secolo scorso (J.C. Maxwell ne e' un esempio).

In compenso bisogna pero' dire che l'esistenza e la neces-

sita' di una *teoria*, sia nella fisica che nella chimica, era già ben stabilita e compresa fin dall'inizio della storia di queste discipline come scienze: nessuno può negare l'elaborazione spiccatamente teorica, in termini di "leggi" o "principi", ottenuta a partire dall'esperienza da parte di scienziati tipicamente "sperimentali" come Galileo o Lavoisier, ad esempio.

Temo invece che sia proprio la mancanza di chiarezza su un possibile significato *scientifico* delle discipline archeologiche, anche ad un livello "paradigmatico" nel senso di T.S. Kuhn, dovuta probabilmente alla loro natura prevalentemente "umanistica" (almeno in Europa) a favorire *artificialmente*, a mio giudizio, possibili divisioni del lavoro.

Si assiste infatti ad un fiorire di generalizzazioni arbitrarie o di frammentazioni particolari o settoriali che limitano in modo drastico le prospettive di studio dei reperti archeologici.

Mi riferisco, come esempio di prospettiva "totalizzante" ad un recente libro di J.C. Gardin (1979), che significativamente si intitola "Une Archéologie Théorique", introducendo quindi un'archeologia come "disciplina universale", al di là dello spazio e del tempo. Secondo il mio punto di vista, la prospettiva da seguire è *esattamente opposta*: l'archeologia è un termine generico che indica sostanzialmente una disciplina sperimentale di raccolta e classificazione di reperti del passato, ed è chiaro anche che esistono differenti archeologie: preistorica, classica, medioevale, ecc., ciascuna con una sua *teoria necessaria* (ed eventualmente un "paradigma"), che ovviamente sarà diversa a seconda del campo d'indagine: basti pensare, ad esempio all'importanza delle fonti storiche per l'Archeologia classica, che mancano invece totalmente nel caso della preistoria.

Per quanto riguarda invece la *settorialità* di certi indirizzi di ricerca, uno degli esempi più clamorosi è quello fornito, secondo me, dai lavori di E.S. Higgs e della sua scuola.

Basta già il titolo ("Papers in economic prehistory") di una raccolta di lavori (Higgs, 1972) a rendere chiaro l'interesse su un elemento particolare ed isolato delle culture preistoriche, mentre, a mio avviso, se si vuole una loro *ricostruzione globale* si dovrà parlare di *economia preistorica* come uno dei tanti elementi che concorrono alla formazione di un quadro generale.

Da quest'ultimo esempio (non sono in grado di parlare dell'archeologia classica o medioevale, per ragioni di competenza, come ho già detto all'inizio) mi pare chiaro quale è, secondo me, la via da intraprendere per stabilire *teorie* o *modelli* di interpretazione globale delle culture preistoriche: lo studio delle prospettive fornite dall'*Antropologia culturale classica* (mi riferisco ovviamente a quella anglosassone ed in particolare a quella statunitense).

Non voglio dire con questo che si debbano necessariamente impiegare i modelli forniti dalle società primitive "attuali" come schema interpretativo rigido di culture preistoriche (ad esempio la cultura degli Esquimesi come modello delle culture del Maddaleniano francese), ma invece sottolineare l'importanza della *metodologia* antropologica come sorgente di idee per la costruzione di modelli e teorie di interpretazione del passato.

E voglio chiudere queste lezioni proprio con una citazione del celebre articolo "L'Antropologia come Storia" di C.M. Arensberg (1957): "Ancora oggi l'antropologia continua a dividersi nelle sue tre branche originarie dell'antropologia culturale, dell'archeologia e dell'antropologia fisica - lo studio cioè dell'evoluzione della specie umana dalle specie animali - che continuano ad arricchirsi vicendevolmente. Proprio questo interesse per i processi e le sequenze temporali, questo confronto del passato con il presente le mantiene unite" (Arensberg, 1957, p.119 della traduzione italiana (1978)).

BIBLIOGRAFIA

- AMMERMANN A.J., CAVALLI SFORZA L.L. (1973) - *A population model for the diffusion of early farming in Europe*, in: "The explanation of Culture Change: models in Prehistory", (ed. da C. Renfrew), Duckworth, pp.343-357.
- ARENSBERG C.M. (1957) - *L'Antropologia come Storia*, in: "Trade and Market in the Early Empires", (a cura di K. Polanyi), Free Press. Traduzione italiana: "Traffici e mercati negli antichi imperi", Einaudi, 1978.
- BENZECRI C.P. (1973) - *L'analyse des données*, 2 vol., Dunod, Parigi.
- BIETTI A. (1974-75) - *Attribute analysis from a "Traditionalist" point of view: an example from an Italian upper Paleolithic assemblage*, "Quaternaria", **18**, pp.45-65.
- BIETTI A. (1976-77) - *Analysis and Illustration of the Epigravettian Industry Collected during the 1955 Excavations at Palidoro (Rome, Italy)*, "Quaternaria", **19**, pp.197-387.
- BIETTI A. (1978) - *The Mesolithic cultures in Italy: New activities in connection with upper Paleolithic cultural traditions*, Atti del II simposium internazionale "Mesolithikum in Europa", Postdam, (in corso di stampa).
- BIETTI A. (1978a) - *Alcune considerazioni sulla tipologia e sulle liste tipologiche per il Paleolitico superiore in Italia*, "Quaternaria", **20**, pp.1-27.
- BIETTI A. (1979) - *Metodi Matematici e Statistici applicati all'Archeologia e alla Paleontologia*, "Contributi del Centro Linceo Interdisciplinare di Scienze Matematiche e lo applicazioni", n°. **47**, Accademia Nazionale dei Lincei, Roma, 50 pp.
- BIETTI A. (1979a) - *Le gisement paléolithique supérieur de Taurisano (Lecce, Italie) et sa position chronologique et culturelle dans l'Epigravettien Italien*, "La fin des Temps Glaciaires en Europe", Colloquio Internazionale. CNRS, n°. **271**, Bordeaux 1977, pp.333-344.
- BIETTI A. (1980) - *The Mousterian Complexes in Italy: a Search for Cultural Understanding of the Industrial Assemblages*, Atti del 20. Congresso della "European Anthropological Association", Brno, Agosto 1980, (in corso di stampa).
- BIETTI A., RAMBALDI A., ZANELLO L. (1978) - *DENDR - Un programma FORTRAN di "Cluster analysis" per applicazioni archeologiche, paleontologiche e paleoecologiche*, "Quaternaria", **20**, pp.49-85.

- BIETTI A., ZANELLO L. (1975) - *CNTG- Un programma FORTRAN V per il calcolo di correlazioni tra attributi di oggetti archeologici con il metodo delle tabelle di contingenza*, Nota C.N.R. "S.S.A. 75.1".
- BIETTI A., ZANELLO L. (1980) - *A project of Pattern Recognition with an interactive system (ISPAHAN) in Prehistoric Archaeology: "Quantitative" typology of the End-Scrapers of an Italian upper Paleolithic industrial assemblage*, "Pattern Recognition in Practice" (ed. da E.S. Gelsema e L.N. Kanal), North Holland, Amsterdam, pp.517-526.
- BIETTI-SESTIERI A.M. (1979) - *Ricerca su una comunita' del Lazio Protostorico: Il sepolcreto dell'Osteria dell'Osa sulla Via Prenestina*, De Luca, Roma, 209 pp.
- BIETTI-SESTIERI A.M., BERGONZI G. (1979) - *La fase piu' antica della cultura laziale*, "Atti della XXI riun.scient. dell'I.I.P.P.", Firenze, ottobre 1977, pp.399-423.
- BINFORD L.R. (1963) - *The Pomranky site. A late archaic burial station*, "Miscellaneous Studies in Typology and Classification", Anthropol. Papers, n^o. 19, pp.149-192, Museum of Anthropology, University of Michigan.
- BINFORD L.R. (1972) - *Contemporary model building: paradigms and the current state of Paleolithic research*, "Models in Archaeology", (ed. da D.L. Clarke), Methuen, Londra, pp.109-166.
- BINFORD L.R. (1973) - *Interassemblage variability the Mousterian and the "functional" argument*, "The Explanation of Culture change: Models in Prehistory" (ed. da C. Renfrew), Duckworth, Londra, pp.227-254.
- BINFORD L.R., BINFORD S.R. (1966) - *A preliminary analysis of functional variability in the Mousterian of Levallois facies*, "Amer. Anthropologist", 68, 2, pp.238-295.
- BLALOCK H.M. (1960) - *Social Statistics*, McGraw-Hill, Londra e New York.
- BORDES F. (1953) - *Essai de classification des industries "mousteriennes"*, "Bull. Soc. Préh. Fr.", 90, pp.457-466.
- BORDES F. (1961) - *Typologie du Paléolithique ancien et moyen*, Delmas, Bordeaux, 85 pp.
- BORDES F. (1967) - *Considérations sur la typologie et les techniques dans le Paléolithique*, "Quartär", 18, pp.25-55.
- BORDES F., LAVILLE H., PAQUERAU M.M. (1966) - *Observations sur le Pléistocène supérieur du gisement de Combe-Grenal (Dordogne)*, "Act. de la Soc. Linn. de Bordeaux", 103B, pp.13-19.
- BORDES F., SONNEVILLE-BORDES D. de (1970) - *The significance of variability in Paleolithic assemblages*, "World Archaeology", 2, pp.61-73.

- BRAINERD G.W. (1951) - *The place of chronological ordering in archaeological analysis*, "Amer. Antiquity", **16**, pp.301-313.
- BROGLIO A. (1971) - *Risultati preliminari delle ricerche sui complessi Epipaleolitici della valle dell'Adige*, "Preist. Alpina", **7**, pp.135-241.
- BROGLIO A. (1975) - *Le passage du Paléolithique supérieur au Néolithique dans la région Vénétie-Trentin-Frioul*, "Epipaléolithique Méditerranéen", Atti del Convegno Internazionale Aix-en Provence, Giugno 1972, pp.5-21.
- BULGARELLI D. (1978) - *Elaborazione elettronica per la correlazione spaziale di dati ecologici e paleontologici*, Tesi di laurea, Istituto di Fisica dell'Università di Roma.
- BULGARELLI G.M. (1972) - *Il Paleolitico della Grotta di Torre Nave (Praia a Mare, Cosenza)*, "Quaternaria", **16**, pp.149-188.
- BULGARELLI G.M. (1974) - *Industrie Mustेरiane della Barma Grande ai Balzi Rossi di Grimaldi (Liguria)*, "Studi di Paleon., Paleoantr., Paleont. e Geol. del Quaternario", Mem.Ist.Ital.Paleont.Um., **II**, Roma, 91-129.
- CANEVA C., FRANGIPANE M., PALMIERI A.M. (1980) - *I metalli di Arslantepe nel quadro dei piu' antichi sviluppi della metallurgia vicino orientale*, Appendice di A. PALMIERI (1980) (in corso di stampa).
- CHADWICK A.J. (1978) - *A Computer simulation of Mycenaean settlement*, in: "Simulation Studies in Archaeology" (ed. da I.Hodder), Cambridge Univ. Press, pp.47-57.
- CLARK P.I., EVANS F.C. (1954) - *Distance to nearest neighbor as a measure of spatial relationship in populations*, "Ecology", **35**, pp.445-453.
- CLARKE L.D. (1968) - *Analytical Archaeology*, Methuen, London, 684 pp.
- CLARKE D.L. (1972) - *Models and paradigms in contemporary archaeology*, "Models in Archaeology" (ed. da D.L.Clarke), Methuen, London, pp.1-60.
- COOLEY W.W., LOHNES P.R. (1971) - *Multivariate Data Analysis*, J. Wiley, New York, 364 pp.
- CRAMER H. (1946) - *Mathematical methods of Statistics*, Princeton Univ.Press.
- DELPORTE H., MAZIERE G., DJINDJIAN F. (1977) - *L'Aurignacien de La Ferrassie. Observations préliminaires à la suite de fouilles récentes*, "Bull. Soc. Preh. Fr.", **74**, pp.343-361.
- DORAN J.E., HODSON F.R. (1975) - *Mathematics and Computers in Archaeology*, Edinburgh University Press, 381 pp.
- FISHER R.A. (1950) - *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh, 354 pp.
- FORD J.A. (1962) - *A quantitative method for deriving cultural chronology*. Technical Manual 1, Dept. of Soc. Affairs, Pan American Union, Washington, D.C.

- FRITZ J.M., PLOG F.T. (1970) - *The nature of archaeological explanation*, "Amer. Antiquity", **35**, pp.405-412.
- FUGAZZOLA-DELPINO A.M. (1976) - *Testimonianze di cultura appenninica nel Lazio*, "Origines", Sansoni, Firenze, pp.318.
- GABORI - CSÁNK V. (1968) - *La Station du Paléolithique Moyen d'Érd - Hongrie*, "Akadémiai Kiadó", Budapest, pp.277.
- GARDIN J.C. (1979) - *Une Archéologie théorique*, Hachette, Parigi, 339 pp.
- GELSEMA E.S. (1980) - *ISPAHAN; An Interactive System for Pattern Analysis: Structures and Capabilities*, "Pattern Recognition in Practice" (ed. da E.S. Gelsema e L.N. Kanal), North-Holland, Amsterdam, pp. 481-491.
- GRECO-PONTRANDOLFO A. (1977) - *Su alcune tombe Pestane: proposte d'una lettura*, "Mélanges de l'éc. Fr. de Rome" (in corso di stampa).
- GREIG-SMITH P. (1964) - *Quantitative plant ecology*, 2^a ed., Plenum Press, New York.
- HARMAN H.H. (1960) - *Modern factor analysis*, University of Chicago Press, Chicago.
- HIGGS E.S. (Ed.) (1972) - *Papers in Economic Prehistory*, Cambridge University Press, 219 pp.
- HODDER I., ORTON C. (1976) - *Spatial analysis in Archaeology*, Cambridge University Press, pp.85-97.
- IHM P. (1976) - *Chronologische Seriation: ein Problem Statistischer Schätzung*, "Atti del IX Congresso UISPP", Nizza, Settembre 1976, Coll. IV, pp.133-140.
- KENDALL D.G. (1963) - *A statistical approach to Flinders-Petrie's sequence-dating*, "Bull. I.S.I. 34th session", Ottawa, pp.657-680.
- KERRICH J.E., CLARKE D.L. (1967) - *Notes on the possible misuse and errors of cumulative percentage frequency graphs for the comparison of prehistoric artefact assemblages*, "Proc. Prehist. Soc.", **33**, pp. 57-69.
- KERSHAW K.A. (1961) - *Association and covariance of plant communities*, "Journal of Ecology", **49**, pp.643-654.
- KUHN T.S. (1970) - *The structure of scientific revolutions*, 2th Ed. "Intern. Encyclopedia of Unif. Science", University of Chicago Press. Traduzione italiana: "La struttura delle rivoluzioni scientifiche", Einaudi, Torino, 4^a ed., 1978.
- LAPLACE G. (1964) - *Essai de typologie systématique*, "Ann. Univ. di Ferrara", N.S., sez. XV, **1**, suppl. II, 85 pp.
- LAPLACE G. (1974) - *De la dynamique de l'analyse structurale ou la typologie analytique*, "Riv. Sci. Preist.", **29**, pp.1-71.

- LAVILLE H., RIGAUD J. Ph. (1973) - *The Perigordian V industries in Périgord: typological variations, stratigraphy and relative chronology*, "World Archaeology", **4**, pp.330-338.
- MCDONALD W.S., RAPP G.R. (1972) - *The Minnesota Messenia Expedition*, University of Minnesota Press.
- MELLARS P. (1970) - *Some comments on the notion of "function variability" in stone-tool assemblages*, "World Archaeology", **2**, pp.74-89.
- MEZZENA L., PALMA di CESNOLA A. (1967) - *L'Epigravettiano della Grotta Paglicci nel Gargano (Scavi F. Zorzi, 1962-63)*, "Riv. Sc. Preist.", **22**, pp.23-156.
- MOVIUS H.L., DAVID N.C., BRICKER H.M., CLAY R.B. (1968) - *The Analysis of Certain Major Classes of Upper Paleolithic Tools*, "Bull. of Am. Sc. Prehist. Res.", n° **26**, Peabody Museum, Harvard.
- MÜLLER-KARPE H. (1962) - *Zur Stadtwerdung Roms*, Kerle, Heidelberg, 108 pp.
- NEWELL R.R., DEKIN A.A. (1978) - *An Integrative Strategy for the Definition of behaviourally meaningful Archaeological Units*, "Palaeohistoria", **20**, pp.7-38.
- NEWELL R.R., VROOMANS A.P.J. (1972) - *Automatic Artifact Registration and Systems for Archaeological Analysis with the Philips P1100 Computer: A Mesolithic Test-case*, Anthropological Publications, Oosterhout, 103 pp.
- PALMA di CESNOLA A. (1975) - *Il Gravettiano della Grotta Paglicci nel Gargano. I: l'industria litica e la cronologia assoluta*, "Riv. Sci. Preist.", **30**, pp.3-177.
- PALMIERI A. (1980) - *Arslantepe, Malatya, Scavi 1976-1979*, (in corso di stampa).
- PERONI R. (1959) - *Per una nuova cronologia del sepolcreto arcaico del Foro - Sequenza culturale e significato storico*, "Civiltà del Ferro", Bologna, pp.461-499.
- PERONI R. (1966) - *Considerazioni ed ipotesi sul ripostiglio di Ardea*, "Bull. Paleon. Ital.", **75**, pp.175-196.
- PETRIE W.M.F. (1899) - *Sequences in prehistoric remains*, "J.Anthrop.Inst.", **29**, pp.295-301.
- PIELOU E.C. (1969) - *An Introduction to Mathematical Ecology*, Wiley, New York.
- PIPERNO M. (1974) - *L'Industria Musteriana su calcare di Grotta Romanelli (Otranto)*, "Studi di Paleon., Paleantr., Paleont. e Geol. del Quaternario", Mem. Ist. Ital. Paleont. Um., **II**, Roma, pp.69-90.
- PIPERNO M., BIDDITTO I. (1978) - *Studio tipologico ed interpretazione dell'industria acheulana e pre-musteriana dei livelli m e d di Torre in Pietra (Roma)*, "Quaternaria", **20**, pp.441-536.

- PRICE T.D. (1978) - *Regional approaches to human adaptation in the Mesolithic of the North European plain*, Atti del II Simp. Internaz. "Mesolithikum in Europa", Potsdam Aprile 1978, (in corso di stampa).
- RADMILLI A.M. (1974) - *Gli scavi nella grotta Polesini a Ponte Lucano di Tivoli e la piu' antica arte nel Lazio*, "Origines", Sansoni, Firenze, p.131.
- RENFREW C. (1972) - *The emergence of civilization: the Cyclades and the Aegean in the third millennium B.C.*, Methuen, Londra.
- RENFREW C. (1977) - *Alternative Models for exchange and spatial distribution*, "Exchange systems in Prehistory" (ed. da T.K. Earle e J.E. Ericson), Academic Press, New York, pp.71-90.
- RITCHIE W.A., MACNEISH R.S. (1949) - *The Pre-Iroquoian pottery of New York State*, "Amer. Antiquity", **15**, pp.97-124.
- ROBINSON W.S. (1951) - *A method for chronologically ordering archaeological deposits*, "Amer. Antiquity", **16**, pp.293-301.
- SACKETT J.R. (1966) - *Quantitative analysis of upper Paleolithic stone tools*, "Amer. Anthropologist", **68**, 2, pp.356-394.
- SIREIX M., BORDES F. (1972) - *Le Mousterien de Chinchon (Gironde)*, "Bull. Soc. Preh. Fr.", **69**, pp.324-336.
- SOKAL R.R., SNEATH P.H.A. (1963) - *Principles of numerical taxonomy*, Freeman, S.Francisco e Londra, 359 pp.
- SONNEVILLE-BORDES D. de, PERROT J. (1953) - *Essai d'adaptation des méthodes statistiques au Paléolithique supérieur. Premiers resultats*, "Bull. Soc. Preh. Fr.", **50**, pp.323-333.
- SONNEVILLE-BORDES D. de, PERROT J. (1954-1956) - *Lexique typologique du Paléolithique supérieur*, "Bull. Soc. Préh. Fr.", **51**, pp.237-355; **52**, pp.76-79; **53**, pp.408-412 e pp.547-559.
- SORDA S. (1975) - *I "ripostigli" di bronzi protostorici dell'Italia Centrale*, "Suppl. Annali dell'Ist. Ital. di Numismatica", **22**, pp.61-74.
- SPAULDING A.C. (1953) - *Statistical techniques for the discovery of artifact types*, "Amer. Antiquity", **18**, pp.305-313.
- SPIEGEL M.R. (1961) - *Theory and problems of Statistics*, Schaum Publishing Co., New York, 359 pp.
- STERUD G. (1973) - *A paradigmatic view of prehistory*, "The Explanation of Culture Change: Models in Prehistory" (ed. da C.Renfrew), Durckworth, Londra, pp.3-17.
- TASCHINI M. (1968) - *La datation au C¹⁴ de l'abri Blanc (Mont Circé). Quelques observations sur le Mésolithique en Italie*, "Quaternaria", **10**, pp.137-165.

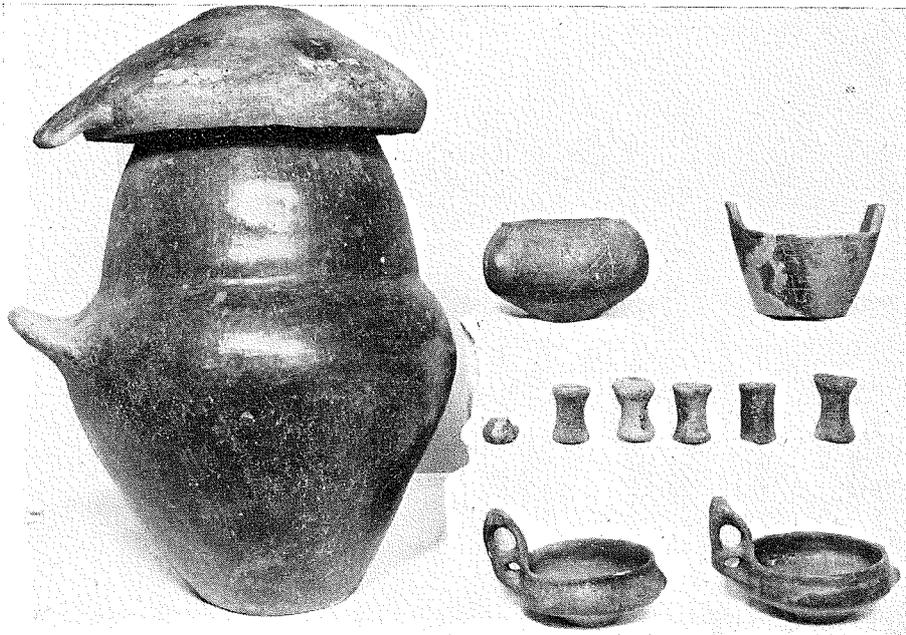
- TASCHINI M. (1979) - *L'industrie lithique de Grotta Guattari au Mont Circé (Latium): définition culturelle, typologique et chronologique du Pontinien*, "Quaternaria", **21**, pp.179-247.
- TIXIER J. (1963) - *Typologie de l'Épipaléolithique du Maghreb*, "Mém. du Centre de rech. anthr. preh. et ethnol.", **2**, Algeri. Parigi.
- TUGBY D.J. (1969) - *Archaeology and Statistics*, "Science in Archaeology", (ed. da D. Brothwell e E.S. Higgs), Thames and Hudson, 2^a ed., pp. 635-648.
- WHALLON R. (1972) - *A new approach to pottery typology*, "Amer. Antiquity", **37**, pp.13-33.
- WHALLON R. (1973) - *Spatial Analysis of Occupation Floors I: Application of Dimensional Analysis of Variance*, "Amer. Antiquity", **38**, pp.266-278.
- WHALLON R. (1974) - *Spatial Analysis of Occupation Floors II: The Application of Nearest Neighbor Analysis*, "Amer. Antiquity", **39**, pp.16-34.
- WOBST H.M. (1974) - *Boundary conditions for Paleolithic social systems: a simulation approach*, "Amer. Antiquity", **39**, pp.147-177.
- YOUNG H.D. (1962) - *Statistical Treatment of Experimental Data*, McGraw-Hill, New York, 172 pp.
- YULE G.U., KENDALL M.G. (1953) - *Introduction to the theory of Statistics*, Griffin and Co., Londra.

INDICE

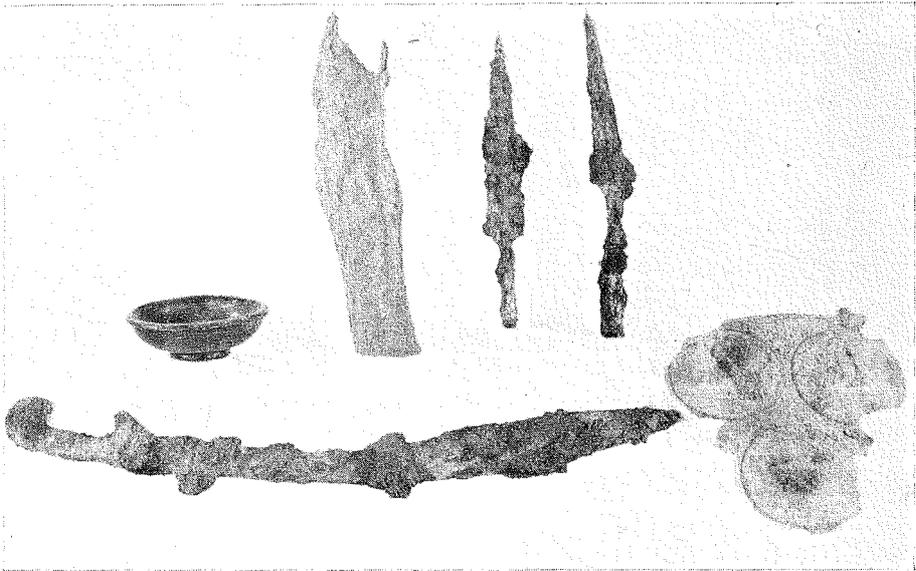
Introduzione.....	Pag.	3
I - Metodi basilari di classificazione dei dati archeologici, mediante statistiche elementari.....	"	7
II - Analisi e confronti tra complessi archeologici. Cenni a metodi di analisi multivariata...	"	34
III - Problemi di interpretazione globale: ricostruzione di "patterns" nello spazio e nel tempo. Cenni a modelli matematici di culture.....	"	110
Bibliografia.....	"	148



a: Osteria dell'Osa (Roma). Corredo ceramico della tomba TU 55.



b: Necropoli di Quattro Fontanili, Veio. Esempio di corredo ceramico.



a: Paestum (Salerno), Necropoli lucana. Armi e corazza della tomba 112.



b: Paestum (Salerno), Necropoli lucana. Corredo ceramico della tomba 11.

AVVERTENZA

Nella presente collezione dei « Contributi del Centro Linceo Interdisciplinare di Scienze Matematiche e loro Applicazioni », sono finora apparse le seguenti pubblicazioni:

1. AGENO M., *Punti di contatto tra fisica e biologia* (con una Prefazione di Beniamino Segre. Corso di dieci lezioni tenute dal 22 al 26 maggio 1972), 1974.
2. ROSSI B., *Astronomia in raggi X* (Lezioni tenute nel febbraio e marzo 1972, raccolte da Bianca Maria Belli), 1974.
3. TOUSCHEK B., *Sull'insegnamento della teoria dei quanti* (Lezioni tenute nell'aprile 1972), 1975.
4. DIRAC P. A. M., *The Development of quantum Mechanics* (Conferenza tenuta il 14 aprile 1972), 1974.
5. FERRARO V. C. A., *Il vento solare ed il campo magnetico interplanetario* (Conferenza tenuta il 17 aprile 1972), 1974.
6. *Seminari su: « La Scienza dei Sistemi »* (con una Prefazione di Beniamino Segre).
Parte Prima (I Seminario: 30 novembre-4 dicembre 1970; II Seminario: 11-15 gennaio 1971), 1975.
Parte Seconda (III Seminario: 8-12 marzo 1971; IV Seminario: 5-9 aprile 1971; V Seminario: 3-7 maggio 1971; VI Seminario: 24-28 maggio 1971), 1975.
7. *Seminario sulla: « Evoluzione Biologica »* (Roma, 10-11 gennaio, 17-19 aprile 1974), 1975.
8. NE'EMAN Y., *Patterns and Symmetry in the Structure of Matter* (Conferenza tenuta il 15 dicembre 1973), 1975.
9. SEIDENBERG A., *Constructions in Algebra* (Riassunto delle lezioni tenute nell'ottobre e novembre 1972), 1975.
10. *Tavola rotonda sul tema: « Problemi matematici ed economici odierni sulle assicurazioni »* (Roma, 24-25 novembre 1972), 1975.
11. CAMPA R., *La guerra e il processo di trasformazione tecnologica* (Conferenza tenuta il 26 maggio 1975), 1975.
12. MEDICI M., *Indirizzi verso motori automobilistici meno inquinanti* (Conferenze tenute nel marzo 1973), 1975.
13. *Colloquio sul tema: « Le tecniche di classificazione e loro applicazione linguistica »* (Firenze, 13 dicembre 1972), 1975.
14. GATTO R. R., *Interazioni elettromagnetiche, invarianza di scala e sue possibili estensioni* (Lezioni tenute nel settembre 1972), 1976.
15. *Seminario sulla: « Evoluzione Biologica »* (Roma, 19-22 febbraio 1975), 1976.
16. DE GIORGI E., *Convergenza in energia di operatori ellittici* (Conferenza tenuta nel febbraio 1974), 1976.
17. MOISIL G. C., *Sur l'emploi des Mathématiques dans les Sciences de l'homme* (Conferenza tenuta il 5 giugno 1972), 1976.
18. ANDREOTTI A., *Lewy Problem for Cauchy-Riemann Equations* (Lezioni tenute nel febbraio 1973), 1976.
19. ALFONSI D., BALLA M. I., DE SANTIS F., GIORGI G. e SCHAEFER M., *Struttura di un sistema informativo per un servizio di documentazione scientifica* (Da una manifestazione tenuta nel febbraio 1976 per iniziativa del Centro Linceo e dell'Università di Roma), 1976.
20. TRUESDELL C. A., *Termodinamica razionale* (Corso di lezioni tenute nel gennaio 1973), 1976.
21. TOGNOLI A., *Introduzione alla teoria degli spazi analitici reali*. (Lezioni tenute nel febbraio 1973, raccolte da D. Smit Ghinelli), 1976.
22. HANSON A., REGGE T. e TEITELBOIM C., *Constrained Hamiltonian Systems* (Ciclo di lezioni tenute dal 29 aprile al 7 maggio 1974), 1976.
23. CHESTNUT H., *Influence of Technology on Modern World Evolution and Use of Dynamic Models of Macro-Economic Systems in Development Planning* (Conferenza tenuta il 21 novembre 1972), 1976.
24. ANDREOTTI A., *Introduzione all'analisi complessa* (Lezioni tenute nel febbraio 1972), 1976.
25. REGGE T. e RASETTI M., *Vortices and Current Algebra* (Conferenze tenute nel giugno 1975), 1976.
26. SANSONE G., *Studi sulle equazioni differenziali ordinarie nell'ultimo cinquantennio* (Lezione tenuta il 12 dicembre 1975), 1976.
27. SEGRÉ E., *Personaggi e scoperte nella Fisica contemporanea* (Ciclo di lezioni tenute dal novembre 1972 fino al marzo 1973), 1976.
28. *Seminario sui: « Sistemi di reperimento e selezione automatica dell'informazione »* (Roma, 17-21 aprile 1972), 1976.

29. *Seminario sulle: « Applicazioni della Scienza dei Sistemi alla Medicina e alla Chirurgia »* (Roma, 22-26 maggio 1972), 1976.
30. *Convegno Internazionale sul tema: « Trends in the Physics and Engineering of Technological Materials »* (Roma, 17-19 ottobre 1973), 1976.
31. *Gruppo di Studio sui: « Fenomeni di alta energia nelle ultime fasi dell'evoluzione stellare »* (Roma-Frascati, 29 maggio-16 giugno 1972), 1976.
32. *III Seminario sulla: « Evoluzione Biologica: il codice genetico »* (Roma, 26-28 febbraio 1976), 1977.
33. *Seminario sul tema: « Una nuova via italiana alla fisica delle alte energie: Ada, Adone.. »* (Roma, 24-25 maggio 1974), 1977.
34. DIRAC G., *Cardinal-determining Subgraphs of infinite Graphs* (Lezione tenuta il 16 aprile 1975), 1977.
35. LEWY H., *On the Boundary Behavior of Holomorphic Mappings* (Lezione tenuta il 3 maggio 1976), 1977.
36. DUBOS R., *The resilience of ecosystems* (Lezione tenuta il 17 dicembre 1976), 1977.
37. *Seminario sul tema: « Rapporti tra Biologia e Statistica »* (Roma, 19-20 dicembre 1975), 1977.
38. BAER R., *Finite Metanilpotent Groups and Finite Sylow Tower Groups* (Corso di lezioni tenute nell'aprile 1976), 1977.
39. CESARI L., *Nonlinear Analysis and Alternative Methods* (Ciclo di lezioni tenute nell'aprile 1974), 1977.
40. *Convegno Internazionale: « Problemi connessi con l'utilizzazione pacifica dell'energia nucleare in Italia »* (Roma, 12-14 aprile 1976), 1977.
41. *IV Seminario sulla: « Evoluzione Biologica »* (Roma, 17-19 febbraio 1977), 1978.
42. ISTRATESCU V. I., *Topics in linear operator theory* (Corso di lezioni tenute nell'aprile 1976), 1978.
43. *Convegno sul tema: « Applicazioni del teorema del punto fisso all'analisi economica »* (Roma, 9-11 marzo 1977), 1978.
44. *Congresso Internazionale su: « L'insegnamento integrato delle Scienze nella scuola primaria »* (Roma, 7-15 gennaio 1976), 1979.
45. MARTINELLI E., *Introduzione alla teoria delle classi caratteristiche: uno sguardo panoramico* (Corso di lezioni tenute nel febbraio e marzo 1978. Redatte da Guido Lupacciolu e Paolo Piccinni), 1979.
46. ANGELINI A. M., *Linee di sviluppo nella utilizzazione della energia solare* (Conferenza tenuta il 9 marzo 1979), 1979.
47. BIETTI A., *Modelli matematici e statistici applicati all'Archeologia e alla Paleontologia* (Conferenza tenuta il 16 giugno 1978), 1979.
48. *V Seminario sulla: « Evoluzione Biologica: evoluzione della sessualità ed evoluzione umana »* (Roma, 23-25 febbraio 1978), 1979.
49. GOLINI A., *Attuali tendenze della popolazione in Italia: problemi e prospettive* (Conferenza tenuta il 9 febbraio 1979), 1979.
50. DE BENEDETTI S., *Dall'universo di Newton a quello di Einstein* (Conferenza tenuta il 28 maggio 1979), 1979.
51. *VI Seminario sulla: « Evoluzione Biologica: Ecologia ed Etologia »* (Roma, 22-24 febbraio 1979), 1980.
52. LAUGWITZ D., *The Theory of Infinitesimals. An Introduction to Nonstandard Analysis* (Ciclo di lezioni tenute nel marzo 1977), 1980.
53. *International Meeting on: « Astrophysics and Elementary Particles, Common Problems »* (Rome, 21st-23rd February 1980), 1980.
54. CARERI G., *Ordine e disordine nella materia. Tre lezioni sugli aspetti interdisciplinari* (Roma, 26, 28 e 30 novembre 1979), 1981.
55. ADKINS W. A., ANDREOTTI A., LEAHY J. V., *Weakly Normal Complex Spaces*, 1981.
56. SAPORETTI C., *Risultati e prospettive dell'analisi dei testi accadici mediante il calcolatore elettronico* (Conferenza tenuta il 14 dicembre 1979), 1981.
57. *VII Seminario sulla: « Evoluzione Biologica e i grandi problemi della Biologia »* (Roma, 28-29 febbraio-1 marzo 1980), 1981.
58. RANZI S., *L'embriologia: recenti studi a livello molecolare* (Conferenza tenuta il 7 marzo 1980), 1981.
59. *Convegno sul tema: « Problemi di popolazione: realtà attuali e prospettive »* (Roma, 13 giugno 1980), 1981.
60. *VIII Seminario sulla: « Evoluzione Biologica e i grandi problemi della Biologia. Aspetti biologici e sociali: parassitismo e simbiosi »* (Roma, 25-27 febbraio 1981), 1982.
61. BIETTI A., *Tecniche matematiche nell'analisi dei dati archeologici* (Ciclo di tre conferenze tenuto nel dicembre 1980), 1982.

