

Numero 7 - 1994

RATIO MATHEMATICA

Rivista di Matematica Applicata all'Economia e all'Ingegneria

a cura di

Franco Eugeni e Antonio Maturo

Comitato Scientifico

Albrecht Beutelispacher, *Giessen*

Franco Eugeni, *Roma*

Angelo Gilio, *Catania*

Mario Gionfriddo, *Catania*

Antonio Maturo, *Chieti*

Bruno Rizzi, *Roma*

Aniello Russo Spena, *L'Aquila*,

Romano Scozzafava, *Roma*

Numero 7 - 1994

RATIO MATHEMATICA

Rivista di Matematica Applicata all'Economia e all'Ingegneria

a cura di

Franco Eugeni e Antonio Maturo

Comitato Scientifico

Albrecht Beutelspacher, *Giessen*

Antonio Maturo, *Chieti*

Franco Eugeni, *Roma*

Bruno Rizzi, *Roma*

Angelo Gilio, *Catania*

Aniello Russo Spena, *L'Aquila*,

Mario Gionfriddo, *Catania*

Romano Scozzafava, *Roma*

Numero 7 - 1994

RATIO MATHEMATICA

Rivista di Matematica Applicata all'Economia e all'Ingegneria

a cura di

Franco Eugeni e Antonio Maturo

Comitato Scientifico

Albrecht Beutelispacher, Giessen

Franco Eugeni, Roma

Angelo Gilio, Catania

Mario Gionfriddo, Catania

Antonio Maturo, Chieti

Bruno Rizzi, Roma

Aniello Russo Spena, L'Aquila,

Romano Scozzafava, Roma

INDICE

PREFAZIONE	pag.	I
UN METODO DI RAGGRUPPAMENTO – I VICINI RECIPROCI Carlo Antonelli	»	1
CONSIDERAZIONI SULL'INFERENZA STATISTICA E LA VERIFICA DI IPOTESI Enzo Ballone - Vittorio Colagrande	»	13
LA DIMENSIONE CAMPIONARIA NELLA VERIFICA DI IPOTESI STATISTICHE Enzo Ballone - Vittorio Colagrande	»	25
DALLA COERENZA ALLA FORMULA DI BAYES: UN PERCORSO TEORICO FONDAMENTALE PER LE APPLICAZIONI Giuseppe Di Biase	»	39
PROBABILITÀ CONDIZIONATE E APPRENDIMENTO DALL'ESPERIENZA: UN'APPLICAZIONE DEL TEOREMA DI BAYES Angelo Gilio	»	57
A REAL LIFE APPROACH TO THE TEACHING OF PROBABILITY Romano Scozzafava	»	75
ALCUNI SPUNTI SULL'INSEGNAMENTO DELLA MATEMATICA CON PARTICOLARE RIFERIMENTO AD ASPETTI ELEMENTARI DELLA NOZIONE DI CONTINUITÀ IN ANALISI MATEMATICA Domenico Lenzi	»	83
NUMERI LATERALI E DIAGONALI Silvio Maracchia	»	91
MATEMATICA E FIGURINE Antonio Maturo	»	103

TEORIA DELLA DIVISIBILITÀ E ALGORITMO EUCLIDEO NELL'INSEGNAMENTO NELLE SCUOLE SUPERIORI Antonio Maturo - Fabiola Paccapelo - Loredana Renzullo	pag.	111
CALCOLO DI AREE E VOLUMI PER MEZZO DI SUCCESSIONI DI NUMERI A CASO Giuseppina Varone	»	125
LINGUAGGIO NATURALE, LINGUAGGIO MATEMATICO Miryam Benvenuto	»	135
MODELLI GEOMETRICI PER Z E Q Raffaele Bruno	»	143
STUDIO DI UNA POLIZZA ALLO SPORTELLLO Filippo Di Leonardo - Giuseppina Pagliuca	»	155
I PROCESSI COGNITIVI DELL'ARITMETICA IN RUSSEL E PIAGET Silvana D'Andrea	»	169
SISTEMI DINAMICI CAOTICI: UN ESEMPIO IN CAMPO NEUROBIOLOGICO Serena Doria	»	175
IL CODICE DI LEON BATTISTA ALBERTI Franco Eugeni - Diana Eugeni	»	179

PREFAZIONE

Il presente volume contiene i testi di alcune selezionate conferenze e comunicazioni tenute in convegni dedicati alla ricerca in Didattica della Matematica quali il 1° Convegno Interregionale su Probabilità, Statistica e Crittografia nell'insegnamento nelle scuole medie superiori, tenutasi ad Isernia nel 1992, i Convegni Nazionali Annuali della Mathesis, Società di Scienze Fisiche e Matematiche del 1992 e del 1993, per i quali, non sono stati pubblicati gli Atti ed il Congresso Internazionale ICME - 7 tenutosi a Québec nel 1992. Inoltre contiene alcuni articoli inviati alla rivista e ritenuti degni di pubblicazioni dal comitato scientifico.

Abbiamo ritenuto importante trovare una maniera per rendere noti alcuni degli interessanti contributi dati dagli insegnanti di Matematica in questi convegni e a tale scopo abbiamo trovato un sostegno da parte dell'Amministrazione dell'Istituto Mecenate di Pescara.

È nostra convinzione, condivisa dai nostri colleghi dell'Università che un Istituto Superiore parificato si possa qualificare ad un livello superiore a varie scuole statali se impegna una parte delle sue energie alla Ricerca Didattica e riesce, in tale settore, a stabilire una collaborazione permanente con l'Università.

L'intera organizzazione dell'Istituto Mecenate, ha mostrato la massima sensibilità verso tali esigenze culturali, assumendo parzialmente l'onere del finanziamento e dell'organizzazione di questo volume.

Esso certamente non sarà un momento isolato di collaborazine fra l'Università e l'Istituto Mecenate ma piuttosto rappresenta un punto di partenza ed un esempio da imitare.

Contributi determinanti alla realizzazione del volume sono dovuti al Dott. Giuseppe Di Biase, ricercatore di Istituzioni di Matematiche presso l'Università, che si è assunto l'impegno di una notevole mole di lavoro organizzativo e alla Prof.ssa Giuseppina Varone, docente di Matematica Applicata presso l'Istituto Mecenate, che ha rappresentato l'Istituto dal punto di vista scientifico.

Ringraziamenti particolari per la realizzazione di quest'opera vanno anche al Prof. Bruno Rizzi, ordinario di Analisi Matematica presso la Terza Università di Roma e Presidente Nazionale della Mathesis fino a tutto il 1993, al Prof. Romano Scozzafava, ordinario di Calcolo delle Probabilità e Statistica presso l'Università La Sapienza di Roma ed al Prof. Camillo Ciarlante, professore presso il Liceo Scientifico di Isernia ed organizzatore del Convegno di Isernia.

Franco Eugeni
professore di Geometria
Terza Università di Roma

Antonio Maturo
professore di Istituzioni di Matematiche
Università "G. D'Annunzio" di Chieti

UNITÀ DIDATTICA: UN METODO DI RAGGRUPPAMENTO – I VICINI RECIPROCI –

Carlo Antonelli

Istituto Tecnico Commerciale "R. De Sterlich" - Chieti Scalo

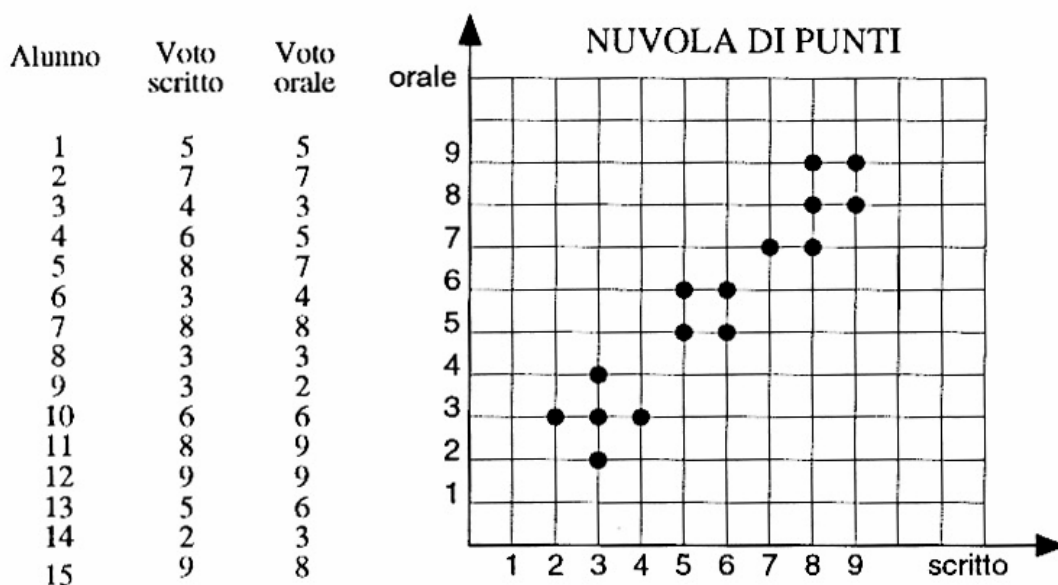
1. DALLA PROBLEMATICA ALLA METODOLOGIA DELL'ANALISI DEI GRUPPI.

In una classe si vogliono individuare tre gruppi di alunni, il più possibile omogenei, rispetto al grado di preparazione in una certa materia.

Si rilevano, per ogni ragazzo, i voti riportati nello scritto e nell'orale in questa materia.

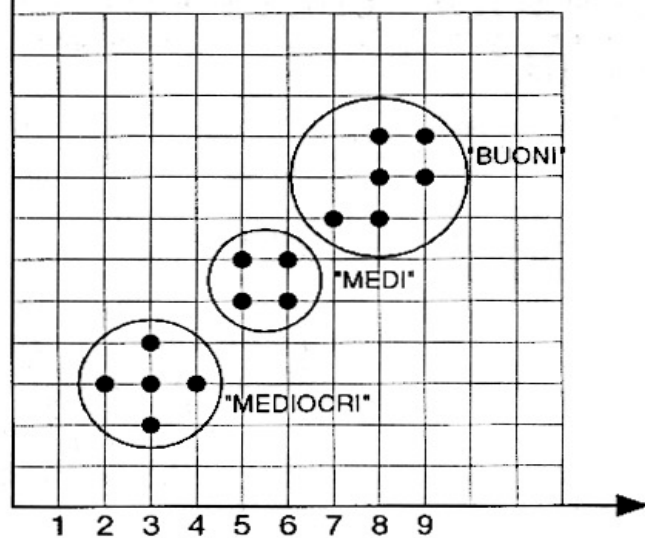
Si riportano le votazioni in un riferimento cartesiano, cosicché ogni alunno sarà individuato da un punto le cui coordinate sono rappresentate dai voti di quel ragazzo nello scritto (x) e nell'orale (y).

Nell'ipotesi in cui i voti rilevati sono i seguenti, avremo la nuvola di punti:



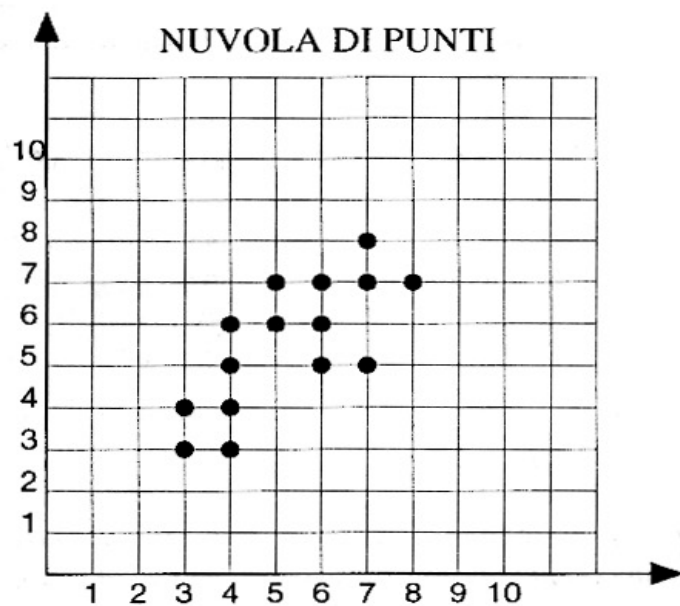
In questa situazione la sola analisi visiva della nuvola dei punti permette agevolmente e in maniera univoca, l'individuazione dei tre gruppi di alunni: gruppo dei "buoni", gruppo dei "medi", e gruppo dei "mediocri".

NUVOLA DI PUNTI CON GRUPPI



Ma situazioni del genere, in cui una popolazione risulta, rispetto ad una o più possibili variabili, "naturalmente" ben suddivisa in gruppi, sono abbastanza rare.

Nelle applicazioni pratiche è molto più facile incontrare situazioni, come quella riportata di seguito, in cui la formazione dei gruppi, per mezzo della sola analisi visiva, non è così facile ed univoca come nel caso precedente ma è lasciata all'arbitrarietà dell'individuo.



Si avverte l'esigenza, quindi, di usare delle metodologie statistiche che permettano, fissati alcuni presupposti che riguardano il tipo di distanza fra unità e la tecnica di raggruppamento, in maniera oggettiva e quindi ripetibile, l'individuazione dei gruppi.

Tali metodologie statistiche di classificazione che servono per raggruppare delle unità vanno sotto il nome di **ANALISI DEI GRUPPI** o **CLUSTER ANALYSIS**.

2. INTRODUZIONE

Lo sviluppo dell'Analisi dei Gruppi ha avuto all'inizio degli anni 60 un grosso impulso grazie soprattutto a ricercatori di scienze naturali quali biologi e zoologi e da antropologi ed archeologi i quali avevano tutti l'esigenza di mettere a punto metodi matematici che permettessero di raggruppare o classificare insiemi di unità.

Queste tecniche di raggruppamento hanno subito trovato un largo campo di applicazione in tutte quelle situazioni dove sorge l'esigenza di raggruppare popolazioni in gruppi (clusters) di unità il più possibile omogenee come per esempio nelle scienze sociali (**individuazione di classi sociali e comportamentali**), nelle ricerche di mercato (**individuazione di stili di vita, segmentazione del mercato**) e nella teoria dei campioni (**individuazione degli strati nei campioni stratificati**).

L'enorme potenziamento dei calcolatori in termini di velocità di calcolo e capacità di memoria ha permesso l'effettivo uso di tali procedure in situazioni reali cioè con popolazioni di migliaia di unità e decine di variabili.

3. TIPOLOGIE DEI METODI DI RAGGRUPPAMENTO

I metodi di classificazione si suddividono in due gruppi: – **metodi scissori** mediante i quali la formazione dei gruppi avviene suddividendo la popolazione di partenza in classi sempre più numerose;

– **metodi aggregativi** mediante i quali la formazione dei gruppi avviene aggregando per passi successivi le singole unità di partenza in gruppi di numerosità crescente.

Si parte considerando ogni unità statistica gruppo a se e per aggregazioni successive, intorno a queste unità iniziali, si formano passo dopo passo numeri decrescenti di gruppi fino a riaggregare le unità statistiche in un solo gruppo che rappresenta la popolazione di partenza.

Il metodo scelto per questa unità didattica è l'algoritmo dei "**VICINI RECIPROCI**" metodo aggregativo che per la sua "velocità" di aggregazione rispetto agli altri algoritmi viene molto usato nelle ricerche sociali dove si ha a che fare con popolazioni molto numerose.

Il principio del metodo dei Vicini Reciproci risale a Mc Quitty (1966), la sua messa a punto è dovuta a C. de Rham (1980) e la sua prima sperimentazione operativa a Benzecri (1982).

Il metodo è stato scelto per questa unità didattica grazie alle seguenti sue caratteristiche:

- è agevole e veloce nell'uso;
- richiede allo studente come prerequisiti, un limitato numero di conoscenze e abilità;
- fa uso anche di strumenti e concetti propri di altre discipline (coordinate, punti, distanze);
- è esaustivo sia per quanto riguarda i concetti base dell'analisi dei gruppi sia per quelli di una completa indagine statistica;
- dà la possibilità di "complicare" la procedura introducendo concetti e strumenti più "sofisticati" (standardizzazione delle variabili, particolari distanze), o di inserirlo in un contesto più ampio (teoria dei campioni), in modo da poterlo trattare proficuamente a vari livelli ed in diversi momenti di un corso di statistica.

4. ALGORITMO DEI "VICINI RECIPROCI"

Esempio di applicazione dell'algoritmo aggregativo dei "Vicini Reciproci" su una popolazione di 6 unità statistiche caratterizzate da due variabili quantitative.

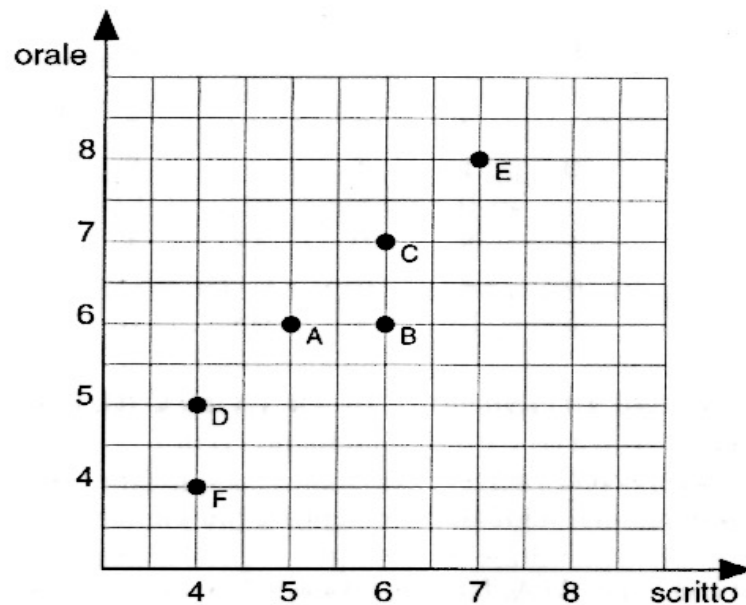
Obiettivo: individuare in un collettivo di 6 alunni tre gruppi di ragazzi omogenei rispetto al grado di preparazione in matematica. Come variabili esplicative del grado di preparazione vengono prese le votazioni riportate dai 6 alunni nello scritto e nell'orale di matematica.

1ª fase: rilevazione dei dati.

Insieme delle osservazioni $I = \{A, B, C, D, E, F\}$

Alunno	Voto scritto	Voto orale
A	5	6
B	6	6
C	6	7
D	4	5
E	7	8
F	4	4

2ª fase: rappresentazione della nuvola di punti



3ª fase: calcolo delle $(6 \times 5) / 2 = 15$ distanze $d(i, j)$. Come distanza è stata scelta il quadrato della distanza euclidea $d = [(x_i - x_j)^2 + (y_i - y_j)^2]$ e costruzione della matrice delle distanze.

$$d(A, B) = [(5 - 6)^2 + (6 - 6)^2] = 1$$

$$d(A, C) = [(5 - 6)^2 + (6 - 7)^2] = 2$$

$$d(A, D) = [(5 - 4)^2 + (6 - 5)^2] = 2$$

$$d(A, E) = [(5 - 7)^2 + (6 - 8)^2] = 8$$

$$d(A, F) = [(5 - 4)^2 + (6 - 4)^2] = 5$$

$$d(B, C) = [(6 - 6)^2 + (6 - 7)^2] = 1$$

$$d(B, D) = [(6 - 4)^2 + (6 - 5)^2] = 5$$

$$d(B, E) = [(6 - 7)^2 + (6 - 8)^2] = 5$$

$$d(B, F) = [(6 - 4)^2 + (6 - 4)^2] = 8$$

$$d(C, D) = [(6 - 4)^2 + (7 - 5)^2] = 8$$

$$d(C, E) = [(6 - 7)^2 + (7 - 8)^2] = 2$$

$$d(C, F) = [(6 - 4)^2 + (7 - 4)^2] = 13$$

$$d(D, E) = [(4 - 7)^2 + (5 - 8)^2] = 18$$

$$d(D, F) = [(4 - 4)^2 + (5 - 4)^2] = 1$$

$$d(E, F) = [(7 - 4)^2 + (8 - 4)^2] = 25$$

MATRICE DELLE DISTANZE

	A	B	C	D	E	F
A	0	1	2	2	8	5
B	1	0	1	5	5	8
C	2	1	0	8	2	13
D	2	5	8	0	18	1
E	8	5	2	18	0	25
F	5	8	13	1	25	0

4ª fase: ricerca per tutti gli elementi "i" di $I = \{A, B, C, D, E, F\}$ della minima distanza $\min d(i)$ fra (i) e gli altri elementi (j) di I.

$$\min d(i) = \inf [d(i, j) / j \in I, j \neq i].$$

Ricerca per ogni i del v (i), cioè dell'elemento il più vicino ad i.

$$\min d(A) = 1 = d(A, B) \Rightarrow v(A) = B$$

$$\min d(B) = 1 = d(B, A), d(B, C) \Rightarrow v(B) = A \text{ e } C$$

$$\min d(C) = 1 = d(C, B) \Rightarrow v(C) = B$$

$$\min d(D) = 1 = d(D, F) \Rightarrow v(D) = F$$

$$\min d(E) = 2 = d(E, C) \Rightarrow v(E) = C$$

$$\min d(F) = 1 = d(F, D) \Rightarrow v(F) = D$$

Schematizzando avremo:

i:	A	B	C	D	E	F
$\min d(i)$:	1	1	1	1	2	1
v (i):	B	A,C	B	F	C	D

5ª fase: individuazione delle coppie dei Vicini Reciproci che andranno a formare l'insieme M.

Definizione: due elementi i, j si definiscono Vicini Reciproci se $\min d(i) = \min d(j)$, e se ogni elemento della coppia è l'elemento più vicino dell'altro elemento.

Deve verificarsi che $i = v(j)$ e $j = v(i)$.

Così gli elementi A e B sono Vicini Reciproci perché $A = v(B)$ e $B = v(A)$, A è l'elemento il più vicino a B e B è l'elemento il più vicino ad A.

Le coppie di Vicini Reciproci saranno quindi:

(A e B) in quanto $v(A) = B$ e $v(B) = A$;

(D e F) in quanto $v(D) = F$ e $v(F) = D$.

Gli elementi C ed E non Formano una coppia di V. R. in quanto:

$$v(C) = B \text{ e } v(E) = C.$$

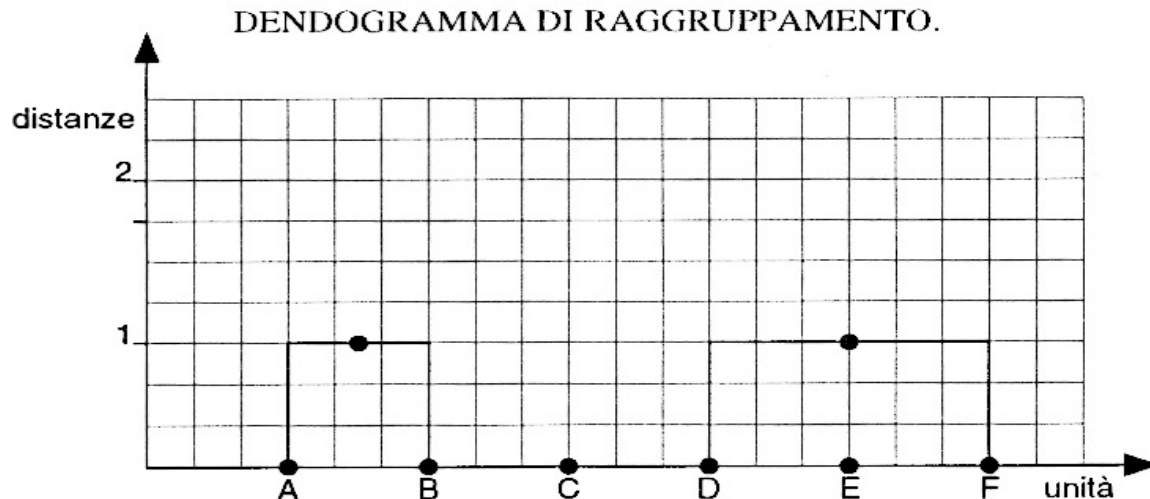
L'insieme M delle coppie dei V. R. sarà $M = \{(A, B), (D, F)\}$.

La prima iterazione è terminata ed ha prodotto i seguenti 4 gruppi:

$$I = \{(A, B), C, (D, F), E\}.$$

Si può fornire una rappresentazione grafica (DENDOGRAMMA) di questa procedura che poi viene aggiornata ad ogni nuova iterazione.

Sull'asse delle ascisse vengono riportate le unità statistiche, sull'asse delle ordinate le distanze a cui avvengono le aggregazioni.



Al livello di distanza 1 l'algoritmo aggrega le unità A e B e le unità D e F in quanto coppie di Vicini Reciproci. I nodi a livello 1 indicano la formazione dei gruppi.

Tornando alla nuvola di punti la prima iterazione ha prodotto 4 gruppi.

II ITERAZIONE

A questo punto si itera la procedura con la novità che le distanze da calcolare non sono più tra elementi ma tra gruppi ed elementi e tra gruppo e gruppo.

Come distanza tra gruppo ed elemento prendiamo la distanza media. La distanza tra l'unità C e la classe (A, B) è data dalla media aritmetica delle distanze tra C e le singole unità delle classe (A, B):

$d[(C, (A, B))] = [d(C, A) + d(C, B)] / 2$. Per quanto riguarda la distanza fra le classi prendiamo la distanza media interclasse, cosicché la distanza fra i gruppi (A, B) e (D, F) sarà data da:

$$d[(A, B), (D, F)] = [d(A, D) + d(A, F) + d(B, D) + d(B, F)] / 4.$$

Calcoliamo le nuove distanze:

$$d[(A, B), C] = [d(A, C) + d(B, C)] / 2 = (2 + 1) / 2 = 1.5$$

$$d[(A, B), E] = [d(A, E) + d(B, E)] / 2 = (8, 5) / 2 = 6.5$$

$$d[(A, B), (D, F)] = [d(A, D) + d(A, F) + d(B, D) + d(B, F)] / 4 \\ = (2 + 5 + 5 + 8) / 4 = 5$$

$$d[(C, (D, F))] = [d(C, D) + d(C, F)] / 2 = (8 + 13) / 2 = 10.5$$

$$d[(D, F), E] = [d(D, E) + d(F, E)] / 2 = (18 + 25) / 2 = 21.5$$

e la nuova matrice delle distanze

	(A, B)	C	(D, F)	E
(A, B)	0	1.5	5	6.5
C	1.5	0	10.5	2
(D, F)	5	10.5	0	21.5
E	6.5	2	21.5	0

Cerchiamo per tutti gli elementi di "i" di $I = \{(A, B), C, (D, F), E\}$ la minima distanza $\min d(i)$ fra (i) e gli altri elementi (j) di I. Ricerchiamo per ogni i il $v(i)$ cioè l'elemento il più vicino ad i.

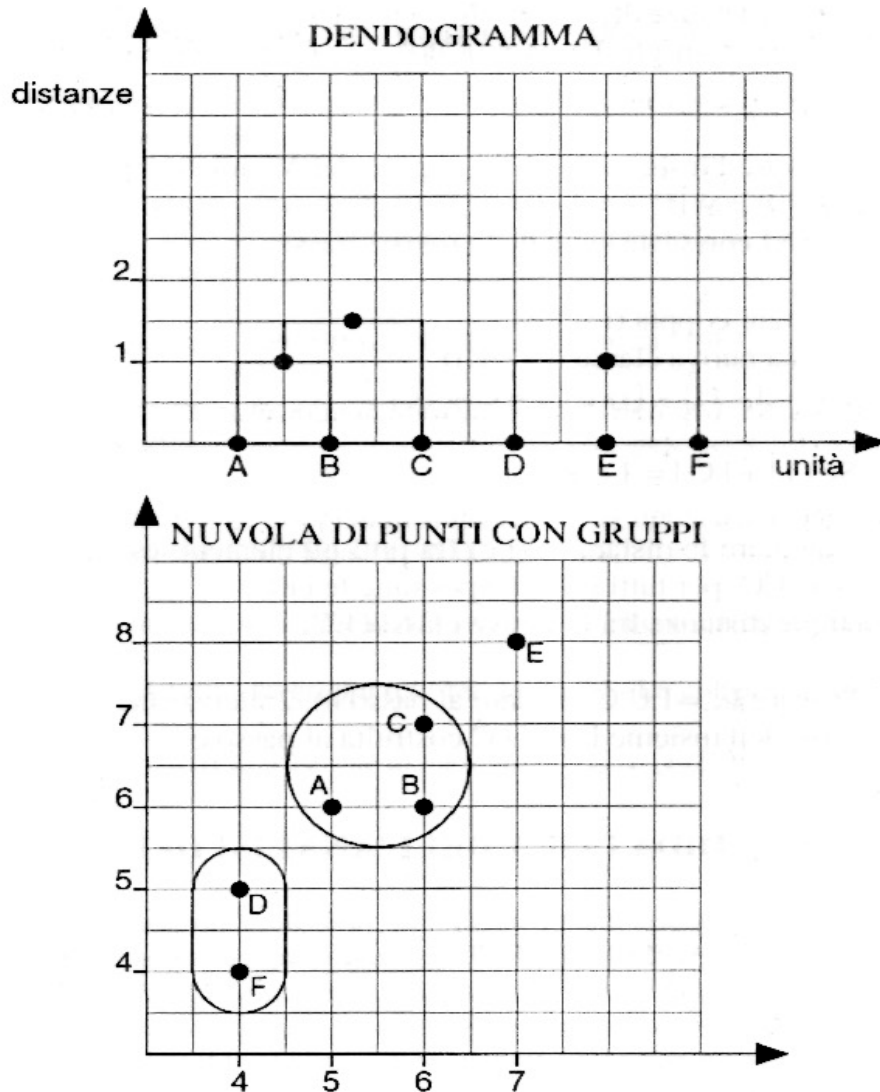
Schematizzando avremo:

i:	(A, B)	C	(D, F)	E
$\min d(i):$	1.5	1.5	5	2
$v(i):$	C	(A, B)	(A, B)	C

Quindi gli elementi (A, B) e C sono Vicini Reciproci in quanto $v(A, B) = C$ e $v(C) = (A, B)$.

La seconda iterazione è terminata ed ha prodotto i seguenti tre gruppi: (A, B, C), (D, F), (E).

Possiamo aggiornare il nostro dendrogramma che ci descrive tutte le fasi della formazione dei gruppi.



Poiché il nostro obiettivo era quello di suddividere la popolazione di partenza in tre gruppi interrompiamo la procedura che, se reiterata, porterebbe alla costituzione di un unico gruppo, la popolazione di partenza.

La scelta del numero ottimale di gruppi, se non viene stabilita a priori, come nel nostro caso, viene di solito effettuata analizzando il dendrogramma o con l'ausilio di appropriati indici.

È possibile formalizzare in "passi" l'Algoritmo dei Vicini Reciproci sia per una trattazione e presentazione di tipo formale sia per tradurre l'algoritmo in procedura informatica.

Passo a) Costruzione della matrice delle distanze.

- 1) Calcolare le distanze $d(i, j)$ per tutti $i, j \in I, i \neq j$.
- 2) Ricercare, per tutti gli $i \in I$, la $\min d(i)$ e $v(i)$ (l'elemento il più vicino ad i).

Passo b) Trovare l'insieme M delle coppie dei Vicini Reciproci $M = \{(i, j) \mid i, j \in I, i \neq j, i = v(j), j = v(i)\}$.

Definire $C := \emptyset$ con C insieme delle nuove classi.

Passo c) Per ogni coppia $(i, j) \in M$

- 1) Formare una nuova classe $c = (i \cup j)$.
- 2) Trovare $I := I - \{i\} - \{j\}$ e $C := C + \{c\}$.

Passo d) Se $|I| + |C| = 1$: stop.

Passo e) Calcolare le distanze $d(i, c)$ fra tutti gli elementi esistenti $i \in I$ e le nuove classi $c \in C$, per tutti gli $i \in I$ e per tutte le $c \in C$.

- 2) Calcolare le distanze fra le nuove classi $c \in C$.

Passo f) Trovare $I := I \cup C$. Andare al passo b). 2^a iterazione sulla matrice delle distanze dell'insieme $I = I \cup C$ costruita al passo e).

5. VALENZA DIDATTICA DELL'ANALISI DEI GRUPPI

Riguardo alle valenze didattiche di tale procedura possiamo dire che: la metodologia dell'ANALISI DEI GRUPPI è una indagine statistica completa in quanto riassume in se tutte le fasi di una ricerca statistica, dalla raccolta dei dati alla tabulazione, dalla rappresentazione grafica alla elaborazione dei dati e all'analisi dei risultati. Pertanto essa può essere trattata proficuamente al termine di un corso di statistica descrittiva.

Si può calibrare il grado di difficoltà, rapportandolo al livello del corso. Infatti partendo da un livello minimale in cui sono richiesti all'alunno pochi prerequisiti, si possono introdurre nella procedura di raggruppamento delle "migliorie" come la standardizzazione delle variabili, l'uso di altri tipi di distanza legate anche alla tipologia delle variabili (centro di gravità, varianza interclasse) o indici per la scelta della partizione migliore.

BIBLIOGRAFIA

- [1] A. Bellacicco *Metodologia e tecnica della classificazione matematica* - La goliardica- Roma (1975)
- [2] J. P. Benzecri *Le Cahiers de l'analyse des Donnée* - Vol. VII (1982).
- [3] S. Zani *Seminario: Due temi di analisi statistica multivariata* - Cleup Bressanone (1978)
- [4] S. ZANI *L'analisi classificatoria - Contributi metodo-logici ed impiego per l'individuazione di aree omogenee*. Istituto di statistica, Università degli studi di Parma (1977)
- [5] G. LETI *Statistica descrittiva* – Il Mulino Bologna (1983)
- [6] M. JAMBU *Classification automatique pour l'anlyse des données*. Dunod Paris (1878)
- [7] S. BOLASCO - R. Coppi *Analisi dei dati e sue applicazioni nella ricerca economica* - ed. Klim Roma (1983)
- [8] AA.VV. *Introduzione allo SPAD* – Centro di calcolo interfacoltà (1984)
- [9] A. RIZZI *Analisi dei gruppi (cluster analysis)* - La goliardica (1981)
- [10] G. MARBACH *Le ricerche di mercato* – UTET (1982)

CONSIDERAZIONI SULL'INFERENZA STATISTICA E LA VERIFICA DI IPOTESI

Enzo Ballone e Vittorio Colagrande***

**Facoltà di Medicina e Chirurgia, Università "G.d'Annunzio" - Chieti*

***Liceo Scientifico "G. Galilei" - Lanciano (CH)*

1. PREMESSA

Nella società odierna la quasi totalità delle indagini statistiche viene effettuata su campioni, diventa allora opportuno introdurre nell'insegnamento secondario superiore le principali tecniche di campionamento e le metodologie statistico-probabilistiche che permettono l'inferenza, ossia la generalizzazione dei risultati ottenuti attraverso una rilevazione per campione a dei collettivi più ampi detti "popolazione".

D'altra parte non si può trascurare la valenza formativa del processo induttivo. La Logica, in senso stretto, offre alle scienze sperimentali strumenti per respingere ("falsificare") una data ipotesi o una teoria quando queste sono in contrasto con un risultato sperimentale. Tuttavia la sola falsificazione è restrittiva per le esigenze delle diverse scienze: è necessario trovare altri strumenti che consentano di "imparare dall'esperienza" con maggiore ampiezza. Si tratta di risalire dalla verifica di un enunciato in uno o più casi particolari alla validità dell'enunciato stesso. È questa l'"induzione" ed è alla base della "logica dell'incerto" e, in particolare, dell'inferenza statistica.

I problemi dell'inferenza statistica sono sostanzialmente due: la stima statistica dei parametri di una popolazione e la verifica delle ipotesi statistiche concernenti la validità o meno di date assunzioni.

Nel presente lavoro, che trae origine dalla relazione tenuta dagli autori nell'ambito del Convegno Interregionale della MATHESIS su "Calcolo delle Probabilità Statistica e Crittografia nell'insegnamento", di Isernia 2-4 Aprile 1992, si esporranno alcune problematiche attinenti la verifica di un'ipotesi sul valore di una frequenza o sulla differenza di due frequenze.

La trattazione vuole essere una proposta didattica che serva di “stimolo” all’approfondimento e di invito alla presentazione “in classe” di alcuni procedimenti di inferenza statistica.

2. IL CAMPIONAMENTO E L’INFERENZA STATISTICA

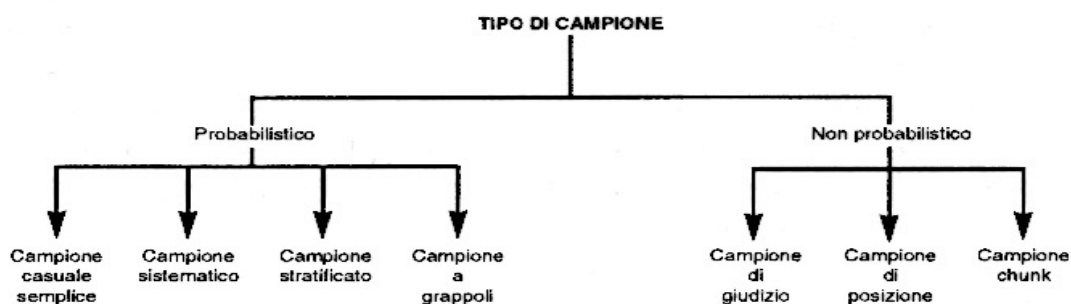
Ogni indagine statistica è mirata a studiare (prendere decisioni su) un carattere X di una ben definita popolazione attraverso alcuni parametri caratteristici (media, varianza, proporzione, ecc.), una parte dei quali o tutti frequentemente sono incogniti.

Supponendo che X sia una variabile casuale (v.c.) e che si distribuisca, nell’ambito della popolazione, secondo una funzione di densità (distribuzione) di probabilità $f(x|\Theta)$ dipendente da un vettore di parametri $\Theta \in R^n$, l’indagine ha come obiettivo quello di stimare o prendere decisioni sul vettore Θ . Nel caso che X sia una v.c. continua con distribuzione normale, $N_{m,s}(x) = 1 / [(2\pi)^{1/2}s] \exp[- (x - m)^2 / (2s^2)]$, il vettore dei parametri ha per componenti il valore medio m e lo scarto quadratico medio s .

L’analisi statistica spesso viene condotta su di un gruppo più o meno ristretto di persone od oggetti appartenenti alla popolazione stessa. Tale gruppo costituisce un campione ed è individuato in modo da consentire la generalizzazione dei risultati campionari all’intera popolazione.

Uno schema dei diversi tipi di campionamento è presentato nella fig. 1 seguente:

Fig. 1



La scelta degli elementi che formeranno un campione può essere effettuata in modo che ogni unità della popolazione abbia la stessa probabilità di entrare a far parte del campione. Questo tipo di estrazione, detto **campionamento casuale semplice**, fa parte delle tecniche randomizzate (probabilistiche) e si può provare che, utilizzando queste, il divario tra le informazioni fornite dal campione e quelle della popolazione è dovuto solo al caso.

Le operazioni di scelta degli elementi, poi, possono essere condotte con ripetizione (campionamento bernoulliano) o senza ripetizione (campionamento esaustivo).

Se ogni unità elementare della popolazione viene identificata con un numero intero progressivo a partire da 1, la scelta di ogni singola unità del campione porta alla definizione della variabile $X_i = \text{"valore di } X \text{ per l'i-ma unità"}$ e, quindi, la n-pla $(X_1, X_2, \dots, X_n)$ costituisce il campione. Dunque ogni campione di dimensione n di X corrisponde a n misure ripetute di X fatte nelle stesse condizioni. Nel caso di campionamento bernoulliano le variabili casuali X_i sono indipendenti.

Dopo aver estratto dalla popolazione un campione che la rappresenti, sui dati campionari si calcolano le statistiche corrispondenti ai parametri della popolazione (media campionaria, deviazione standard, proporzione campionaria, ecc.). Attraverso "la teoria della stima" e "i test statistici" è possibile, con diversi livelli di probabilità di errore, "inferire" dai risultati campionari a quelli dell'intera popolazione, vero oggetto di ricerca.

I motivi che spingono a condurre analisi su campioni vanno ricondotti principalmente a problemi organizzativi, di costi e di tempi per l'acquisizione ed elaborazione dei dati. D'altra parte le indagini campionarie permettono un'analisi più particolareggiata di quanto consentirebbe uno studio di tutte le unità della popolazione oggetto di interesse.

Bisogna ovviamente analizzare se e quando è possibile estendere le conclusioni campionarie a tutto il collettivo. Il campione dovrebbe riflettere le caratteristiche della popolazione ("riprodurre in miniatura" l'intera popolazione) oggetto di studio, ma la mancata conoscenza di tali caratteristiche non consente di giudicare

della rappresentatività del campione stesso. Si deve allora ripiegare su un campione che almeno non sia “distorto” e ciò è possibile attraverso la scelta di una tecnica di campionamento appropriata e della dimensione minima ottimale del campione. Stabilita la tecnica di campionamento più appropriata e fissata la dimensione, i risultati dell’indagine campionaria possono essere estesi poi all’intera popolazione, sia pure con un margine di errore prefissato (espresso in termini probabilistici).

Difficilmente, in realtà, si ottengono campioni che riproducano esattamente le caratteristiche della popolazione, tuttavia **nell’Universo dei campioni** (collezione di tutti i possibili campioni di data dimensione che si potrebbero estrarre da una popolazione) ci saranno certamente campioni maggiormente rappresentativi. In generale può affermarsi che quanto più il campione è numeroso, tanto più sarà rappresentativo della popolazione di provenienza; d’altra parte, però, al crescere della dimensione del campione crescono i costi e le difficoltà tecniche di organizzazione dell’indagine.

3. CARATTERISTICHE DELLA VERIFICA DI IPOTESI STATISTICHE

La verifica di ipotesi in una ricerca statistica è compiuta al fine di confrontare un’ipotesi sulla “composizione” relativa ad un carattere X , in una popolazione oggetto di studio, con i risultati sperimentali che scaturiscono dall’osservazione di un campione estratto dalla popolazione stessa.

In generale un’indagine di tipo statistico parte con la formulazione di una ipotesi H_0 (ipotesi nulla). Essa postula frequentemente la completa accidentalità dell’esito sperimentale. Dall’analisi dei dati campionari, secondo regole di “comportamento induttivo”, si arriva alla decisione di accettare o di rifiutare H_0 . L’ipotesi nulla quindi è un’assunzione sui parametri della popolazione, ossia:

$$H_0 : \Theta = \Theta_0 \quad \text{con } \Theta_0 \text{ particolare valore di } \Theta$$

Tale ipotesi viene posta a confronto con una alternativa H_1 che viene formulata sul valore che lo stesso parametro può assumere in alternativa a quello considerato con H_0 . Ad esempio:

$$H_1 : \Theta \neq \Theta_0$$

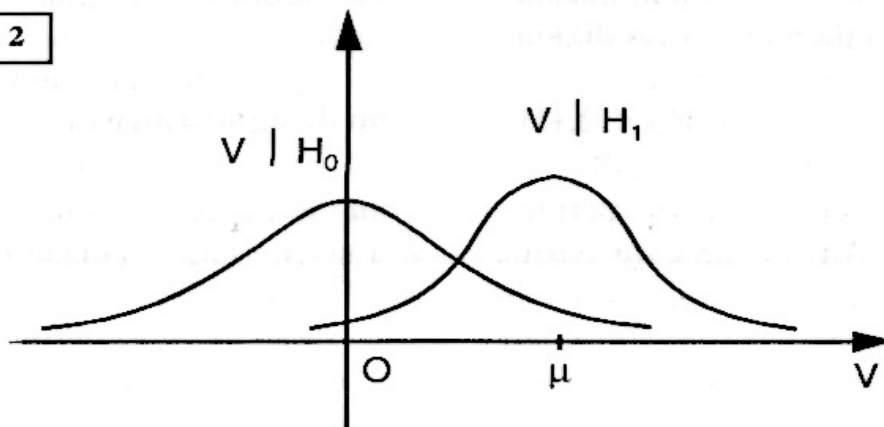
Lo strumento di decisione è un test statistico basato su una opportuna funzione definita sull'universo dei campioni (**funzione test**). Poiché gli elementi costituenti un campione sono delle variabili casuali, è v. c. anche ogni funzione test. Esiste quindi la possibilità di derivare la distribuzione di probabilità di tale variabile in funzione della distribuzione degli elementi campionari.

Al variare delle n -ple campionarie, la funzione test è tale che:

- 1) quando H_0 è vera, è descritta da una variabile aleatoria V condizionata ad H_0 ($V|H_0$) avente un'opportuna distribuzione (densità) di probabilità;
- 2) quando H_1 è vera, è descritta da una variabile aleatoria $V|H_1$ avente una distribuzione (densità) della stessa classe della variabile $V|H_0$, ma con valori diversi dei parametri.

La fig. 2 illustra una possibile situazione concernente la funzione test sotto l'assunto che $V|H_0$ e $V|H_1$ siano distribuite normalmente rispettivamente con medie zero e μ e varianze $\text{var}(V|H_0)$ e $\text{var}(V|H_1)$.

Fig. 2



La funzione test spesso è costruita definendo una “distanza”, espressa in un’opportuna unità di misura, tra la distribuzione (densità) di probabilità che l’ipotesi nulla permette di assegnare ai “punti” dell’universo campionario e un punto campionario scelto opportunamente (valor medio, mediana, moda, ecc).

Convenzionalmente si decide di rifiutare l’ H_0 se la distanza tra il campione osservato x e il punto “privilegiato” supera un prefissato valore-soglia. Per la definizione di quest’ultimo, l’insieme Ω dei punti campionari viene suddiviso in due sottoinsiemi, Σ e $\Omega \setminus \Sigma$, in modo tale che si decide di rifiutare H_0 se $x \in \Sigma$ e di accettarla (non rifiutarla) se $x \in \Omega \setminus \Sigma$. Lo spazio Σ di rifiuto di un’ipotesi viene detto “regione critica” e l’appartenenza di un punto campionario a Σ viene stabilita in base al valore assunto su di esso dalla funzione test “opportunamente” scelta.

Ovviamente la decisione presa può essere corretta o no e non sarà dato sapere con “certezza” se la particolare ipotesi formulata sia vera o falsa. Questo perché la prova di verifica di ipotesi statistiche discende da un processo induttivo che utilizza quantità limitate di informazioni.

Vanno introdotti allora tutti i possibili errori che si possono commettere nell’estendere le conclusioni che possono emergere su base campionaria alla popolazione da cui proviene il campione x .

Un errore, errore di prima specie, consiste nel rifiutare l’ipotesi nulla quando, in realtà, questa è vera; per definirlo bisogna considerare la probabilità condizionata

$$\alpha = \text{Prob}(x \in \Sigma | H_0) = \text{livello di significatività}$$

e la quantità $1 - \alpha$ è interpretabile come la capacità del test statistico scelto di indicare correttamente come vera l’ipotesi nulla sulla base del campione o anche come livello di “fiducia” che si dà al test.

Bisogna considerare anche un secondo errore, errore di seconda specie, che si può commettere se l’ipotesi nulla viene considerata vera, mentre in realtà non lo è. La probabilità di questo tipo di errore

è espressa da

$$\beta = \text{Prob}(x \in \Omega \setminus \Sigma | H_1)$$

e la quantità $1 - \beta$, detta **potenza del test**, esprime la capacità del test di far rilevare la falsità dell'ipotesi nulla.

Il tutto può essere schematizzato nella seguente fig. 3

		SITUAZIONE REALE NELLA POPOLAZIONE	
		È vera H_0	Non è vera H_0
SITUAZIONE STIMATA IN BASE AL CAMPIONE	Si considera vera H_0	Decisione corretta Prob = $1 - \alpha$	Errore di II specie Prob = β
	Si considera non vera H_0	Errore di II specie Prob = α	Decisione corretta Prob = $1 - \beta$

Se alle ipotesi H_i vengono assegnate delle probabilità iniziali $p_i = \text{Prob}(H_i)$, la probabilità degli eventi legati a comportamenti corretti è data dalla

$$\text{Prob}((\{x \in \Sigma\} \cap H_1) \cup (\{x \in \Omega \setminus \Sigma\} \cap H_0)) = p_1 (1 - \beta) + p_0 (1 - \alpha)$$

e quella di comportamenti erronei dalla

$$\text{Prob}((\{x \in \Sigma\} \cap H_0) \cup (\{x \in \Omega \setminus \Sigma\} \cap H_1)) = p_0 \beta + p_1 \alpha.$$

Si cerca, ragionevolmente, di rendere massima la probabilità

degli eventi legati a comportamenti corretti e minima quella relativa a comportamenti errati. Pertanto se, come usualmente si procede, il livello di significatività α viene fissato e si suppone l'equiprobabilità delle ipotesi ($p_i = 1/2$), il problema di ottimizzazione consiste nel far sì che β sia minimo. Operativamente, fissato α , si cerca di individuare un test che minimizzi β o, in altri termini, di individuare una regione critica che massimizzi la potenza del test.

Il livello di significatività, comunque, non può essere troppo piccolo in quanto una diminuzione dell'errore di prima specie provoca un aumento di quello di seconda specie. Si può arrivare ad una corretta valutazione solo conoscendo la natura del problema in questione e bilanciando i pro e i contro degli errori e delle loro conseguenze.

Ovviamente quanto appena visto è la problematica che si presenta prima dell'osservazione del campione. Dopo l'estrazione di quest'ultimo, la probabilità che sia vera H_0 avendo osservato $x \in \Sigma$ e quella che sia vera H_1 avendo osservato $x \in \Omega \setminus \Sigma$ sono date rispettivamente, per il teorema di Bayes, da

$$\text{Prob}(H_0 \mid x \in \Sigma) = \frac{p_0 \alpha}{p_0 \alpha + p_1 (1 - \beta)} = \frac{\alpha}{\alpha - \beta + 1},$$

$$\text{Prob}(H_1 \mid x \in \Omega \setminus \Sigma) = \frac{p_1 \beta}{p_1 \beta + p_0 (1 - \alpha)} = \frac{\beta}{\beta - \alpha + 1}.$$

La condizione di ottimo richiesta, allora, premette di rendere minima sia la probabilità di verità dell'ipotesi nulla quando l'esito campionario cade nella regione critica, sia quella dell'ipotesi alternativa quando l'esito si trova nella regione di accettazione.

Si osservi, infine, che trattandosi di analisi fatte su campioni, in tutto questo processo inferenziale deve entrare in gioco anche la numerosità campionaria. Infatti risulta immediato che per un campione di numerosità pari a quella della popolazione verrebbero ad annullarsi in due tipi di errori e per uno di piccole dimensioni α e β potrebbero risultare troppo grandi.

4. INFERENZA SU PERCENTUALI

Problemi che spesso si incontrano nella realtà quotidiana e che possono costituire interesse da un punto di vista didattico sono quelli legati a inferenze su **“popolazioni dicotomiche”**.

Una popolazione dicotomica è un collettivo le cui unità sono classificate secondo due modalità di un carattere, ad esempio malato o sano, guarito non guarito, presente o assente, promosso o respinto, ecc. In tal caso interessa analizzare la frequenza relativa di una delle modalità del carattere.

La verifica d'ipotesi sul valore π_0 di una frequenza relativa o proporzione π di una data modalità di una variabile X dicotomica, consiste nel provare statisticamente se le osservazioni campionarie siano conformi o meno alla congettura fatta sul valore della frequenza. Il valore può essere “stimato” sulla base di conoscenze precedenti, di “intuizioni” sulla incidenza o prevalenza del fenomeno, attraverso un'indagine pilota, ecc.

Si supponga quindi di estrarre dalla popolazione dicotomica un campione casuale bernoulliano di dimensione n . Indicato con f il numero di unità che nel campione presentano la modalità oggetto d'interesse e con $p = f/n$ la corrispondente frequenza relativa, si prova che, nell'ambito dell'universo campionario, la v.c. frequenza assoluta F si distribuisce in modo binomiale e che per n grande (la regola empirica generalmente usata suggerisce di prendere n tale che $n\pi > 5$ e $n(1 - \pi) > 5$) la distribuzione della proporzione campionaria $P = F/n$ può essere approssimata da una gaussiana con media π_0 e varianza $\pi_0(1 - \pi_0)/n$.

La verifica dell'ipotesi nulla consiste nel confrontare il valore trovato p con il valore ipotizzato π_0 : piccole divergenze potranno essere attribuite a fluttuazioni campionarie, divergenze notevoli, invece, denotano che le osservazioni campionarie non si conformano al valore ipotizzato per la frequenza della popolazione.

Fissata l'ipotesi nulla

$$H_0 : \pi = \pi_0 \quad (\delta = \pi - \pi_0 = 0)$$

occorre considerare anche l'ipotesi alternativa H_1 che presume una

popolazione caratterizzata da una frequenza relativa π_1 e può essere quindi:

bidirezionale: $H_1 : \pi_1 \neq \pi_0 \quad (\delta = \pi_1 - \pi_0 \neq 0)$;

unidirezionale: $H_1 : \pi_1 > \pi_0 \quad (\delta = \pi_1 - \pi_0 > 0)$ oppure

$H_1 : \pi_1 < \pi_0 \quad (\delta = \pi_0 - \pi_1 > 0)$.

La quantità δ rappresenta la differenza minima tra frequenze che si ritiene rilevante e di cui, quindi, l'indagine deve essere in grado di evidenziare la significatività statistica. In pratica la reale differenza tra frequenze non è nota prima dell'indagine e la sua valutazione rappresenta, spesso, proprio lo scopo dello studio statistico.

Per la verifica di ipotesi si può far ricorso alla funzione test

$$Z = \frac{p - \pi}{[p_0(1 - \pi_0) / n]^{1/2}}$$

la cui distribuzione può essere approssimata con una normale standardizzata.

Parimenti, supponendo di voler confrontare le proporzioni π_1 e π_2 relative ad una stessa modalità di un carattere X in due popolazioni dicotomiche distinte, si estraggono due campioni bernoulliani di dimensioni n_1 e n_2 dalle rispettive popolazioni e si analizza la differenza tra le proporzioni di soggetti p_1 e p_2 che presentano la modalità nei campioni da confrontare. Sempre relativamente a campioni di grandi dimensioni la v.c. $P_1 - P_2$ si distribuisce normalmente con media $\pi_1 - \pi_2$ e varianza

$$[\pi_1(1 - \pi_1)/n_1] + [\pi_2(1 - \pi_2)/n_2].$$

In questo caso le ipotesi da prendere in esame sono:

$$H_0 : \delta = \pi_1 - \pi_2 = 0$$

contro una delle alternative:

$$H_1 : \delta = \pi_1 - \pi_2 \neq 0;$$

$$H_1 : \delta = \pi_1 - \pi_2 > 0 \quad \text{oppure} \quad H_1 : \delta = \pi_2 - \pi_1 > 0$$

e, assumendo:

$$p^* = (n_1 p_1 + n_2 p_2) / (n_1 + n_2)$$

come stima della proporzione incognita $\pi = \pi_1 = \pi_2$ sotto l'ipotesi H_0 , la funzione test utilizzabile sarà:

$$Z = \frac{p_1 - p_2}{[p^* (1 - p^*) (1/n_1 + 1/n_2)]^{1/2}}. \quad (1)$$

Sia nel caso di confronto tra proporzione campionaria e quella della popolazione che in quello tra due proporzioni, il rifiuto o meno dell'ipotesi nulla è stabilito in base al valore z assunto dalla funzione test Z utilizzata sui dati campionari. Se z , in valore assoluto, risulta maggiore di un opportuno "valore critico", l'ipotesi nulla viene rifiutata; se risulta minore o uguale "in pratica" l' H_0 viene accettata. Il valore critico considerato, nel caso di "test unidirezionali", è dato dalla determinazione Z_α della variabile normale standardizzata tale che $\text{Prob}(|Z| > Z_\alpha) = \alpha$ (livello di significatività). Per "test bidirezionali" il valore critico viene individuato sostituendo a z_α il valore di $z_{\alpha/2}$.

Si consideri il seguente esempio. Da due popolazioni, una di non vaccinati (NV) e l'altra di vaccinati (V) con un vaccino antinfluenzale, si estraggono casualmente 800 e 700 soggetti rispettivamente. Nel primo campione si riscontrino 150 malati ($p_1 = 0.19$) e

nel secondo 100 ($p_2 = 0.14$); ci si chiede la capacità immunitaria del vaccino.

L'ipotesi nulla postula nessuna efficacia del vaccino ($\pi_1 = \pi_2$) e l'ipotesi alternativa una significativa attività immunitaria ($\pi_1 \neq \pi_2$). Si tratta quindi di applicare un "test bidirezionale".

Data l'alta numerosità campionaria si può ricorrere alla funzione test (1).

Una stima di π può essere data da $p^* = 250/1500 = 0.17$, che esprime l'incidenza "cumulata" della malattia nell'ambito dei due campioni (NV e V) in esame. Il denominatore della (1) (detto, in termini tecnici, errore standard di una proporzione) vale $[0.17 \cdot 0.83 \cdot (1/800 + 1/700)]^{1/2} = 0.02$ e quindi $|z| = 2.85$. Pertanto, supposto un livello di significatività $\alpha = 5\%$, si ha $|z| > z_{\alpha/2} = 1.96$ e, quindi, si è indotti a rifiutare H_0 ed accettare H_1 , ossia ad assumere l'efficacia del vaccino antinfluenzale per ridurre l'incidenza dei casi.

Qualora la stessa indagine fosse stata effettuata su due campioni rispettivamente di 80 e 70 soggetti con 15 e 10 malati rispettivamente, applicando la (1) si otterrebbe un valore di z uguale a 0.83, valore chiaramente non significativo (essendo minore di $z_{\alpha/2}$). Ciò mette in evidenza l'importanza della numerosità campionaria. Infatti la stessa incidenza percentuale di malati (19% nei non vaccinati e 14% nei vaccinati) risulta significativa (a favore del vaccino) quando i campioni hanno numerosità di 800 e 700 unità e di nessuna significatività statistica qualora l'indagine venga svolta su campioni di piccole dimensioni (80 e 70).

BIBLIOGRAFIA

- [1] M.L. BERENSON - D.M. LEVINE (1989), *Statistica per le Scienze Economiche*, Zanichelli, Bologna.
- [2] F. DI ORIO (1990), *Statistica Medica*, La Nuova Italia Scientifica, Roma.
- [3] G. PRODI (1992), *Metodi Matematici e Statistici*, McGrawHill Italia, Milano.
- [4] C. ROSSI - G. SERIO (1990), *La metodologia statistica nelle applicazioni biomediche*, Springer-Verlag, Berlin.
- [5] M. R. SPIEGEL (1992), *Statistica*, Etas Libri, Milano.

LA DIMENSIONE CAMPIONARIA NELLA VERIFICA DI IPOTESI STATISTICHE

Enzo Ballone e Vittorio Colagrande***

**Facoltà di Medicina e Chirurgia, Università "G. d'Annunzio" - Chieti*

***Liceo Scientifico "G. Galilei" - Lanciano (CH)*

1. INTRODUZIONE

Nel presente lavoro si esporranno alcune problematiche attinenti la dimensione minima "ottimale" di un campione per la verifica di ipotesi sul valore di una proporzione o sul confronto di due proporzioni. Si presenterà anche un software per il calcolo della dimensione stessa.

La trattazione è uno sviluppo della relazione presentata dagli autori al Convegno Interregionale della MATHESIS su "Calcolo delle Probabilità, Statistica e Crittografia nell'insegnamento", tenutosi a Isernia nei giorni 2-4 aprile 1992. Essa vuole essere una proposta didattica nell'ambito dell'insegnamento della Statistica nelle scuole secondarie superiori. La stessa scelta di riportare in appendice un software sviluppato da C. DI MASCIO (Facoltà di Medicina e Chirurgia, Università di Chieti), secondo la metodologia della programmazione strutturata, ha una valenza espressamente didattica.

Si supporranno noti concetti e tecniche concernenti il campionamento e la verifica di ipotesi statistiche.

2. PRESENTAZIONE DELLA PROBLEMATICHE

Il processo inferenziale della verifica di ipotesi statistiche consiste nel confrontare un'ipotesi relativa ad una popolazione con i risultati sperimentali che scaturiscono da un campione estratto dalla popolazione stessa.

Problemi importanti (perché spesso si incontrano nella realtà quotidiana) sono quelli relativi a inferenze statistiche su "popolazioni dicotomiche".

A tal proposito si consideri il seguente esempio. Si vuole studiare la prevalenza della tossicofilia (uso occasionale di droghe leggere) nell'ambito degli studenti delle scuole superiori di una città (popolazione) e si decide di condurre un'indagine campionaria.

Si stabilisce allora di scegliere casualmente, con opportune tecniche, un campione di studenti che possa ritenersi "rappresentativo" di tutta la popolazione studentesca oggetto di studio e su di esso si effettua l'indagine statistica.

Se dall'indagine risultasse una percentuale di tossicofilia del 9%, qualora valutazioni a livello nazionale ritengano verosimile che circa il 15% dei giovani delle scuole medie superiori facciano uso occasionale di droghe leggere, un problema potrebbe essere quello di verificare, attraverso l'applicazione di un appropriato test statistico, se nella popolazione scolastica della città in esame la prevalenza del fenomeno sia statisticamente inferiore.

Un secondo problema potrebbe essere quello di stabilire se la tossicofilia aumenti con l'età. Ciò si potrebbe verificare, ad esempio, confrontando statisticamente le percentuali di prevalenza tra le prime due classi e le ultime due degli istituti superiori in esame.

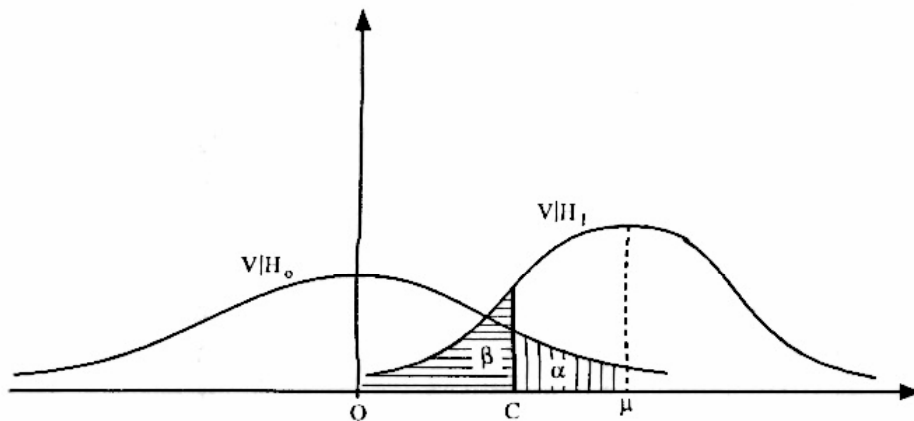
Per ottenere risultati attendibili da indagini campionarie e poter applicare quindi i processi dell'inferenza statistica, è necessario che il campione da analizzare sia "rappresentativo" della popolazione di provenienza, cioè rifletta il più possibile le sue caratteristiche.

In generale può affermarsi che quanto più il campione è numeroso, tanto più sarà rappresentativo; d'altra parte, però, al crescere della dimensione del campione crescono i costi e le difficoltà tecniche di organizzazione dell'indagine. La determinazione della dimensione campionaria costituisce, quindi, uno dei momenti di maggiore incertezza e riflessione da parte di un ricercatore che si appresti a realizzare un'indagine o un esperimento. Egli deve chiedersi, preventivamente, di quale grandezza debba essere il campione da esaminare; una scarsa attenzione alle dimensioni lascia il ricercatore nell'incertezza circa l'attendibilità dei risultati campionari e quindi può vanificare tutta la ricerca.

3. RELAZIONE FONDAMENTALE PER IL CALCOLO DELLA DIMENSIONE

Con riferimento alla problematica attinente la verifica di ipotesi nella ricerca statistica, si supponga che una variabile casuale test V , utilizzata per la verifica stessa, sia di tipo gaussiano con media m e varianza s^2 , cioè che la sua densità di probabilità sia $N_{m,s}(v)$. Essa abbia densità di probabilità $N_{0,s_1}(v)$ sotto l'ipotesi " H_0 : media = 0 e varianza = s_1^2 " e $N_{\mu,s_2}(v)$ sotto l'ipotesi alternativa " H_1 : media = μ e varianza = s_2^2 ".

Si può avere così, per $\mu > 0$ (ipotesi H_1 unidirezionale) una situazione come quella illustrata in figura seguente:



Nella figura sono individuate anche le aree relative agli errori di prima (Prob. = α) e seconda specie (Prob. = β), per un particolare valore c di V detto valore soglia, determinato in modo tale da minimizzare entrambi gli errori. È stato preso $0 < c < \mu$ e in quel che subito segue, si manterrà tale assunto.

Supponendo vera H_0 , il livello di fiducia è espresso dalla probabilità

$$1 - \alpha = \int_{-\infty}^c N_{0,s_1}(v) dv = 1/s_1 \int_{-\infty}^c N_{0,1}(v/s_1) dv = \int_{-\infty}^{c/s_1} N_{0,1}(t) dt,$$

Dunque c/s_1 è la determinazione numerica $z_{1-\alpha}$ della variabile normale standardizzata che soddisfa la relazione

$\text{Prob}(z < z_{1-\alpha}) = 1 - \alpha$ e, quindi, $c = z_{1-\alpha} s_1$.

Supponendo vera H_1 si ha, invece, che l'errore di seconda specie è misurato dalla β

$$\begin{aligned}\beta &= \int_{-\infty}^c N_{\mu, s_2}(v) dv = 1/s_2 \int_{-\infty}^c N_{0,1}((v - \mu)/s_2) dv = \\ &= \int_{-\infty}^{t_c} N_{0,1}(t) dt,\end{aligned}$$

avendo posto $t_c = (z_{1-\alpha} s_1 - \mu)/s_2$.

Pertanto t_c è proprio quel valore $-z_{1-\beta}$ della variabile normale standardizzata tale che $\text{Prob}(z < -z_{1-\beta}) = \beta$. Si può scrivere allora

$$-z_{1-\beta} = \frac{z_{1-\alpha} s_1 - \mu}{s_2}$$

da cui, osservato che $\mu = |\mu|$, si ricava la seguente

$$z_{1-\alpha} s_1 + z_{1-\alpha} s_2 = |\mu|. \quad (1)$$

La precedente formula ha validità anche nel caso $\mu < 0$.

Si osservi, comunque, che qualora l' H_1 fosse bidirezionale ($\mu \neq 0$) si prova che la relazione (1) rimane invariata nella sua "struttura", è sufficiente sostituire $z_{1-\alpha}$ con $z_{1-\alpha/2}$.

4. DETERMINAZIONE DELLA NUMEROSITÀ PER PROPORZIONI.

Nel caso si voglia verificare un'ipotesi statistica sul valore π_0 della frazione π di unità, appartenenti ad una data popolazione, che presenta una delle due modalità di un carattere dicotomico, l'ipotesi nulla H_0 può essere formulata come:

$$H_0: \pi = \pi_0 \quad (\delta = \pi - \pi_0 = 0)$$

e quella alternativa H_1 può esprimersi con una delle due ipotesi unidirezionali

$$H_1: \pi > \pi_0 \quad (\delta = \pi - \pi_0 > 0),$$

oppure

$$H_1: \pi < \pi_0 \quad (\delta = \pi - \pi_0 < 0)$$

Il parametro π della popolazione assume il significato di frequenza relativa delle unità che in essa possiedono l'attributo preso in esame. Nel seguito esso verrà denominato indifferentemente con termini quali "proporzione", "frequenza", "percentuale".

Per la verifica dell'ipotesi nulla occorre sottoporre a test la differenza tra il valore p della frequenza in un campione e quello ipotizzato π_0 della frequenza nella popolazione di riferimento.

Il valore $\delta = \pi - \pi_0$ che figura nella formulazione delle ipotesi può essere preso come quella differenza minima tra proporzioni che il ricercatore ritiene "rilevante" ai fini dell'indagine e di cui vuole verificare la significatività statistica. Esso va fissato prima di intraprendere l'indagine, pertanto l'ipotesi alternativa H_1 viene a presumere una popolazione caratterizzata da una frequenza $\pi_1 = \pi + \delta$.

Facendo riferimento ad un'estrazione campionaria di tipo casuale bernoulliano (con n si denoterà la dimensione di un campione) e indicata con P la variabile casuale proporzione campionaria, la variabile test V è data, nel caso di n "grande", dalla

$$Z = \frac{P - \pi_0}{[\pi_0 (1 - \pi_0) / n]^{\frac{1}{2}}}$$

Essa ha densità di probabilità con parametri

$$\mu = \frac{\delta}{[\pi_0 (1 - \pi_0) / n]^{\frac{1}{2}}}, \quad s_1 = 1, \quad s_2 = \frac{[\pi_1 (1 - \pi_1) / n]^{\frac{1}{2}}}{[\pi_0 (1 - \pi_0) / n]^{\frac{1}{2}}}.$$

Sostituendo queste formule nella (1) si ottiene:

$$z_{1-\alpha} + z_{1-\beta} \frac{[\pi_1 (1 - \pi_1) / n]^{\frac{1}{2}}}{[\pi_0 (1 - \pi_0) / n]^{\frac{1}{2}}} = \frac{|\delta|}{[\pi_0 (1 - \pi_0) / n]^{\frac{1}{2}}},$$

da cui si ricava il valore minimo della dimensione campionaria:

$$n \approx \left[\frac{z_{1-\alpha} [\pi_0 (1 - \pi_0)]^{\frac{1}{2}} + z_{1-\beta} [\pi_1 (1 - \pi_1)]^{\frac{1}{2}}}{\delta} \right]^2 . \quad (2)$$

La condizione di minimo va intesa nel senso che ogni altro valore di n superiore a quello determinato con la (2) permetterà di ottenere risultati statisticamente più attendibili.

La (2) esprime la dimensione ottimale del campione in dipendenza della variabilità del fenomeno, del valore di δ e dei livelli di probabilità di errore α e β .

Riprendendo l'esempio delle tossicofilie introdotto nel par. 2, se a livello nazionale si stima la frequenza del fenomeno, nei giovani delle scuole superiori, intorno al 15% (π_0) e si ritiene che nelle scuole superiori della città esaminata il fenomeno sia meno frequente qualora si riscontrasse un valore della differenza tra frequenze minore o uguale al 6% ($\delta = - 0.06$), si può calcolare il numero di studenti da includere nel campione. Dalla (2), per valori di $\alpha = 0.05$ e $1 - \beta = 0.90$ cui corrispondono $z_{1-\alpha} = 1.645$ e $z_{1-\beta} = 1.282$ (come si ricava consultando opportune tavole della distribuzione normale o utilizzando un software specifico come quello proposto nel successivo par. 6), si ottiene

$$n = \text{ROUND} \left[\left[\frac{1.645 \cdot (0.15 \cdot 0.85)^{\frac{1}{2}} + 1.282 \cdot (0.09 \cdot 0.91)^{\frac{1}{2}}}{0.06} \right]^2 \right] = 253 ,$$

dove con ROUND[...] si è indicato il minimo intero maggiore o uguale della quantità entro parentesi.

Per realizzare le condizioni suddette occorre, quindi, un campione di studenti di almeno 253 unità.

Qualora si volesse una differenza tra frequenze minore o uguale al 9% ($\delta = - 0.09$), si può ricavare che, agli stessi livelli di errori di prima e seconda specie, la dimensione campionaria minima sarebbe di 98 unità. Sempre a titolo di esempio, qualora risultasse a livello nazionale un valore di $\pi_0 = 0.25$ e si volesse una differenza

minore o uguale al 10% ($\delta = -0.10$) agli stessi livelli di errore, la numerosità campionaria richiesta sarebbe di $n = 137$.

Nel caso si desideri verificare se esiste una diversa prevalenza della tossicofilia tra le prime due classi e le ultime due, è necessario fissare la differenza δ di percentuale di casi che si ritiene rilevante tra i due gruppi e determinare quindi le dimensioni minime dei due campioni.

Le ipotesi da saggiare sono quelle relative al confronto tra due proporzioni π_1 e π_2 di una stessa modalità di un carattere in due popolazioni dicotomiche distinte.

L'ipotesi nulla può essere formulata come:

$$H_0: \delta = \pi_1 - \pi_2 = 0$$

e quella alternativa prendendo in esame una delle ipotesi unidirezionali:

$$H_0: \delta = \pi_1 - \pi_2 > 0 \text{ oppure } H_1: \delta = \pi_1 - \pi_2 < 0.$$

Considerando anche in questo caso campioni casuali bernoulliani (n_1 e n_2 indicheranno le dimensioni di due campioni) e indicata con $P_1 - P_2$ la variabile casuale delle proporzioni campionarie, la funzione test da utilizzare nel caso di "grandi campioni" è la seguente:

$$Z = \frac{P_1 - P_2}{[p^* (1 - p^*) (1/n_1 + 1/n_2)]^{\frac{1}{2}}}.$$

Z è distribuita con parametri

$$\mu = \frac{\delta}{[p^* (1 - p^*) (1/n_1 + 1/n_2)]^{\frac{1}{2}}}, \quad s_1 = 1,$$

$$s_2 = \frac{[p_1 (1 - p_1)/n_1 + p_2 (1 - p_2)/n_2]^{\frac{1}{2}}}{[p^* (1 - p^*) (1/n_1 + 1/n_2)]^{\frac{1}{2}}},$$

dove p_1, p_2 sono opportune frequenze campionarie note assunte come stime delle frequenze π_1 e π_2 delle due popolazioni (studenti delle prime due classi e delle ultime due nel problema in esame) e

la proporzione p^* può essere data dalla media ponderata delle due frequenze campionarie p_i ($p^* = (n_1 p_1 + n_2 p_2) / (n_1 + n_2)$), dalla prevalenza registrata in altre indagini simili o essere desunta dalla bibliografia.

Sostituendo i parametri, la (1) diventa

$$z_{1-\alpha} + z_{1-\beta} \frac{[p_1 (1 - p_1) / n_1 + p_2 (1 - p_2) / n_2]^{\frac{1}{2}}}{[p^* (1 - p^*) (1/n_1 + 1/n_2)]^{\frac{1}{2}}} =$$

$$= \frac{|\delta|}{[p^* (1 - p^*) (1/n_1 + 1/n_2)]^{\frac{1}{2}}}.$$

Un primo dimensionamento campionario lo si può ottenere assumendo che i due campioni abbiano la stessa numerosità, ossia $n_1 = n_2 = n$. In tal caso, ricavando n dalla formula precedente, si ha

$$n \approx \left[\frac{z_{1-\alpha} [2p^* (1 - p^*)]^{\frac{1}{2}} + z_{1-\beta} [p_1 (1 - p_1) + p_2 (1 - p_2)]^{\frac{1}{2}}}{\delta} \right]^2.$$

Nell'esempio della tossicofilia, allora, ipotizzando che le percentuali di studenti interessati dal fenomeno siano, ad esempio, $p_1 = 0.07$ e $p_2 = 0.13$, con un livello di significatività $\alpha = 0.05$ e una potenza $1 - \beta = 0.90$, dalla (3) si ottiene

$$n = \text{ROUND} \left[\left[\frac{1.645 \cdot (2 \cdot 0.10 \cdot 0.90)^{\frac{1}{2}} + 1.282 \cdot (0.07 \cdot 0.93 + 0.13 \cdot 0.87)^{\frac{1}{2}}}{0.06} \right]^2 \right] = 426,$$

per cui vanno indagati almeno 426 ragazzi delle prime e seconde e 426 delle ultime due classi.

Se si ritiene che questa ricerca, tendente a scoprire una eventuale differenza tra due sottogruppi di studenti, sia più particolare, potrebbe imporsi un livello di significatività "più rigoroso", ad esempio $\alpha = 0.01$ e, a parità di potenza, la numerosità di ciascuno dei due campioni dovrebbe essere di almeno 649, come risulta dal calcolo seguente

$$n = \text{ROUND} \left[\left[\frac{2.328 \cdot (2 \cdot 0.10 \cdot 0.90)^{\frac{1}{2}} + 1.282 \cdot (0.07 \cdot 0.93 + 0.13 \cdot 0.87)^{\frac{1}{2}}}{0.06} \right]^2 \right] = 649.$$

Si fa osservare che la numerosità campionaria è inversamente proporzionale al quadrato della differenza δ prefissata, a parità di errori di prima e seconda specie. Così se si dimezza la differenza nella percentuale di casi da scoprire nel campione (rispetto alla popolazione) o tra due campioni, è necessario quadruplicare le dimensioni dell'indagine o ricerca. D'altra parte n aumenta al diminuire dei due tipi di errori, condizione questa abbastanza ovvia in quanto, qualora si voglia minimizzare la probabilità di errore, necessita aumentare la numerosità del campionamento.

5. CONSIDERAZIONI CONCLUSIVE

Con questa breve nota non si ritiene di aver esaurito l'argomento, ma di avere, se non altro, presentato il quesito fondamentale che tutti i ricercatori dovrebbero porsi nel pianificare la loro indagine, e cioè quanti soggetti o oggetti sono necessari per soddisfare gli obiettivi della ricerca. Ciò affinché i risultati della ricerca stessa siano attendibili e, principalmente, permettano di fare dell'inferenza statistica, ossia di effettuare la verifica delle ipotesi sulla base dei risultati campionari.

Si fa rilevare che il criterio adottato nel lavoro è quello di assegnare una differenza minima tra frequenze relative che si ritiene rilevante, differenza definita in termini di $\delta = \pi_0 - \pi_1$ oppure $\delta = \pi_1 - \pi_2$. Si potrebbe specificare un valore δ_0 diverso da δ , al fine di cautelarsi maggiormente rispetto agli errori di campionamento. Infatti non si può mai "garantire" una differenza significativa dal momento che le fluttuazioni campionarie possono portare ad un valore di $p - \pi_0$ oppure $p_1 - p_2$ molto più piccolo di δ e, quindi, ad un non-rifiuto dell'ipotesi nulla.

È opportuno far presente, infine, che la determinazione della numerosità campionaria è stata sviluppata nella presente nota secondo un criterio di "ottimalità" ma che tuttavia esistono diverse situazioni pratiche che possono limitare la dimensione ottimale. Ad esempio, per ragioni di carattere economico, etico, ecc., può essere desiderabile che il numero di unità da sottoporre ad esperimento sia il minimo possibile. Si pensi a quando l'esperimento comporta la distruzione dell'unità da controllare oppure quando esso presuppone la somministrazione di un farmaco del quale non è ancora del tutto nota la tossicità sull'uomo.

6. SOFTWARE PER IL CALCOLO DELLA DIMENSIONE CAMPIONARIA. (a cura di CLAUDIO DI MASCIO)

Si riporta di seguito il listato di un programma in Turbo Pascal per il dimensionamento di campioni al fine di effettuare verifiche d'ipotesi su frequenze relative.

La metodologia utilizzata per la costruzione del programma è quella della programmazione strutturata, in quanto molto efficace da un punto di vista didattico.

Le varie procedure sono facilmente interpretabili. Va osservato comunque che il programma utilizza la Unit esterna "STAT" per la determinazione dei valori $z_{1-\alpha}$ e $z_{1-\beta}$ della variabile normale standardizzata. La Unit è stata costruita con la tecnica della Programmazione Orientata agli Oggetti (OOP). In essa viene definito l'oggetto "normale", risultato dell'"incapsulamento" di dati (i parametri di una distribuzione normale) e metodi (uno per il calcolo dell'area al di sotto della curva normale e un secondo per la determinazione del valore x_0 della variabile normale tale che la Prob ($a < x < x_0$) sia uguale ad un valore prefissato). Tale Unit può essere vista come una specie di "cassetta di strumenti" statistici incrementabile volta volta e utilizzabile per la soluzione di problemi.

Il programma richiede in input i seguenti dati:

- livello di significatività α e potenza del test $1 - \beta$;
- differenza δ che si ritiene rilevante;
- frequenza relativa π_0 nota o stimata nella popolazione, se trattasi di inferenza su un campione; frequenze relative (π_1 e π_2) delle due popolazioni analizzate, se trattasi di inferenza su due campioni (in questo caso $\delta = \pi_1 - \pi_2$).

PROGRAM DIMENSIONE_CAMPIONARIA;

```
uses crt, STAT;
var   N           :normale;
      alpha,potenza,beta   :real;
      z1,z2,po,p,d,p1,p2   :real;
      num          :integer;
      ch           :char;

Function SCELTA (scritta:string; ch1, ch2: char): char;
var   x,y:word;
      ch :char;
begin
write(scritta); x:= wherex; y:= wherey;
repeat
  gotoxy (x, y); clrcol; ch:= readkey; ch:= upcase (ch); write (ch)
until (ch=ch1) or (ch=ch2);
SCELTA:= ch;
end;
Procedure RIGHE (R:WORD);
var i:word;
begin
  for i:= 1 to R do writeln
end;
Procedure ERRORI;
var x,y: word;
begin
  RIGHE (2); write ('Livello di significatività  $\alpha$  = ');
  x:= wherex; y:= wherey;
  repeat
    gotoxy (x, y); clrcol; read (alpha)
  until (alpha < 0.5) and (alpha > 0);
  write(' Potenza del test  $1 - \beta$  = ');
  x:= wherex; y:= wherey;
  repeat
    gotoxy (x, y); clrcol; read (potenza);
  until (potenza > 0.5) and (potenza < 1);
  RIGHE (1); x:= wherex; y:= wherey;
  gotoxy (1,24); write ('Attendere Prego...!');
  z1:= N.z (-4*N.s + N.m, 4*N.s + N.m, 1 - alpha);
  z2:= N.z (-4*N.s + N.m, 4*N.s + N.m, potenza);
  gotoxy (1,24); delline; gotoxy (x,y);
  writeln ('z1-  $\alpha$  = ',z1 : 6 : 3); writeln ('z1 -  $\beta$  = ',z2 : 6 : 3);
end;
Procedure CAMPIONE;
var   x, y :word;
      ch   :char;
begin
  RIGHE (3); writeln ('CASO DI UN CAMPIONE' :45);
writeln; writeln;
  write ('Frequenza relativa  $\pi_0$  = '); readln (po);
  write ('Differenza "rilevante"  $\delta$  = '); readln (d);
```

```

        RIGHE (1);
        p:= po + d;
    end;
    Procedure DIM_CAMPIONE;
    var calc: real;
    begin
        RIGHE (1); d:= ABS (d);
        calc:= (z1*sqrt (po*(1 - po)) + z2*sqrt (p*(1-p))) / d;
        num:= ROUND      (calc*calc);
        gotoxy (23,22); write ('DIMENSIONE CAMPIONARIA n = ', num);
    end;
    Procedure CAMPIONI;
    var  x, y  : word;
        ch  :char;
    begin
        RIGHE (3); writeln ('CASO DI DUE CAMPIONI': 45);
    writeln; writeln;
        write ('Frequenza nella Prima Popolazione  $\pi_1$  = '); readln (p1);
        write ('Frequenza nella Seconda Popolazione  $\pi_2$  = ');
    readln (p2);
        d:= p1 - p2;
        write ('Differenza che si ritiene "rilevante"  $\delta$  = ', d : 4 : 2);
    end;
    Procedure DIM_CAMPIONI;
    var calc, media: real;
    begin
        RIGHE (1); media:= (p1 + p2) / 2; d:= ABS (d);
        calc:= (z1*sqrt (2*media*(1 - media)) + z2*sqrt (p1*(1 - p1)
            + p2*(1 - p2))) / d;
        num:= ROUND (calc*calc);
        gotoxy (23, 22); write ('DIMENSIONE CAMPIONARIA n = ', num);
    end;

    {MAIN BODY}
    begin
        repeat
            clrscr; N.m:= 0; N.s:= 1;
            writeln ('CALCOLO DELLA DIMENSIONE MINIMA DI UN CAMPIONE': 60);
            ERRORI; RIGHE (1);
            ch:= SCELTA ('Inferenza su (U)no o (D)ue campioni : ', 'U', 'D');
            if ch = 'U' then
                begin CAMPIONE; DIM_CAMPIONE; end
            else
                begin CAMPIONI; DIM_CAMPIONI; end;
            gotoxy (1, 25);
        until SCELTA ('Vuole Continuare ? (S/N) ', 'S', 'N') = 'N'
    end.

```

Il listato della Unit STAT è il seguente.

UNIT STAT;

Interface

```
type
    normale = object
        m,s: real; {parametri della distribuzione normale}
        function f (x: real) : real;
        function area (a, b: real) : real;
        function z (a, b, alpha: real): real;
    end;

Implementation

function pot (base, esponente: word): real;
var    i: word;
        p: real;
begin
    p:= 1; for i:= 1 to esponente do p:= p*base;
    pot:= p
end;

function normale.f(x:real):real;
begin
    f:= exp (- (x - m)*(x - m) / (2*s*s)) / (s*sqrt (2*Pi));
end;

function normale.area;
{Calcolo dell'area al di sotto della curva normale utilizzando
 il metodo di integrazione di Simpson}
var
    h, r, l, p, int1      : real;
    n                     : word;
begin
    n:= 100; h:= (b - a) / n; r:= a; int1:= 0;
    repeat
        l:= r + 2*h; p:= r + h;
        int1:= int1 + ((l - r) / 6)*(f (r) + f (l) + 4*f (p));
        r:= l
    until r>=(b-h);
    area:=int1
end;

function normale.z;
{Risoluzione dell'equazione Prob (a < x < x0) = alpha nell'incognita x0 col metodo di
Newton}
var
    y, y1, e: real;
    i: word;
        function g (x: real): real;
        begin
            g:= area (a, x) - alpha
        end;
```

```
begin
  i := 1; y1 := (a + b) / 2;
  repeat
    y := y1 - g(y1) / f(y1); e := abs(y - y1);
    y1 := y; i := i + 1
  until (i > 100) or (e < 1 / pot(10, 3));
  z := y
end;
end.
```

BIBLIOGRAFIA

- [1] M.L. BERENSON - D.M. LEVINE (1989), "Statistica per le Scienze Economiche", Zanichelli, Bologna.
- [2] J. BISHOP (1990), "Pascal - Corso di programmazione", Addison-Wesley Masson, Milano.
- [3] B.M. CESANA - E. MARUBINI, "Dimensione del campione nella ricerca bio-medica", Collana "Centro Zambon" n. 8, Milano.
- [4] B.J. COX (1990), "Programmazione Orientata agli Oggetti", Addison-Wesley Masson, Milano.
- [5] M. R. SPIEGEL (1992), "Statistica", Etas Libri, Milano.

DALLA COERENZA ALLA FORMULA DI BAYES: UN PERCORSO TEORICO FONDAMENTALE PER LE APPLICAZIONI

Giuseppe Di Biase
Università "G. D'Annunzio" - Facoltà di Farmacia
Palazzo de Pasquale, Via Vicoli - 66100 Chieti

Sommario. Dopo aver illustrato il concetto di coerenza, essenziale per rendere operativa in modo corretto la definizione di probabilità come grado di fiducia, viene mostrato come da esso discendono i due teoremi fondamentali del Calcolo delle probabilità: probabilità totali e composte. Si perviene, quindi, alla formula di Bayes ottenuta come rilettura del secondo dei teoremi citati. Infine due esempi significativi, uno di natura applicativa, che coinvolge anche l'uso di una procedura automatica recentemente sviluppata da Di Biase in [5] per verificare la coerenza di assegnamenti di probabilità condizionate, e l'altro di natura didattica, ne consentono la comprensione per un uso appropriato.

1. INTRODUZIONE

La probabilità di un evento E è la misura del grado di fiducia che un individuo ha sul verificarsi di E . La sua valutazione dipende dallo stato di informazione dell'individuo e può, quindi, essere giudicata diversamente da un altro individuo. Inoltre, uno stesso individuo può cambiare la propria opinione in seguito all'acquisizione di nuove informazioni.

Ad esempio oggi un tifoso (piuttosto ottimista) può valutare uguale a 0.8 la probabilità che il Pescara rimanga in serie B, un altro, con argomentazioni altrettanto valide, potrebbe valutarla uguale a 0.1. È ragionevole pensare che il primo, se venisse a conoscenza del fatto che i dirigenti sono in completo disaccordo sui premi salvezza, diminuirebbe il suo grado di fiducia e che il secondo, se venisse a sapere che la squadra ha acquistato un bravo giocatore, lo aumenterebbe. Per rendere operativa questa definizione di probabilità esistono vari criteri (cfr. ad es. de Finetti in [3], Scozzafava in [10] e Gilio in [7]). Brevemente qui si illustra quello della *scommessa coerente*.

La probabilità di un evento E è l'importo p che un individuo è disposto a pagare, effettuando una *ipotetica* scommessa, per ricevere 1 se E si verifica oppure 0 se E non si verifica (perdendo quindi l'importo p). È ragionevole pensare che la scommessa è "corretta" se l'individuo è disposto, oltre che ad effettuarla, anche a riceverla, ovvero a ricevere p con l'impegno di pagare 1 se E si verifica oppure 0 se E non si verifica (guadagnando l'importo p).

Una scommessa in cui sia stata introdotta la clausola di scambiabilità delle parti è detta coerente.

Alterando i fattori di scala si può parlare di quota di scommessa coerente pS (S arbitrario, eventualmente negativo in caso di ricezione della scommessa) per ricevere (pagare) S se E si verifica e niente in caso contrario.

All'inché la scommessa avvenga in maniera coerente occorre che l'importo p sia "giusto" e quindi che la probabilità di E , indicata anche con $p(E)$, sia "corretta". Ciò avviene se $p = p(E)$ è tale che nessuno dei due scommettitori può effettuare scommesse che gli diano *a priori* una vincita od una perdita certe, a prescindere dall'esito di E .

Nel caso di scommessa su un solo evento ciò comporta che p non può essere negativo, né maggiore di 1, cioè che $p \in [0,1]$.

Per realizzare tecnicamente una scommessa o una combinazione di scommesse coerenti vi sono varie maniere (cfr. ad es. de Finetti in [3], Scozzafava in [9] e [10], Maturo in [8]). Si illustra di seguito quella dovuta a quest'ultimo autore.

2. REALIZZAZIONE DI UNA COMBINAZIONE DI SCOMMESSE COERENTI

Se un individuo A vuole scommettere con un individuo B su n eventi qualsiasi $E_1, E_2, \dots, E_n$ delle somme $p_1S_1, p_2S_2, \dots, p_nS_n$ in maniera che $p_1, p_2, \dots, p_n$ siano probabilità coerenti degli eventi $E_1, E_2, \dots, E_n$, la scommessa deve avvenire con le seguenti regole.

Per ogni $i \in \{1, 2, \dots, n\}$:

- (R1) A sceglie le p_i , $p_i \in [0,1]$
- (R2) B sceglie le S_i , (in particolare il segno).

La scelta del segno delle S_i implica la scelta di chi tra i due individui fa da banco o da scommettitore.

In sostanza A deve scegliere "bene" le $p_i = p(E_i)$ in maniera che B non possa fissare le S_i in modo da avere un guadagno sicuro.

Il guadagno aleatorio relativo ad una scommessa su un evento E è univocamente definito dalla formula:

$$G_E = (|E| - p)S$$

ove $|E|$ è l'indicatore dell'evento E così definito:

$$|E| = \begin{cases} 1 & \text{se } E \text{ si verifica} \\ 0 & \text{se } E \text{ non si verifica} \end{cases}$$

Il guadagno relativo ad una scommessa su un evento E è aleatorio in quanto il suo valore non è noto per carenza di informazione su E. Per l'individuo che effettua la scommessa sarà $G_E = S - pS$ in caso di vincita (E si è verificato) oppure $G_E = -pS$ in caso di perdita (E non si è verificato).

Per realizzare una scommessa coerente è necessario che i due possibili valori del guadagno non siano dello stesso segno.

Nel caso di una combinazione di scommesse su n eventi il guadagno aleatorio assume la forma:

$$G_{E_1, \dots, E_n} = (|E_1| - p_1)S_1 + (|E_2| - p_2)S_2 + \dots + (|E_n| - p_n)S_n.$$

I due individui A e B realizzeranno una combinazione di scommesse che è coerente se avranno fissato rispettivamente le p_i e le S_i in modo da non avere *a priori*, cioè prima di vedere quali degli eventi E_i si siano verificati, un guadagno (perdita o vincita) che sia certo.

Dalla condizione di coerenza discendono in modo naturale, come vedremo nei prossimi paragrafi, il Teorema delle probabilità totali:

$$(2.1) \quad A \cap B = \emptyset \Rightarrow p(A \cup B) = p(A) + p(B)$$

e quello delle probabilità composte:

$$(2.2) \quad p(A \cap B) = p(A) p(B/A).$$

Si può dimostrare inoltre, vedi ad esempio Bruno e Gilio in [1], che la coerenza è condizione anche sufficiente, oltre che necessaria, per la validità delle (2.1) e (2.2).

3. TEOREMI DELLE PROBABILITÀ TOTALI E COMPOSTE

Supponiamo di fare una scommessa coerente simultanea su due eventi incompatibili E_1 ed E_2 , ovvero una combinazione di scommesse su E_1 ed E_2 con importi rispettivi p_1 e p_2 in modo che tali valori siano valutazioni di probabilità coerenti per gli eventi incompatibili E_1 ed E_2 .

Il guadagno aleatorio relativo a tale scommessa è:

$$(3.1) \quad G_{E_1, E_2} = (|E_1| - p_1) 1 + (|E_2| - p_2) 1 = \begin{cases} 1 - (p_1 + p_2) & \text{se si vince} \\ - (p_1 + p_2) & \text{se si perde} \end{cases}$$

Se, invece, facciamo una scommessa coerente sull'evento E che risulti essere proprio la somma logica di E_1 e E_2 , di importo $p = p(E)$, il guadagno assume la forma:

$$(3.2) \quad G_E = (|E| - p) 1 = \begin{cases} 1 - p & \text{se si vince} \\ -p & \text{se si perde} \end{cases}$$

Essendo E_1 e E_2 incompatibili in realtà la prima scommessa equivale alla seconda. Infatti, poiché $E = E_1 \cup E_2$ vinceremo la scommessa su E se si verifica E_1 oppure se si verifica E_2 e la perderemo se non si verifica nè E_1 nè E_2 . D'altra parte vinceremo la prima scommessa se si verifica E_1 oppure se si verifica E_2 , guadagnando in entrambi i casi la somma $1 - (p_1 + p_2)$, la perderemo se non si verificano nè E_1 nè E_2 . Pertanto i guadagni aleatori relativi alle due scommesse debbono essere uguali e dalle (3.1) e (3.2)

segue:

$$p = p_1 + p_2$$

ovvero la (2.1).

La conclusione si può trarre ancora più semplicemente considerando che se E_1 ed E_2 sono incompatibili, l'indicatore dell'evento $E = E_1 \cup E_2$ gode della proprietà:

$$|E| = |E_1 \cup E_2| = |E_1| + |E_2|$$

e quindi essendo:

$$G_E = (|E| - p) = |E_1| + |E_2| - p, \quad G_{E_1, E_2} = |E_1| + |E_2| - (p_1 + p_2),$$

si ha:

$$G_E = G_{E_1, E_2} \Rightarrow p = p_1 + p_2.$$

Si dice probabilità condizionata $p(E/H)$ di un evento E subordinato all'evento H l'importo p che un individuo è disposto a scommettere coerentemente per ricevere:

- 1 se E si verifica, una volta che si sia verificato anche H ;
- 0 se E non si verifica, una volta che si sia verificato H ;
- p se H non si verifica (in questo caso cioè si tratta di una finta scommessa in quanto p viene restituito).

Naturalmente se si vuol vincere una somma S bisogna fissare come quota della scommessa l'importo pS .

Per dimostrare il teorema delle probabilità composte supponiamo dapprima di effettuare una scommessa coerente sull'evento H con l'obiettivo di vincere la somma S nel caso in cui quest'evento si verifichi. La relativa quota è pS con $p = p(H)$ ed il guadagno aleatorio è dato da:

$$G_H = (|H| - p)S = \begin{cases} S - pS & \text{se } H \text{ si verifica} \\ -pS & \text{se } H \text{ non si verifica} \end{cases}$$

Successivamente supponiamo di fare un'altra scommessa coerente sull'evento subordinato E / H di quota $p_1 = p(E / H)$ per vincere la somma unitaria. Il relativo guadagno aleatorio è:

$$G_{E/H} = |H| (|E| - p_1) = \begin{cases} 1 - p_1 & \text{se si verificano } H \text{ ed } E \\ -p_1 & \text{se si verificano } H \text{ ed } \bar{E} \\ 0 & \text{se } H \text{ non si verifica} \end{cases}$$

Ipotizzando che siamo noi a scommettere, S è positivo.

Supponiamo ora di fare simultaneamente tutte e due le scommesse precedenti, il guadagno aleatorio è dato dalla somma dei guadagni relativi alle singole scommesse:

$$\begin{aligned} G_{H, E/H} &= G_H + G_{E/H} = (|H| - p)S + |H| (|E| - p_1) = \\ &= |H| S - pS + |H| |E| - |H| p_1. \end{aligned}$$

Essendo S arbitrario, ma positivo, lo possiamo porre uguale proprio a p_1 , ottenendo:

$$G_{H, E/H} = |H| |E| - p p_1 = |H| |E| - p(H) p(E/H).$$

Osserviamo che

$$|H| |E| = \begin{cases} 1 & \text{se } H \text{ ed } E \text{ si verificano} \\ 0 & \text{altrimenti} \end{cases}$$

per cui $|H| |E| = |H \cap E|$. Il guadagno aleatorio relativo alle due scommesse coerenti su H e su E / H si può, quindi, scrivere nella seguente maniera:

$$G_{H, E/H} = |H \cap E| - p(H) p(E/H) = \begin{cases} 1 - p(H) p(E/H) & \text{se } H \cap E \text{ è vero} \\ -p(H) p(E/H) & \text{se } H \cap E \text{ è falso} \end{cases}$$

Tale espressione rappresenta il guadagno relativo ad una scommessa coerente sull'evento $H \cap E$ se e solo se $p(H) p(E/H)$ è

proprio la probabilità dell'evento $H \cap E$. Pertanto deve risultare:

$$(3.3) \quad p(H \cap E) = p(H) p(E / H)$$

che esprime il teorema delle probabilità composte.

Ovviamente, scambiando il ruolo di E ed H , la (3.3) si può anche scrivere:

$$(3.4) \quad p(E \cap H) = p(E) p(H / E).$$

Osserviamo che i costituenti associati all'evento condizionato E/H sono $C_1 = E \cap H$, $C_2 = \bar{E} \cap H$, $C_3 = \bar{H}$. Essi costituiscono una classe completa di eventi incompatibili, infatti:

$$C_1 \cap C_2 = \emptyset, \quad C_1 \cap C_3 = \emptyset, \quad C_2 \cap C_3 = \emptyset, \quad C_1 \cup C_2 \cup C_3 = \Omega.$$

Allora, dette x_1, x_2, x_3 le probabilità dei costituenti si ha:

$$x_1 + x_2 + x_3 = 1.$$

Inoltre la (3.3) si può esprimere anche nella seguente maniera:

$$(3.5) \quad \sum_{C_h \subseteq E \cap H} x_h = p(E / H) \sum_{C_h \subseteq H} x_h$$

ossia

$$(3.6) \quad x_1 = p(E / H) (x_1 + x_2).$$

4. LA FORMULA DI BAYES

Per il teorema delle probabilità composte si ha che:

$$(4.1) \quad p(\bar{H} \cap E) = p(\bar{H}) p(E / \bar{H})$$

Sommando membro a membro la (3.3) e la (4.1) e tenendo conto della proprietà commutativa del prodotto logico, si ha:

$$p(E \cap H) + p(E \cap \bar{H}) = p(H) p(E / H) + p(\bar{H}) p(E / \bar{H});$$

ma $E \cap H$ e $E \cap \bar{H}$ sono due eventi incompatibili e quindi per la (2.1) il primo membro è uguale a $p[(E \cap H) \cup (E \cap \bar{H})]$.

Essendo inoltre $(E \cap H) \cup (E \cap \bar{H}) = E$, si ottiene:

$$(4.2) \quad p(E) = p(H) p(E / H) + p(\bar{H}) p(E / \bar{H}) .$$

Dalla (3.4) si ricava, se $p(E) > 0$, che:

$$p(H / E) = \frac{p(E \cap H)}{p(E)} = \frac{p(H \cap E)}{p(E)} = \frac{p(H) p(E / H)}{p(E)}$$

e, quindi, per la (4.2):

$$(4.3) \quad p(H/E) = \frac{p(H) p(E / H)}{p(H) p(E / H) + p(\bar{H}) p(E / \bar{H})} .$$

Quest'ultima è la formula di Bayes che consente, in base al verificarsi di un evento E che sintetizza i dati osservati, di rivalutare con $p(H/E)$ la probabilità di un evento H , probabilità che inizialmente, cioè prima della raccolta dei dati, era stata valutata uguale a $p(H)$.

In base alle nostre conoscenze sull'evento H abbiamo fatto una certa valutazione $p(H)$ della sua probabilità (classica, frequentista, soggettiva) ma nutriamo forti dubbi che se avessimo su H maggiori informazioni, la nostra valutazione sarebbe più veritiera.

Andiamo ad osservare, allora, un altro fenomeno, descritto dall'evento E sul quale è possibile fare delle prove (o acquisirle), il cui esito è influenzato dal verificarsi o meno di H . Conosciuto il risultato di questa osservazione, modifichiamo $p(H)$ sostituendola con $p(H / E)$ calcolata con la (4.3), purchè sia possibile calcolare le probabilità che compaiono a secondo membro, cioè valutare in quale misura il verificarsi o meno di H influisce sul verificarsi di E .

Questa interpretazione giustifica l'impiego della formula di Bayes in molteplici campi (medico, legale, strutturale, farmacologico, geologico, etc.) per misurare il grado di validità di una ipotesi

formulata (evento H), mediante la conoscenza di dati sperimentali (evento E). Nei prossimi paragrafi si effettuano due esempi adatti a mostrare l'uso consapevole e la validità scientifica di tale formula.

Approfondimenti e molti altri esempi si possono trovare in Daboni, [2], Di Biase - Maturo, [6] e Scozzafava, [9].

Il primo, mutuato dai giochi con le carte, è particolarmente utile agli insegnanti delle scuole secondarie che facciano un po' di calcolo delle probabilità. Il secondo riguarda un'applicazione nell'ambito della diagnostica clinica.

5. UN ESEMPIO

Supponiamo di avere tre mazzi di carte napoletane A , B , C , identici all'apparenza, ma tali che nel mazzo B le dame sono state sostituite scambiandole con gli assi del mazzo C in modo che B ha otto figure (e otto assi) e C ne ha sedici (e nessun asso).

Ci proponiamo in primo luogo di calcolare la probabilità che, estraendo una carta da un mazzo scelto a caso, esca una dama ed in secondo luogo di fare un'ipotesi su quale sia il mazzo da cui è stata presa. Consideriamo in proposito gli eventi:

E = "la carta estratta è una dama",

H_j = "è stato scelto il mazzo j " con $j \in \{A, B, C\}$.

Poiché i mazzi apparentemente sono uguali è naturale assumere:

$$p(H_j) = \frac{1}{3}, \quad \forall j \in \{A, B, C\}.$$

Inoltre, dalle informazioni in nostro possesso, si ha:

$$p(E / H_A) = \frac{4}{40}, \quad p(E / H_B) = 0, \quad p(E / H_C) = \frac{8}{40}.$$

Poiché $\{A, B, C\}$ è una partizione di Ω , E si può scindere nella somma di eventi incompatibili:

$$E = (E \cap H_A) \cup (E \cap H_B) \cup (E \cap H_C).$$

Pertanto dalla (2.1) si deduce:

$$p(E) = p(E \cap H_A) + p(E \cap H_B) + p(E \cap H_C)$$

e dalla (3.4):

$$p(E) = \sum_J p(H_J) p(E / H_J) = \frac{1}{10}$$

che coincide con la media delle $p(E / H_J)$ ponderata con le $p(H_J)$.

Per valutare in maniera più soddisfacente le

$$p(H_J), \forall j \in \{A, B, C\}$$

possiamo fare ricorso alla formula di Bayes utilizzando delle informazioni aggiuntive.

Osserviamo l'esito di alcune estrazioni sul mazzo prescelto ed indichiamo con D tale evento. In base al risultato possiamo correggere le $p(H_J) = 1/3$ con i valori $p(H_J / D), \forall J \in \{A, B, C\}$.

Sia $D =$ "in due estrazioni successive è uscita una figura".

Si ha:

$$p(D / H_A) = \frac{12}{40} \cdot \frac{11}{39} = \frac{11}{130}; \quad p(D / H_B) = \frac{8}{40} \cdot \frac{7}{39} = \frac{7}{195};$$

$$p(D / H_C) = \frac{16}{40} \cdot \frac{15}{39} = \frac{2}{13}; \quad p(D) = \frac{1}{3} \cdot \frac{11}{130} + \frac{1}{3} \cdot \frac{7}{195} + \frac{1}{3} \cdot \frac{2}{13} = \frac{107}{1170}$$

da cui applicando la formula di Bayes si ottiene:

$$p(H_A / D) = \frac{33}{107}, \quad p(H_B / D) = \frac{14}{107}, \quad p(H_C / D) = \frac{60}{107}.$$

Il procedimento può essere iterato un numero qualsivoglia di volte effettuando altre estrazioni che ci consentono di perfezionare progressivamente le valutazioni di probabilità riguardo le ipotesi sul mazzo scelto a caso, sostituendo nella formula di Bayes le probabilità a priori con quelle a posteriori ottenute al passo precedente.

6. UN'APPLICAZIONE

Un esempio classico di tipo applicativo, molto sfruttato in letteratura, a volte anche in modo erroneo, è quello di un medico che deve diagnosticare una malattia sulla base di alcuni sintomi accusati da un paziente.

Qui si considera il caso in cui il paziente, Carlo, ed il dottore, Franco, sono due miei amici. Carlo, durante un aperitivo in compagnia, ci comunica di soffrire da un po' di giorni un certo malessere. Franco pensa subito che i sintomi denunciati possano derivare da una malattia H , piuttosto diffusa in quel periodo nella nostra città. Infatti sa che due malati su tre in città sono affetti dal virus che provoca H . Conoscendo la mia passione per la probabilità soggettiva ed avendo appreso, durante le numerose chiacchierate sull'argomento, che è lecito assumere come probabilità le frequenze relative, Franco, nel considerare $p(H) = 0,6\bar{6}$, è propenso a diagnosticare a Carlo la malattia H .

Ma, pungolato dalla mia presa in giro sulla scarsa professionalità della categoria dei medici, indaga più approfonditamente sulla sintomatologia presentata da Carlo e lo invita in laboratorio per una visita.

Durante la visita, individuando la presenza di altri sintomi non palesati in precedenza, consulta l'archivio del suo personal computer e verifica che, nella popolazione da lui assistita, la frequenza relativa con cui tali sintomi si manifestano in presenza della malattia H è pari a 0,15. Inoltre osserva che, in assenza di H , la sintomatologia riscontrata è presente con una frequenza relativa pari a 0,72.

Fornitimi questi dati invito il dottore ad usare la formula di Bayes.

Identificando probabilità e frequenza, indicato con E l'evento "presenza dei nuovi sintomi", essendo:

$$p(E / H) = 0,15, \quad P(E / \bar{H}) = 0,72$$

dalla (4.3) si ottiene:

$$(6.1) \quad p(H / E) = \frac{0,6\bar{6} \times 0,15}{0,6\bar{6} \times 0,15 + 0,3\bar{3} \times 0,72} = 0,29,$$

e, di conseguenza, $p(\bar{H} / E) = 0,71$.

A questo punto stupore e perplessità traspaiono nello sguardo di Franco che si rende conto immediatamente dell'errore che avrebbe commesso nel diagnosticare la malattia H .

Incuriosito ed affascinato dal semplice, ma allo stesso tempo potente, modello probabilistico utilizzato, Franco mi invita ad approfondire l'argomento.

Allo scopo di fare una diagnosi più attendibile gli propongo di seguire uno dei seguenti metodi che, volendo, possono essere usati anche simultaneamente.

- a) Se è possibile, si prendono in considerazione tutte le malattie $H_1, H_2, \dots, H_n$, che sono causa dei sintomi riscontrati, tali che:

$$\bigcup_{i=1}^n H_i = \Omega \quad \text{e} \quad H_i \cap H_j = \emptyset \quad \forall i \neq j \in \{1, 2, \dots, n\}.$$

In tal caso, scomponendo E in somma logica di eventi incompatibili, si deduce che la (4.3) viene riformulata nella seguente maniera:

$$(6.2) \quad p(H_i / E) = \frac{p(H_i) p(E / H_i)}{\sum_{i=1}^n p(H_i) p(E / H_i)}, \quad i \in \{1, 2, 3, \dots, n\}.$$

Tale formula ci consente, di calcolare la probabilità di ciascuna malattia possibile *a posteriori*, cioè dopo osservato E , purchè siano valutabili le probabilità *probative* $p(E / H_i)$, che esprimono il grado di fiducia sul verificarsi di E , in seguito all'informazione (eventualmente solo supposta) della verità della possibili malattie ipotizzate.

- b) Si sottopone il paziente a dei test clinici in grado di convalidare la malattia che il medico ritiene più probabile, oppure di scartare quella che egli ritiene meno probabile.

Per il primo scopo esistono test ad alta sensibilità e, per il secondo, test ad alta specificità. La *sensibilità* di un test da indicazioni sulla probabilità che un soggetto malato presenti un risultato positivo al test, la *specificità* che un soggetto non malato ne

presenti uno negativo.

Per chiarire le idee consideriamo la seguente tabella:

	TEST +	TEST -	TOTALI
MALATI	VP	FN	VP + FN
NON MALATI	FP	VN	FP + VN
TOTALI	VP+FP	FN + VN	VP + FN + FP + VN

ove:

TEST + = risultato positivo del test

TEST - = risultato negativo del test

VP = Veri Positivi (numero di individui positivi al test ed effettivamente malati)

FN = Falsi Negativi (numero di individui negativi al test ma risultati malati)

FP = Falsi Positivi (numero di individui positivi al test ma risultati non malati)

VN = Veri Negativi (numero di individui risultati negativi al test ed effettivamente non malati).

Indicato con s_e ed s_p rispettivamente la sensibilità e la specificità di un test si ha:

$$s_e = \frac{VP}{VP + FN}, \quad s_p = \frac{VN}{VN + FP}.$$

Fissato un test, se risulta $s_e > s_p$, esso si dice ad alta sensibilità; al contrario se $s_p > s_e$ il test è ad alta specificità.

Inoltre esistono dei criteri di tipo statistico (non esenti da critiche, ma abitualmente accettati), che consentono di valutare l'errore commesso nel considerare un test ad alta sensibilità o specificità.

Nell'esempio, essendo $p(H / E) = 0,29$ e $p(\bar{H} / E) = 0,71$, è ragionevole pensare che il medico suggerisca al paziente di effet-

tuare dei test idonei a ricusare la malattia “indagata”, ossia dei test ad alta specificità.

Sia $T = \{T_1, T_2, T_3\}$ una batteria di tali esami di laboratorio. Indichiamo con T_i l'evento “un soggetto ha avuto un risultato negativo rispetto all' i -esimo test” $i \in \{1, 2, 3\}$.

Le informazioni fornite da un laboratorio specializzato vengono sintetizzate in una tabella che contiene le frequenze relative di risultati negativi della batteria T , associata all'assenza della malattia H , con un errore fissato, rilevata su un campione di individui.

In tale maniera, identificando probabilità e frequenze possiamo assumere note le $p_i = p(T_i / H)$, $i \in \{1, 2, 3\}$ che rappresentano le probabilità che un soggetto ha di presentare un risultato negativo al test T_i , non avendo contratto la malattia H .

Poichè tali esami hanno un costo elevato, supponiamo che il dottore decida di prescriberne solo due, ad esempio T_1 e T_2 .

Dopo aver osservato l'evento $E = T_1 \cap T_2 =$ “il paziente è risultato negativo a T_1 e a T_2 ”, possiamo nuovamente applicare la formula di Bayes nella forma:

$$(6.3) \quad p(\bar{H} / E) = \frac{p(\bar{H}) p(E / \bar{H})}{p(\bar{H}) p(E / \bar{H}) + p(H) p(E / H)}$$

sostituendo le probabilità *a priori* $p(\bar{H})$ con quella *a posteriori* ottenuta con la (6.1).

Non essendoci ragioni per considerare indipendenti stocasticamente gli eventi T_1 e T_2 (vedi ad esempio Scozzafava in [11] e Di Biase [4]), in genere si avrà:

$$\begin{aligned} p(E / \bar{H}) &= p(T_1 \cap T_2 / \bar{H}) \neq p(T_1 / \bar{H}) p(T_2 / \bar{H}) \\ p(E / H) &= p(T_1 \cap T_2 / H) \neq p(T_1 / H) p(T_2 / H). \end{aligned}$$

Pertanto, dai dati raccolti presso il laboratorio, il dottore che conosce il calcolo delle probabilità non è in grado di valutare le probabilità che compaiono a secondo membro della (6.3).

A questo punto o si approfondiscono le indagini presso il laboratorio, deducendo tabelle di frequenza in cui i test della

batteria T vengono considerati anche due a due, oppure, applicando in maniera estrema la teoria soggettiva, il dottore esprime, in base alla sua esperienza, il proprio grado di fiducia nei riguardi di $p(E/\bar{H})$ e $p(E/H)$.

Supponiamo che egli ritenga ragionevoli le seguenti valutazioni.

$$(6.4) \quad p(E/\bar{H}) = 0.80 \quad , \quad p(E/H) = 0.12.$$

Per procedere in maniera corretta è necessario, però, verificare la *coerenza* di tali valutazioni.

Assegnata la coppia di eventi $E/\bar{H}$ ed E/H i costituenti ad essa associati si possono trovare eseguendo tutti i possibili prodotti logici, diversi dall'evento impossibile $\emptyset$, tra i costituenti associati ad $E/\bar{H}$ e quelli associati ad E/H . Si ottiene:

$$C_1 = E \cap H \quad , \quad C_2 = \bar{E} \cap H \quad , \quad C_3 = E \cap \bar{H} \quad , \quad C_4 = \bar{E} \cap \bar{H},$$

Se indichiamo con x_1, x_2, x_3, x_4 le loro rispettive probabilità, che ovviamente sono incognite si ha:

$$(6.5) \quad x_1 + x_2 + x_3 + x_4 = 1.$$

Allora, affinché valga il teorema delle probabilità composte nella forma (3.5), si deve avere:

$$(6.6) \quad x_3 = 0.80(x_3 + x_4),$$

$$(6.7) \quad x_1 = 0.12(x_1 + x_2).$$

Si dimostra (vedi ad esempio Gilio in [6]) che condizione sufficiente per la coerenza dell'assegnamento $p(E/\bar{H}) = 0,80$ e $p(E/H) = 0,12$ è che esistano x_1, x_2, x_3, x_4 tutti positivi e soddisfacenti le (6.5), (6.6) e (6.7). Allo scopo in Di Biase, [5], è stata implementata una procedura automatica per verificare la coerenza di un numero n qualsiasi (finito e compatibile con la capacità di calcolo e la memoria dell'elaboratore usato) di eventi condizionati.

Poichè ad esempio i numeri:

$$x_1 = 0.084, \quad x_2 = 0.616, \quad x_3 = 0.240, \quad x_4 = 0.06$$

soddisfano le condizioni richieste, possiamo inserire i valori (6.4) nella (6.3).

Si ottiene:

$$P(\bar{H} / E) = \frac{0.71 \times 0.80}{0.71 \times 0.80 + 0.29 \times 0.12} = 0.94$$

$$P(H / E) = \frac{0.29 \times 0.12}{0.29 \times 0.12 + 0.71 \times 0.80} = 0.06.$$

In conclusione (per così dire) essendo la probabilità di assenza della malattia H , avendo osservato il risultato negativo dei test T_1 e T_2 , molto maggiore di quella di presenza di H in seguito alla stessa osservazione, è ragionevole scartare la malattia H dal numero delle possibilità e, quindi, il mio amico Franco dovrà sudare sette camice, nel continuare le sue indagini!

BIBLIOGRAFIA

- [1] BRUNO G. e GILIO A. (1979). *La probabilità dal punto di vista soggettivo: breve introduzione*. Quaderni dell'Istituto di Comunicazioni Elettriche Università La Sapienza di Roma.
- [2] DABONI L. (1993) *Calcolo delle Probabilità ed elementi di statistica*. Utet, Torino.
- [3] DE FINETTI B. (1979) *Teoria della Probabilità*. Vol I, Einaudi, Torino.
- [4] DI BIASE G. (1990) *Sul concetto di indipendenza tra eventi*. Atti del Convegno Nazionale Mathesis: "Matematica negli anni novanta", Isco (BS), Aprile '90, pp. 161-172.
- [5] DI BIASE G. (1994). *Sulla ricerca di un insieme di probabilità condizionate coerenti nei sistemi esperti*. Atti del XVIII Convegno Amases, Modena, sett. 1994, pp. 357-368.
- [6] DI BIASE G. e MATURO A. (1992) *La formule de Bayes dans l'induction statistique*. Plot n. 60, Editeur: Associations régionales de l'APMEP de Poitiers, pp. 7-13.
- [7] GILIO A. (1990) *Criterio di penalizzazione e condizioni di coerenza nella valutazione soggettiva della probabilità*. Bollettino Umi (7) 4-B, pp. 645-660.
- [8] MATURO A. (1992) *Una introduzione alla probabilità soggettiva*. Periodico di Matematiche, serie VI, vol. 68, n. 3 - 1992, pp. 19-36.
- [9] SCOZZAFAVA R. (1993) *Le probabilità soggettiva e le sue applicazioni*. III edizione, Masson, Milano.
- [10] SCOZZAFAVA R. (1991) *Le probabilità condizionate: de Finetti o Kolmogoroff?* Scritti in omaggio a Luciano Daboni, ed. Lint, Trieste, pp. 223-237.
- [11] SCOZZAFAVA R. (1990) *Aspetti soggettivi del concetto di indipendenza*. Induzioni n. 0, 1990, ed. Giardini, Firenze, pp. 15-18.

PROBABILITÀ CONDIZIONATE E APPRENDIMENTO DALL'ESPERIENZA: UN'APPLICAZIONE DEL TEOREMA DI BAYES

Angelo Gilio

Dipartimento di Matematica - Università di Catania

Viale A. Doria, 6 - 95125 Catania

1. INTRODUZIONE

La matematica entra in molte situazioni della vita quotidiana, fornendoci concetti e metodi che permettono di analizzare problemi, al fine di prendere decisioni in modo razionale. Essendo in generale la nostra informazione incompleta, i problemi reali sono spesso caratterizzati da incertezza e di norma, prima di prendere delle decisioni, è opportuno ridurre tale incertezza predisponendo dei metodi per aumentare la nostra informazione. In tali situazioni la matematica ci mette a disposizione gli strumenti della statistica e del calcolo delle probabilità, al fine di esprimere, in base allo stato di conoscenza, le nostre valutazioni nei riguardi delle diverse ipotesi possibili. Lo strumento fondamentale per rappresentare in modo dinamico la conoscenza è quello della probabilità condizionata e del teorema di Bayes, che inquadra sul piano formale il meccanismo di "apprendimento dall'esperienza".

Il teorema di Bayes è lo strumento matematico con il quale, modificandosi lo stato di informazione, si calcolano le probabilità finali di determinate ipotesi, a partire da quelle iniziali delle stesse. Questa metodologia può essere sviluppata nella sua piena generalità se l'insegnamento della probabilità nella scuola viene basato sull'approccio soggettivo, secondo il quale, la probabilità di un evento, per un certo individuo, è la misura numerica del suo "grado di fiducia" nel verificarsi dell'evento stesso (cfr. de Finetti, 1970, Scozzafava, 1993). Tale grado di fiducia dipende strettamente dallo stato di informazione dell'individuo e si modifica sulla base di nuove informazioni. L'adozione della visione soggettiva nell'insegnamento della probabilità si rivela una scelta valida anche sotto

l'aspetto formativo, che è un obiettivo fondamentale della scuola (cfr. Gilio, 1990, 1993, Scozzafava, 1990).

In questo lavoro il meccanismo di apprendimento dall'esperienza viene illustrato applicando la formula di Bayes con riferimento al caso di un acquirente che deve decidere se accettare o rifiutare un lotto di articoli, che il fornitore garantisce essere tutti esenti da difetti. L'acquirente, sospettando che vi sia un articolo difettoso, prima di accettare il lotto vuole effettuare dei controlli in modo tale che, subordinatamente all'ipotesi che un certo numero di articoli scelti a caso siano risultati esenti da difetti, la probabilità che il lotto contenga un pezzo difettoso sia minore o uguale di un prefissato valore α , giudicato accettabile dall'acquirente.

Il lavoro include un'Appendice in cui è riportato un programma in Turbo Pascal 6.0 realizzato da Angelo Bossi, studente iscritto al secondo anno di Ingegneria Chimica dell'Università "La Sapienza" di Roma (a.a. 1992/93).

Tale programma permette di:

- 1) scegliere la dimensione del lotto e il valore della soglia probabilistica α ;
- 2) calcolare i valori e tracciare il diagramma delle funzioni che, al variare del numero di controlli effettuati, descrivono la relazione fra la probabilità finale e quella iniziale che nel lotto vi sia un articolo difettoso;
- 3) tabulare, in funzione del valore α e del numero di articoli da controllare, gli intervalli di valori di appartenenza della probabilità iniziale in corrispondenza dei quali la probabilità finale risulta minore o uguale ad α , determinando così l'accettazione del lotto da parte dell'acquirente.

L'uso di tali intervalli consente un approccio "qualitativo", nel senso che l'acquirente, per stabilire il numero di controlli da effettuare, non ha bisogno di specificare esattamente il valore della probabilità iniziale, ma soltanto l'intervallo di appartenenza.

2. FORMULAZIONE DEL PROBLEMA

Indichiamo con N la dimensione del lotto, cioè il numero totale di articoli, e con n il numero di controlli. Consideriamo gli eventi:

H = “il lotto contiene un articolo difettoso”,

H^c = “il lotto non contiene articoli difettosi”,

E_i = “l’ i -esimo articolo controllato è non difettoso”, $i = 1, 2, \dots, n$.

Posto $P(H) = p$, introduciamo le funzioni $f_1(p)$, $f_2(p)$, ..., $f_n(p)$ così definite:

$$f_k(p) = P(H | E_1 E_2 \dots E_k), \quad k = 1, 2, \dots, n,$$

dove il simbolo $E_1 E_2 \dots E_k$ rappresenta l’evento intersezione degli eventi $E_1, E_2, \dots, E_k$.

Il valore p è la probabilità iniziale secondo l’acquirente che sia vero l’evento H , mentre $f_k(p)$ rappresenta in funzione di p la probabilità finale di H dopo k controlli, cioè la probabilità di H condizionata all’ipotesi che siano stati controllati k articoli e che siano risultati tutti non difettosi.

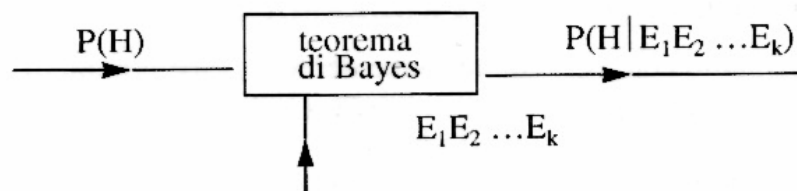
La regola in base alla quale l’acquirente decide di accettare o rifiutare il lotto è la seguente: procedendo con estrazioni a caso (senza restituzione) si effettuano controlli successivi su articoli del lotto. Il lotto viene rifiutato se uno degli articoli controllati risulta difettoso. Se, invece, gli articoli man mano esaminati risultano tutti non difettosi la procedura viene interrotta e il lotto viene accettato, in corrispondenza del valore n tale che:

$$f_n(p) = P(H | E_1 E_2 \dots E_n) \leq \alpha,$$

dove la soglia probabilistica α è scelta dall’acquirente.

Ovviamente, se $p \leq \alpha$ l’acquirente accetta il lotto senza effettuare controlli.

Il meccanismo di aggiornamento delle probabilità è rappresentato nello schema seguente:



Esso si basa sull'uso iterato della seguente formula:

$$P(H | E_1 E_2 \dots E_k) = \frac{P(H) P(E_1 E_2 \dots E_k | H)}{P(H) P(E_1 E_2 \dots E_k | H) + P(H^c) P(E_1 E_2 \dots E_k | H^c)},$$

che, applicata per $k = 1, 2, \dots, n$, fornisce ad ogni passo le probabilità finali dell'evento H date da:

$$P(H | E_1) = \frac{P(H) P(E_1 | H)}{P(H) P(E_1 | H) + P(H^c) P(E_1 | H^c)},$$

$$P(H | E_1 E_2) = \frac{P(H) P(E_1 E_2 | H)}{P(H) P(E_1 E_2 | H) + P(H^c) P(E_1 E_2 | H^c)},$$

.....

$$P(H | E_1 E_2 \dots E_n) = \frac{P(H) P(E_1 E_2 \dots E_n | H)}{P(H) P(E_1 E_2 \dots E_n | H) + P(H^c) P(E_1 E_2 \dots E_n | H^c)}.$$

Nel nostro caso si ha:

$$P(E_1 | H) = \frac{N-1}{N}, \quad P(E_1 E_2 | H) = \frac{N-1}{N} \frac{N-2}{N-1} = \frac{N-2}{N},$$

.....

$$P(E_1 E_2 \dots E_n | H) = \frac{N-1}{N} \frac{N-2}{N-1} \dots \frac{N-n}{N-n+1} = \frac{N-n}{N},$$

$$P(E_1 | H^c) = P(E_1 E_2 | H^c) = \dots P(E_1 E_2 \dots E_n | H^c) = 1,$$

da cui si ottiene:

$$P(H|E_1) = \frac{p \frac{N-1}{N}}{p \frac{N-1}{N} + (1-p)} = \frac{(N-1)p}{N-p} = f_1(p) ,$$

$$P(H|E_1 E_2) = \frac{p \frac{N-2}{N}}{p \frac{N-2}{N} + (1-p)} = \frac{(N-2)p}{N-2p} = f_2(p) ,$$

.....

$$P(H|E_1 E_2 \dots E_n) = \frac{p \frac{N-n}{N}}{p \frac{N-n}{N} + (1-p)} = \frac{(N-n)p}{N-np} = f_n(p) .$$

Nel caso in esame, la condizione $f_n(p) = P(H|E_1 E_2 \dots E_n) \leq \alpha$

equivale alla disuguaglianza $\frac{(N-n)p}{N-np} \leq \alpha$, che è soddisfatta se e

solo se $p \leq \frac{N\alpha}{N-n(1-\alpha)}$.

Ad esempio, per $N = 6$ si ha

$$P(H|E_1) = f_1(p) = \frac{5p}{6-p} , \quad P(H|E_1 E_2) = f_2(p) = \frac{4p}{6-2p} ,$$

$$\dots , \quad P(H|E_1 E_2 \dots E_5) = f_5(p) = \frac{p}{6-5p} .$$

Se l'acquirente giudica accettabile una soglia probabilistica

$\alpha = 0.2$, utilizzando la disuguaglianza $p \leq \frac{N\alpha}{N-n(1-\alpha)}$ con $N = 6$,

$\alpha = 0.2$, $n = 1, 2, \dots$, ai fini dell'accettazione del lotto sarà sufficiente un controllo se la probabilità iniziale p è tale che $f_1(p) \leq 0.2$,

cioè se $p \leq \frac{3}{13} \cong 0.2308$.

Se la probabilità iniziale p è maggiore di $\frac{3}{13}$ saranno sufficienti due controlli se risulta $f_2(p) \leq 0.2$, cioè se è soddisfatta la condizione $p \leq \frac{3}{11} \cong 0.2727$.

Nel caso $p > \frac{3}{11}$ la procedura termina dopo tre passi se si ha $f_3(p) \leq 0.2$, cioè se $p \leq \frac{1}{3} \cong 0.3333$.

Nel caso $p > \frac{1}{3}$ la procedura termina dopo quattro passi se $f_4(p) \leq 0.2$, cioè se $p \leq \frac{3}{7} \cong 0.4286$.

Nel caso $p > \frac{3}{7}$ la procedura termina dopo cinque passi se $f_5(p) \leq 0.2$, cioè se $p \leq \frac{3}{5} = 0.6$.

Osserviamo infine che nel caso $p > 0.6$ la condizione $f_n(p) \leq \alpha$ è soddisfatta solo per $n = N = 6$, cui corrisponde $f_6(p) = 0$. In questo caso infatti il “sospetto” iniziale dell’acquirente, rappresentato dalla probabilità iniziale $P(H) = p$, è troppo alto e, affinché la probabilità di H scenda al di sotto di α , è necessario controllare tutti gli articoli. Al contrario, nel caso $p \leq 0.2$ l’acquirente accetta il lotto senza effettuare controlli.

In questo esempio, quindi, l’intervallo $[0, 1]$ dei possibili valori di p viene suddiviso nei sottointervalli

$$I_0 = [0, 0.2], \quad I_1 = (0.2, \frac{3}{13}], \quad I_2 = (\frac{3}{13}, \frac{3}{11}],$$

$$I_3 = (\frac{3}{11}, \frac{1}{3}], \quad I_4 = (\frac{1}{3}, \frac{3}{7}], \quad I_5 = (\frac{3}{7}, \frac{3}{5}], \quad I_6 = (\frac{3}{5}, 1].$$

Il numero (minimo) di controlli che l’acquirente deve predisporre è pari all’intero n tale che $p \in I_n$.

In generale, per un lotto di N pezzi, utilizzando la disuguaglianza

$p \leq \frac{N\alpha}{N - n(1 - \alpha)}$ si ottengono i sottointervalli:

$$I_0 = [0, \alpha], \quad I_1 = (\alpha, \frac{N\alpha}{N - (1 - \alpha)}], \quad I_2 = (\frac{N\alpha}{N - (1 - \alpha)}, \frac{N\alpha}{N - 2(1 - \alpha)}],$$

$$\dots, I_n = (\frac{N\alpha}{N - (n - 1)(1 - \alpha)}, \frac{N\alpha}{N - n(1 - \alpha)}], \dots, I_N = (\frac{N\alpha}{N - (N - 1)(1 - \alpha)}, 1]$$

e si ha $f_n(p) \leq \alpha$ se $p \in I_n$.

3. ASPETTI GENERALI

Il problema fin qui esaminato costituisce una versione semplificata di una situazione più generale. Infatti abbiamo supposto che il lotto contiene al più un pezzo difettoso, cioè il numero aleatorio X di articoli difettosi nel lotto può assumere soltanto il valore 1 (corrispondente all'ipotesi H) oppure il valore 0 (corrispondente all'ipotesi H^c). In generale, però, i valori possibili di X sono gli interi da 0 fino ad N e, introdotti gli eventi

$$H_k = \text{"nel lotto vi sono } k \text{ pezzi difettosi"}, \quad k = 0, 1, \dots, N,$$

l'acquirente dovrebbe valutare le probabilità iniziali

$$P(H_k), \quad k = 0, 1, \dots, N$$

per poi calcolare, per $n = 1, 2, \dots$, le probabilità finali

$$P(H_k | E_1 E_2 \dots E_n), \quad k = 0, 1, \dots, N,$$

utilizzando la formula di Bayes.

La versione semplificata si può adottare nei casi in cui le ipotesi $H_2, H_3, \dots, H_N$ vengono trascurate perché si ritengono "molto piccole" le loro probabilità. In effetti in molte situazioni si ha $P(H_k) \cong 0$ per $k \geq 2$ e quindi $P(X \geq 1) \cong P(X = 1)$.

Si consideri ad esempio il caso in cui X rappresenta il numero di bombe nascoste nei bagagli di un aereo, oppure X rappresenta il numero di componenti guasti del motore di tale aereo, e così via.

Nel caso specifico, inoltre, la versione semplificata si può giustificare sulla base di un atteggiamento probabilistico ben preciso dell'acquirente. Infatti, la probabilità di trovare pezzi difettosi durante i controlli (e quindi di rifiutare il lotto) è tanto più alta quanti più pezzi difettosi vi sono nel lotto. Quindi, dal punto di vista dell'acquirente, è ragionevole ritenere che il fornitore, ammesso che abbia deciso di inserire qualche articolo difettoso nel lotto, ha in tal caso interesse ad inserirne un numero molto basso. Pertanto, sotto l'aspetto probabilistico, il modo migliore di premunirsi contro la presenza di articoli difettosi è quello di supporre che al più ve ne sia uno, il che corrisponde ad assumere:

$$P(X \geq 1) = P(X = 1), \quad \text{cioè} \quad P(X \geq 2) = 0.$$

Questo atteggiamento si può caratterizzare in modo formale. Infatti, posto:

$$P(X = 0) = 1 - p, \quad P(X = k) = \lambda_k, \quad k = 1, 2, \dots, N,$$

$$P(X \geq 1) = p = \lambda_1 + \lambda_2 + \dots + \lambda_N,$$

con p assegnato, si può verificare che, fra tutte le distribuzioni di probabilità iniziali del tipo

$$\pi = (1 - p, \lambda_1, \dots, \lambda_N),$$

quella che richiede il numero più grande di controlli, affinché sia soddisfatta la condizione (di accettazione del lotto)

$$P(X \geq 1 | E_1 E_2 \dots E_n) \leq \alpha,$$

è la distribuzione π_0 tale che $\lambda_1 = p, \lambda_2 = \dots = \lambda_N = 0$, cioè

$$\pi_0 = (1 - p, p, 0, \dots, 0),$$

per la quale si ha $P(X \geq 1) = P(X = 1)$.

Infatti, tenendo conto che $P(E_1 E_2 \dots E_n | X = 0) = 1$, si ha:

$$\begin{aligned} P(X = 0 | E_1 E_2 \dots E_n) &= \frac{P(X = 0) P(E_1 E_2 \dots E_n | X = 0)}{\sum_{k=0}^N P(X = k) P(E_1 E_2 \dots E_n | X = k)} = \\ &= \frac{(1 - p)}{(1 - p) + \sum_{k=1}^N \lambda_k P(E_1 E_2 \dots E_n | X = k)} . \end{aligned}$$

Osserviamo, inoltre che la probabilità $P(E_1 E_2 \dots E_n | X = k)$ è pari al rapporto $\binom{N-n}{k} / \binom{N}{k}$ se l'evento $E_1 E_2 \dots E_n$ è compatibile con l'evento $(X = k)$, mentre è nulla se tali eventi sono incompatibili. Pertanto, risulta

$P(E_1 E_2 \dots E_n | X = 1) \geq P(E_1 E_2 \dots E_n | X = 2) \geq \dots \geq P(E_1 E_2 \dots E_n | X = n)$,
da cui segue

$$\begin{aligned} P(X = 0 | E_1 E_2 \dots E_n) &\geq \frac{(1 - p)}{(1 - p) + \left(\sum_{k=1}^N \lambda_k \right) P(E_1 E_2 \dots E_n | X = 1)} = \\ &= \frac{(1 - p)}{(1 - p) + p \frac{N-n}{N}} = \frac{N(1 - p)}{N - np} \end{aligned}$$

e quindi

$$\begin{aligned} P(X \geq 1 | E_1 E_2 \dots E_n) &= 1 - P(X = 0 | E_1 E_2 \dots E_n) \leq \\ &\leq 1 - \frac{N(1 - p)}{N - np} = \frac{(N - n)p}{N - np} = P(H | E_1 E_2 \dots E_n) = f_n(p) . \end{aligned}$$

In conclusione, subordinatamente all'ipotesi che gli articoli man mano esaminati siano tutti esenti da difetti, la distribuzione di probabilità iniziale $\pi = (1 - p, \lambda_1, \dots, \lambda_N)$ che "ritarda" il più possibile l'accettazione del lotto, cioè per la quale la probabilità $P(X \geq 1 | E_1 E_2 \dots E_n)$ scende più lentamente al disotto della soglia α , è data da $\pi_0 = (1 - p, p, 0, \dots, 0)$.

Tale distribuzione concentra il valore $p = \lambda_1 + \lambda_2 + \dots + \lambda_N$ sull'evento $(X = 1)$, corrispondente all'ipotesi H nella versione semplificata del problema.

BIBLIOGRAFIA

- [1] DE FINETTI B. (1970), *Teoria delle probabilità*. Einaudi Editore, Torino.
- [2] GILIO A. (1990), *Aspetti algebrici e geometrici nell'insegnamento della probabilità soggettiva*, Archimede, 1, Genn. - Marzo, 32 - 41.
- [3] GILIO A. (1993), *Probabilità di eventi come previsioni di numeri aleatori*, Induzioni, 7, Luglio - Dicembre, 39 - 48.
- [4] SCOZZAFAVA R. (1990), *Per un insegnamento "fusionista" della probabilità e della statistica nelle scuole secondarie*, Matematica Oggi: dalle idee alla scuola (a cura di Fulvia Furinghetti), Cidi di Genova, Irrsae Liguria, Edizioni Scolastiche Bruno Mondadori, 251 - 256.
- [5] SCOZZAFAVA R. (1993), *La probabilità soggettiva e le sue applicazioni*, Ed. Veschi, Masson (Milano).

APPENDICE

Programma per calcolatore in Turbo Pascal 6.0

In questa Appendice viene riportato un programma in Turbo Pascal 6.0 realizzato dallo studente Angelo Bossi, che consente di analizzare il problema con l'aiuto del calcolatore e permette in particolare di:

- 1) scegliere il valore dei parametri N, α ;
- 2) calcolare i valori e tracciare il grafico delle funzioni $f_1(p), f_2(p), \dots, f_{N-1}(p)$;
- 3) determinare gli intervalli $I_1, I_2, \dots, I_n$.

Il programma consiste in una serie di procedure, alcune delle quali riguardano la presentazione grafica, mentre altre si riferiscono agli aspetti probabilistici e numerici del problema.

```
Program Tabula Guasti;
Uses Crt, Dos, Graph;
Const
  MaxNum=300;{Modificare questo valore per variare la dimensione dell'array}
  CoordX=60;
  CoordY=450;
  Triangle1: array[1..4] of PointType =
    ((x: 620; y: 450 ),(x: 600 ; y:440 ),
     (x: 600 ; y:460 ),(x: 620 ; y:450 ));
  Triangle2: array[1..4] of PointType =
    ((x: 60; y:0 ),(x: 50; y:20 ),
     (x: 70; y:20),(x: 60; y:0 ));
Var n,l,m,h,k,conta,x,alfa,ind,probind,g,inizio,fine,ind,i1,i2:Integer;
    GrDriver,GrMode,ErrCode:Integer;
    y1,alfa,y,prob:Real;
    tab:array[1..MaxNum] of Integer;
    ch,risposta,scelta,modo,ris:Char;
    label 1,2,3,4;
Procedure Intro; {Pagina introduttiva}
begin
  for conta:=1 to 8 do WriteLn;
  WriteLn('          Programmato in Turbo Pascal 6.0 da Angelo Bossi');
  for conta:=1 to 8 do WriteLn;
  WriteLn('          ');
  WriteLn('N.B.=E" possibile graficare le N-1 curve solo per N < ',MaxNum);
  WriteLn(' Per entrare in ambiente grafico e" necessaria una scheda');
  WriteLn(' grafica VGA o EGA');
```

```

WriteLn('          Premi <INVIO> per continuare...');
for conta:=1 to 3 do WriteLn;
TextColor(0);ReadLn;TextColor(15);
end;

Procedure Tabula;
{Questa procedura calcola le probabilita' al variare di alfa e n
e ne stampa il valore;per migliorare la leggibilita' della
tabella e' stato introdotta una visualizzazione a pagina}
label 1;
begin
1: ClrScr;
TextColor(10);
WriteLn(' 1) Per tutti i possibili valori della soglia');
WriteLn(' 2) Solo per un particolare valore della soglia');
WriteLn;TextColor(15);
Write(' Qual e" la tua scelta? ');
ReadLn(ris);
if ris='1' then begin i1:=1; i2:=19; end;
if ris='2' then begin
WriteLn;Write(' Immetti il valore della soglia:=');
ReadLn(alfa);
if (alfa<0) or (alfa>1) then goto 1;
i1:=1;
i2:=1;
end;
if (ris<>'1') and (ris<>'2') then goto 1;
l:=1;      {Inizializzazione per la tabulazione}
repeat
ClrScr;
TextColor(10);
Writeln('          TABULAZIONE SU ',n,' PEZZI');
if ris='1' then Writeln else for k:=1 to 10 do WriteLn;
TextColor(15);
Write(' ');
{Condizione per la tabulazione a pagine}
if (l+7-n)<0 then m:=l+7 else m:=n-1;
for conta:=l to m do begin
if conta<10 then Write(' ',conta)
else if conta<100 then Write(' ',conta)
else Write(' ',conta);
end;
TextColor(15);
Writeln;
{Questi 2 cicli FOR sono la parte centrale di questa procedura;
tramite una ricorsione infatti viene calcolato ,memorizzato
in un array e poi stampato il valore di alfa e di n}
for k:=i1 to i2 do begin
if ris='1' then alfa:=k/20;
g:=Round(alfa*100); {normalizzazione di alfa}
if g<10 then Write('0.0',g,'; ') else Write('0.',g,'; ');
for h:=1 to m do begin
tab[h]:=Round((10000*alfa*n)/((n-h)+alfa*h));
if tab[h]<1000 then Write('0.0',tab[h],'; ');

```

```
        else Write('0.',tab[h],'; ');
    end;
    Writeln;
end;
if ris='2' then for k:=1 to 10 do Writeln;
Write('Premi <INVIO>');
TextColor(0);
ReadLn;
l:=m+1;
until l>n-1; {Condizione per la fine della tabulazione a pagine}
TextColor(10);
Write('                Fine Tabulazione');
Sound(880);Delay(150);NoSound;
ReadLn;
end;
```

Procedure Tabula1;

{Questa procedura calcola la F(p) al variare della probabilità
e del numero dei componenti del lotto e ne stampa il valore;
per migliorare la leggibilità della tabella e' stata introdotta
una visualizzazione a pagina}
label 1;
begin
1: ClrScr;
TextColor(10);
Writeln(' 1) Per tutti i possibili valori della probabilità');
Writeln(' 2) Solo per un particolare valore della probabilità');
Writeln;TextColor(15);
Write(' Qual e' la tua scelta? ');
ReadLn(ris);
if ris='1' then begin i1:=1; i2:=19; end;
if ris='2' then begin
Writeln;Write(' Immetti il valore della probabilità:');
ReadLn(prob);
if (prob<0) or (prob>1) then goto 1;
i1:=1;
i2:=1;
end;
if (ris<>'1') and (ris<>'2') then goto 1;
l:=1; {Inizializzazione per la tabulazione}
repeat
ClrScr;
TextColor(10);
Writeln(' TABULAZIONE SU ',n,' PEZZI');
if ris='1' then Writeln else for k:=1 to 10 do Writeln;
TextColor(15);
{Condizione per la tabulazione a pagine}
if (l+7-n)<0 then m:=l+7 else m:=n-1;
if m<100 then Write(' p ') else Write(' p ');
for conta:=1 to m do begin
if conta<10 then Write(' F',conta,'(p)')
else if conta<100 then Write(' F',conta,'(p)')
else Write(' F',conta,'(p)');
end;
end;

```

TextColor(15);
Writeln;
for k:=i1 to i2 do begin
  if ris='1' then prob:=k/20;
  g:=Round(prob*100); {normalizzazione della probabilita'}
  if g<10 then Write('0.0',g,' '); else Write('0.',g,' ');
  for h:=1 to m do begin
    tab[h]:=Round((10000*prob*(n-h))/(n-prob*h));
    if tab[h]<10 then Write('0.000',tab[h],'; ');
    else if tab[h]<100 then Write('0.00',tab[h],'; ');
    else if tab[h]<1000 then Write('0.0',tab[h],'; ');
    else Write('0.',tab[h],'; ');
  end;
  Writeln;
end;
if ris='2' then for k:=1 to 10 do Writeln;
Write('Premi <INVIO>');
TextColor(0);
ReadLn;
l:=m+1;
until l>n-1; {Condizione per la fine della tabulazione a pagine}
TextColor(10);
Write('                Fine Tabulazione');
Sound(880);Delay(150);NoSound;
ReadLn;
end;

```

Procedure Curve;

{I 2 cicli FOR ,come nella procedura TABULA, sono qui usati
per calcolare le coordinate(esprese in pixel) delle curve}

begin

```

  for h:=inizio to fine do begin
    SetColor(h+8);SetLineStyle(0,0,1);
    if (h mod 8)=0 then SetColor(h+1);
    MoveTo(CoordX,CoordY); {sposto il cursore nell'origine degli assi}
    for x:=1 to 100 do begin
      y:=((n-h)*x/100)/(n-(h*x/100));
      LineTo(CoordX+x*5,Round(CoordY-(y*400)));
    end;
    SetColor(15);
    SetLineStyle(1,0,1);
  end;

```

{Tracciano l'indicatore di alfa e le relative probabilita' e viceversa}

```

  if ch='1' then begin
    alfaind:=60+Round((500*alfa*n)/((n-h)+alfa*h));
    Line(60,GetMaxY-Round(alfa*400)-29,
      alfaind,GetMaxY-Round(alfa*400)-29);
    Line(alfaind,450,alfaind,GetMaxY-Round(alfa*400)-29);
  end;
  if ch='2' then begin
    y1:=((n-h)*prob)/(n-(h*prob));
    probind:=GetMaxY-Round((y1*400)+28);
    if h=inizio then ind:=probind;
    Line(Round((prob*500)+60),450,Round((prob*500)+60),ind);
    Line(60,probind-1,Round((prob*500)+60),probind-1);
  end;

```

```

    end;
  end;
end;
Procedure Grafica;
{Questa procedura, sfruttando i dati di TABULA o TABULA1 permette la
visualizzazione grafica delle curve rappresentanti le probabilita'
in un sistema cartesiano(F(p),P)}
label 2,4;
begin
  TextColor(15);
  WriteLn;
4: if ch='1' then begin
  Write(' Inserisci la soglia probabilistica alfa:=');
  ReadLn(alfa);
  if (alfa>1) and (alfa<0) then goto 4;
end else begin
Write(' Inserisci la probabilita":=');
  ReadLn(prob);
  if (prob>1) and (prob<0) then goto 4;
end;
ClrScr;
GrDriver := Detect;
InitGraph(GrDriver,GrMode,""); {Inizializzazione grafica}
ErrCode := GraphResult;
if ErrCode = grOk then
begin
  if (GrDriver<>9) and (GrDriver<>4) then goto 2;
  SetColor(15);
  SetLineStyle(0,0,1);
  Line(30,450,600,450);
  Line(60,20,60,480); {Assi Cartesiani}
  Line(60,50,560,50);
  Line(560,50,560,450);
  FillPoly(SizeOf(Triangle1) div SizeOf(PointType), Triangle1);
  FillPoly(SizeOf(Triangle2) div SizeOf(PointType), Triangle2);
{Scala di valori}
  for conta:=1 to 20 do
    Line(12+conta*25+35,448,12+conta*25+35,452);
  for conta:=1 to 20 do
    Line(conta*25+35,445,conta*25+35,455);
  for conta:=1 to 11 do
    Line(conta*50+10,442,conta*50+10,458);
  for conta:=1 to 20 do
    Line(58,GetMaxY-(10+conta*20+9),62,GetMaxY-(10+conta*20+9));
  for conta:=1 to 20 do
    Line(55,GetMaxY-(conta*20+9),65,GetMaxY-(conta*20+9));
  for conta:=1 to 11 do
    Line(52,GetMaxY-(conta*40-11),68,GetMaxY-(conta*40-11));
  SetColor(10);
  SetTextStyle(2,0,6);OutTextXY(25,460,'0.0');
  OutTextXY(100,460,'0.1');OutTextXY(25,400,'0.1');
  OutTextXY(150,460,'0.2');OutTextXY(25,360,'0.2');
  OutTextXY(200,460,'0.3');OutTextXY(25,320,'0.3');
  OutTextXY(250,460,'0.4');OutTextXY(25,280,'0.4');

```

```

    OutTextXY(300,460,'0.5');OutTextXY(25,240,'0.5');
    OutTextXY(350,460,'0.6');OutTextXY(25,200,'0.6');
    OutTextXY(400,460,'0.7');OutTextXY(25,160,'0.7');
    OutTextXY(450,460,'0.8');OutTextXY(25,120,'0.8');
    OutTextXY(500,460,'0.9');OutTextXY(25, 80,'0.9');
    OutTextXY(550,460,'1.0');OutTextXY(25, 40,'1.0');
    SetColor(14);SetTextStyle(1,0,2);
    OutTextXY(10,0,'F(p)');OutTextXY(600,457,'P');
    Curve;      {Richiamo la procedure che traccia le curve}
    SetLineStyle(0,0,1);
    Line(60,50,560,50);
    Line(560,50,560,450);
    SetColor(14);SetTextStyle(2,0,6);
    OutTextXY(250,5,'Premi <INVIO>');
    Sound(880);Delay(150);NoSound;
    TextColor(0);
    ReadLn;
2:   CloseGraph;
    end;
end;

{PROGRAMMA PRINCIPALE}
begin
    ClrScr;
    TextBackGround(0);
    TextColor(15);
    Intro;
1:   ClrScr;
    TextColor(15);
    Write(' Inserisci il numero dei componenti del lotto :=');
    ReadLn(n);
    if (n<2) or (n>MaxNum) then goto 1;
4:   ClrScr;
    TextColor(10);
    WriteLn(' 1) Tabulazione');
    WriteLn(' 2) Grafico');
    WriteLn(' 3) Modifica i valori');
    WriteLn(' 4) Uscita');
    TextColor(15);
    WriteLn;
    Write(' Qual e'' la tua scelta? ');
    ReadLn(scelta);
    if scelta='1' then goto 2 else
        if scelta='2' then goto 3 else
            if scelta='3' then goto 1 else
                if scelta='4' then begin ClrScr;Exit; end else goto 4;
    Sound(831);Delay(150);NoSound;
2:   ClrScr;
    TextColor(10);
    WriteLn('1) Calcolo della probabilita'' data la soglia probabilistica');
    WriteLn('2) Calcolo della F(p) data la probabilita''');
    WriteLn('3) Uscita');
    TextColor(15);
    WriteLn;

```

```
Write('Qual e'' la tua scelta? ');
ReadLn(modo);
if modo='1' then Tabula else
  if modo='2' then Tabula1 else
    if modo='3' then begin ClrScr;Exit; end else goto 2;
  goto 4;
3: ClrScr;
  TextColor(10);
  WriteLn(' 1) Indicatore di F(p)');
  WriteLn(' 2) Indicatore di p');
  WriteLn;
  TextColor(15);
  Write(' Qula e'' la tua scelta ? ');
  ReadLn(ch);
  if (ch<>'1') and (ch<>'2') then goto 3;
  ClrScr;
  TextColor(10);
  WriteLn(' 1) Grafico di tutte le curve');
  WriteLn(' 2) Grafico di un sottinsieme di tutte le curve');
  WriteLn(' 3) Grafico di una sola curva');
  WriteLn(' 4) Uscita');
  WriteLn;
  WriteLn;TextColor(15);
  Write(' Qual e'' la tua scelta ? ');
  ReadLn(risposta);
  if risposta='1' then begin inizio:=1; fine:=n-1; Grafica; end;
  if risposta='2' then begin
    WriteLn;
    Write(' Componente iniziale del lotto:=');
    ReadLn(inizio);
    if (inizio<0) or (inizio>n-1) then goto 3;
    WriteLn;
    Write(' Componente finale del lotto:=');
    ReadLn(fine);
    if (fine>n-1) or (fine<inizio) then goto 3;
    Grafica;
    end;
  if risposta='3' then begin
    WriteLn;
    Write(' Componente del lotto:=');
    ReadLn(inizio);
    if (inizio<0) or (inizio>n-1) then goto 3;
    fine:=inizio;
    Grafica;
    end;
  if risposta='4' then begin
    ClrScr;
    exit;
  end;
  if (risposta<>'1') and (risposta<>'2') and
    (risposta<>'3') and (risposta<>'4') then goto 3;
  goto 4;
end.
```

A REAL LIFE APPROACH TO THE TEACHING OF PROBABILITY (*)

Romano Scozzafava
Dipartimento Metodi e Modelli Matematici
Università "La Sapienza" - Roma (Italy)

What are the challenges and perspectives for probability and statistics education, apart from the need of providing a professional tool for working on practical problems? Since classical mathematics is still a source of intriguing and stimulating ideas, it seems inadequate to subscribe to the conception that the role of probability and statistics is that of making pupils acquainted with the existence of new fundamental topics besides classical mathematical theories. In other words, we maintain that the main role of this teaching is essentially that of introducing a *new way of thinking* (i.e. inductive, to handle random situations) more than that of acquiring some interesting and less usual mathematical notions or that of emphasizing a utilitarian point of view. The conceptual framework of inductive reasoning, which is fundamental for the scientific knowledge, lies in the possibility of measuring the expectation of "future" events on the basis of observed "past" events (usually, statistical data). In the various real-life situations in which uncertainty is present, gathering and interpreting statistical data leads in general to a decrease of uncertainty with respect to the initial situation: the measurement of this uncertainty can be quantitatively carried out by the same tool used for the measurement of the uncertainty associated with random phenomena such as coin tossing or dice throwing, i.e. through the concept of *probability*. The outlined process may be called, in a vague but expressive way, "learning from experience." Statistics and probability proceed so at the same pace, the former providing techniques for the normalization and the synthesis of data, the latter interpreting them through "conclusions" which in general *are not certain* (as those of ordinary, i.e. deductive, logic), but only

(*) Contributed paper at ICME-7, Québec, 1992.

more or less probable. In order to fully grasp the richness of this merged approach to probability and statistics, an overcoming of barriers created by prevailing opinions is needed: for example, many elementary approaches to these subjects rely essentially upon a “combinatorial” *assessment of probability* (assuming equal probability of all possible cases) and upon the possibility of introducing the *probability of an event* through the *frequency observed in the past* for other events that are considered, in a sense, “equal” to that of interest. Yet it is not generally underlined that the choice of these events (as the choice, in the combinatorial approach, of the outcomes for which equal probability is assumed) is necessarily *subjective*.

For example, an insurance company that needs to evaluate the probability of dying within the year of a given person can base its assessment on data referring to individuals of the same town (or region, or district) as the given person, or of the same age, or sex, or civil status, or of the same profession, or income, or having an analogous bodily constitution (height, weight, etc.), and so on, *grouping in many different ways some or all of the preceeding characteristics, or possibly others*, and to each of these (subjective) choices there corresponds in general a *different frequency*.

In other words, it is essential to give up any artful limitation to particular events (not even clearly definable) and try to ascribe to probability a more general meaning, which after all should be a sensible way to cope with real situations: in fact a concrete and not stereotyped teaching should start from the subjective and intuitive “real life” meaning of probability as *degree of belief* in the occurrence of an event. Last but not least, *from a didactic point of view* the introduction, through a natural condition of *coherence*, of *subjective probability* is very simple and quick. It is in fact easy to show (as we are going to do in what follows) that subjective probability satisfies the usual and classical properties: it is a function whose range is *between zero and one*, these two extreme values being taken, in particular, by the *impossible* and *certain* event respectively, and which is *additive* for mutually exclusive events. These properties constitute the starting point in the axiomatic

approach: they are well-known for the two most “popular” cases, i.e. when the event of interest is embedded in a set of possible outcomes which are considered, in some (necessarily subjective) sense, “symmetric”, or when the event is seen as a further repetition, under “similar” (subjective judgement!) conditions, of a given phenomenon.

An *event* E is any unambiguous proposition that can only assume the two “values” TRUE or FALSE (denoted by 1 or 0, if E is regarded as a simple *random variable*). The lack of information on the actual value of E paves the way to the introduction, as an ersatz, of the concept of probability.

A value $p = P(E)$ (or pS , for arbitrary $S \neq 0$) is regarded as an amount to be paid to bet on E , with the proviso of winning a unit amount (or an amount S) of money if E occurs and nothing if E does not occur (in other words, p is the amount to be paid to get an amount *equal to the value assumed by* E): *coherence* is defined by the requirement that *the choice of* p *would not make the player a sure loser or winner*. If E is different from Ω and $\emptyset$ (certain and impossible event), the two possible “gains” are

$$G_1 = (-p + 1)S \quad (\text{if } E \text{ occurs}),$$

$$G_2 = -pS \quad (\text{if } E \text{ does not occur}),$$

and so, since coherence requires that they must not be both negative or both positive, p must satisfy the inequality $-p(1-p)S^2 < 0$, which is, for any $S \neq 0$, the same as

$$0 < p < 1.$$

When $E = \Omega$ or $E = \emptyset$ there is *no uncertainty* on the outcome of the corresponding bet: the *only* (certain!) value of the gain is

$$G(\Omega) = (-p + 1)S \quad \text{or} \quad G(\emptyset) = -pS$$

respectively, and so coherence requires that the gain is *equal to zero*, which gives $p = 1$ for $E = \Omega$ and $p = 0$ for $E = \emptyset$. In conclusion,

if the subjective probability of E (our *degree of belief* on E) is defined as an amount $p = P(E)$ that makes coherent a bet on E , then

$$(1) \quad 0 \leq P(E) \leq 1,$$

with, in particular,

$$(2) \quad P(\Omega) = 1, \quad P(\emptyset) = 0.$$

Now, given a finite *partition* of $\Omega = E_1 \vee E_2 \vee \dots \vee E_n$ and n simultaneous bets, let $P(E_i)$ be the amount paid for a coherent bet on E_i (with $i = 1, 2, \dots, n$): clearly, these n bets can be regarded as a single bet on Ω with amount $P(E_1) + P(E_2) + \dots + P(E_n)$, and so the first of (2) implies

$$(3) \quad P(E_1) + P(E_2) + \dots + P(E_n) = 1.$$

So the usual “axioms” (1), (2), (3) of probability are easily obtained in a very simple way, as necessary and sufficient conditions for coherence. Notice that this result is based on *hypothetical* bets: the force of the argument *does not depend on whether or not one actually intends to bet*, since a method of assessing probabilities making one a sure loser or winner *if he had to gamble* (whether or not he really will act so) would be suspicious and unreliable for any purposes whatsoever.

The “combinatorial” and “frequentist” methods of evaluation of probabilities can be easily embedded into the general concept of subjective probability. In fact, given n possible outcomes represented by the events $E_1, E_2, \dots, E_n$ of a partition of Ω , and an event E which is a union of r among the E_i ’s, the classical evaluation $P(E) = r/n$ follows easily from (3) (and from its extension to the case in which the events are still mutually exclusive but not necessarily constituting a partition) *through the subjective opinion that a symmetry exists and that it implies equality of probabilities*, namely $P(E_i) = 1/n$. Moreover, for a sequence $A_1, A_2, \dots, A_{n+m}$ of events (“trials” of a given

phenomenon), assume that the outcome of the first m is known (i. e., the corresponding “past” frequency X is, say, k/m) and consider the “future” frequency

$$y = \frac{A_{m+1} + A_{m+2} + \dots + A_{m+n}}{n}.$$

The *prevision* of a random variable like Y can be interpreted *similarly to the probability of an event*: $\mathcal{P}(Y)$ is the amount that must be paid (in n simultaneous coherent bets on the events $A_{m+1}, A_{m+2}, \dots, A_{m+n}$ of amounts $P(A_i)S$, where $S = 1/n$) to get an amount *equal to the value that will assume* Y , i.e.

$$\mathcal{P}(Y) = \frac{P(A_{m+1}) + P(A_{m+2}) + \dots + P(A_{m+n})}{n}.$$

If the above events are judged *equally probable* (subjective opinion!) we get, denoting by p this common probability

$$(4) \quad \mathcal{P}(Y) = p.$$

Assuming (subjective opinion!) that *the probability distribution of the “future” frequency Y is equal to that of the “past” frequency X* , whose value is known and equal to k/m (so that $\mathcal{P}(X) = k/m$), from (4) it follows that $p = k/m$, i.e. the “frequentist” evaluation of $P(A_i)$ for $i \geq m+1$.

It is important to point out that this approach puts in the right perspective all the subjective aspects hidden in the so-called “objectivistic theories”. Probability cannot be simply seen as a “physical” property of an event and may be considered also for “unique” situations, going beyond the two particular cases corresponding to “*equally probable outcomes*” and “*events that can be repeated under similar conditions*”.

Let us now face the problem of choosing, for each event E , a suitable value of its probability p *among all coherent evaluations*, i.e. those satisfying (1), (2), (3): it is not enough directing our attention only toward the event E , but we need taking into

account also *other events* which contribute in determining our information on E . To this end two fundamental tools are *conditional probability and Bayes' theorem*.

A bet on a *conditional event* $E|H$, with $H \neq \emptyset$, is a bet on E which is *called off* when H does not occur: so, if an amount p is paid to bet on $E|H$, we get, when H turns out to be true, an amount 1 if E is true and an amount 0 if E is false, and we get back the amount p if H turns out to be false. Then, putting $p = P(E|H)$, if such a bet is *coherent* the function

$$P_H(\cdot) = P(\cdot | H)$$

verifies (1), (2), (3) for any given $H \neq \emptyset$. In this case $P(E|H)$ is said *conditional probability* (of E given H). Recall that, for *unconditional bets*, p is the amount to be paid to get an amount equal to the value assumed by E (i.e. 1 or 0), or, more generally, the *prevision* $\mathcal{P}(X)$ of a random variable X is the amount to be paid to get an amount equal to the value that will assume X . It follows that $E|H$ can be regarded as a *three-valued event*, or else as a random variable X assuming one of the three values 1, 0, p , corresponding respectively to *won, lost, called off* bet, according to whether $H = E = 1$, or $H = 1$ and $E = 0$, or $H = 0$, that is

$$X = E H + p(1 - H).$$

Then the prevision of X must coincide with $p = P(E|H)$, i.e.

$$(5) \quad p = \mathcal{P}(X) = P(E \wedge H) + p - p P(H).$$

From (5) it follows easily

$$(6) \quad P(E \wedge H) = P(H) P(E|H).$$

Notice that this fundamental (and well-known) relationship involving conditional probability has been deduced, like (1), (2), (3), *assuming only coherence* (i.e. the obvious condition that no bet is sensible if one knows in advance to be a sure loser or winner!); moreover, *no restriction has been imposed on the kind of events involved* (such as equally probable, or “repeatable”, or anything else).

By re-writing eq. (6) with H^c (the *opposite* of H) in place of H and summing (6) itself to the relation so obtained, we get, since $(E \wedge H) \vee (E \wedge H^c) = E$,

$$(7) \quad P(E) = P(H) P(E|H) + P(H^c) P(E|H^c).$$

Interchanging the role of E and H in (6), we may eliminate the left-hand side from these two equations; then, assuming that $P(E) > 0$ and recalling (7), we obtain

$$(8) \quad P(H|E) = \frac{P(H) P(E|H)}{P(H) P(E|H) + P(H^c) P(E|H^c)}.$$

This formula constitutes *Bayes' theorem*, that allows, given an event E (representing, for example, observed data), evaluating the conditional probability $P(H|E)$ of an other event H (often called *hypothesis*), which *initially* (i.e., before collecting data) had been assessed equal to $P(H)$. In other words, the true problem for any event H may be considered *not* simply that of evaluating $P(H)$, which may be possibly regarded just as a provisional assessment (a *prior*, in a more technical jargon), but that of evaluating $P(H|E)$ taking into account all the relevant information carried by some other event E .

So a “merging” of probability and statistics can be easily attained without being entangled with involved or complicated arguments. Thanks to the conditional probability $P(H|E)$ it is possible to give a different *probability evaluation* of H for each different “state of information” expressed by E (usually corresponding to *statistical data*): almost all Bayesian statistical procedures are, essentially, extensions of this way of thinking. Moreover, Bayes' theorem may be applied repeatedly, taking into account new statistical data and assuming now $P(H|E)$ (i.e., the *final* probability “of yesterday”) as the *initial* probability “of today”. And so on: there is no distinction at all between initial (or *prior*) and final (or *posterior*) probabilities, except that *they refer to a different state of information*.

In the so-called “ortodox” approach to statistics, there is no possibility to speak of “probable” conclusions concerning events such as H , just because this would involve probability as “degree of belief” and not as a frequency. So the attempt is made to introduce concepts such as that of *accepting* or *rejecting* a given hypothesis H , on the basis of the conditional probability $P(E|H)$ only: for example, if this probability (in the sense of observed frequency in the past) is “very large” (i.e. E is highly probable *given* H) and E in fact occurs, this may be considered as a good reason for accepting H . Notice that this argument corresponds to *ignore* not only *other* hypotheses but also the *initial probability* $P(H)$; moreover, the acceptance of the hypothesis H is not based on any judgement of plausibility, but it is a sort of a mechanical act, based on the frequency of success of the method on which the acceptance criterion is based. Accepting (or rejecting) a hypothesis is a concept that after all is alien both to the *logic of certainty* (i.e., the usual, ordinary logic) and to probability theory. Nevertheless *an interpretation in Bayesian terms is possible* assuming *equal* initial probabilities of all possible hypotheses and also when an experiment can be performed a large number of times, so that the observed data do influence the evaluation of probability in a prevailing way as compared to initial probabilities: of course, in both circumstances the statement “*accepting a hypothesis*” may be ... accepted only as a way of saying that the given hypothesis has a “very large” subjective probability.

ALCUNI SPUNTI SULL'INSEGNAMENTO DELLA MATEMATICA CON PARTICOLARE RIFERIMENTO AD ASPETTI ELEMENTARI DELLA NOZIONE DI CONTINUITÀ IN ANALISI MATEMATICA

Domenico Lenzi
Dipartimento di Matematica
Università di Ancona

Questo vuole essere un intervento in chiave didattica e trova la sua ragione nelle difficoltà che molte persone hanno nello studio della matematica.

Tali difficoltà, che a volte sono dovute a motivi che un pò semplicisticamente vengono espressi con la locuzione "mancanza di predisposizione", spesso sono conseguenza di cause del tutto diverse tra cui il fatto che il linguaggio tipico della matematica non sempre è in grado di esprimere in maniera sufficientemente comprensiva e penetrante i concetti trattati. Ciò non per inadeguatezza del linguaggio stesso ma perchè esso non riesce a trasmettere, quando non è ben padroneggiato, tutte le sfumature di cui la matematica è ricca e che una persona normale dovrebbe teoricamente essere in grado di recepire. In definitiva ci sembra che si verifichi un fenomeno simile a quello che si ha anche con la lingua italiana. Infatti c'è chi a volte, volendo dare più incisività a un suo messaggio, è costretto a ricorrere al dialetto.

Diamo un esempio di "prosa" tratto da un lavoro scientifico di Analisi Matematica:

Sia $f: \mathcal{R} \rightarrow \mathcal{R}$ una funzione, $x_0 \in \mathcal{R}$ e $P_0 = (x_0, f(x_0))$, poniamo, per $x \in \mathcal{R}$ ed $a \in \mathcal{R}$:

- i) $P(x, \alpha) = (x, f(x_0) + \alpha(x - x_0))$
- ii) $G_f(x) = \begin{cases} \{(y, f(y)) : x_0 \leq y \leq x\} & \text{se } x_0 \leq y \leq x \\ \{(y, f(y)) : x \leq y \leq x_0\} & \text{se } x \leq y \leq x_0 \end{cases}$

Lo scritto sarebbe senz'altro più incisivo se, pur conservando le formule i) e ii), $P(x, \alpha)$ fosse preannunciato come il punto di ascissa x che è situato sulla retta passante per $(x_0, f(x_0))$ ed avente coefficiente angolare α . Inoltre sarebbe bene preannunciare ciascuno dei due insiemi di ii) come quella parte del grafico della funzione f che è compresa tra i punti $(x_0, f(x_0))$ e $(x, f(x))$.

Un'altra situazione significativa può essere ricavata dall'algebra. Infatti, se $(\mathcal{A}, +, \cdot)$ è un anello ed a, b, c , sono tre elementi distinti di $\mathcal{A}$, possiamo dire che essi costituiscono una "terna di Steiner" se godono della seguente proprietà: comunque si prendano due elementi distinti tra a, b, c , il loro prodotto o l'opposto di questo coincide col terzo elemento. Questa proprietà tradotta in "matematiche" diventa:

$$\begin{aligned} \text{iii)} \quad & \forall x, y \in \{a, b, c\}: x \neq y \Rightarrow \\ & \Rightarrow (x \cdot y \in \{a, b, c\} - \{x, y\} \text{ o } -x \cdot y \in \{a, b, c\} - \{x, y\}). \end{aligned}$$

Tuttavia la iii), pur presentando lo stesso concetto è meno espressiva.

Certo qualcuno potrebbe dire che il problema della comprensione della matematica non è così grave, e tutto sommato potrebbe risolversi facendo studiare questa disciplina solo chi non vive in un rapporto conflittuale con essa. Ed uno studioso eminente come S. Banach era proprio di questo avviso (si veda [P]).

Tuttavia pensiamo che un atteggiamento del genere sia riduttivo. Infatti una limitazione nello studio della matematica priverebbe i nostri ragazzi di uno strumento formidabile di promozione culturale e, soprattutto, di formazione sul piano della razionalità. In tale ottica la "dimensione matematica" è una delle acquisizioni fondamentali dell'uomo e relegare questa materia ad un ruolo secondario, riservandola solo a pochi eletti, ricalcherebbe quell'errore che la nostra scuola ha già fatto con l'abolizione del latino. Per cui le difficoltà incontrate da molti studenti in quest'altra disciplina, che potevano essere un sintomo da valutare attentamente, sono state rimosse semplicemente eliminandone il veicolo e non – come sarebbe stato opportuno – eliminandone le cause.

Inoltre, un'educazione matematica opportunamente realizzata avrebbe un'importanza fondamentale non solo sul piano sociale, ma anche sul piano pratico, contribuendo ad attenuare quella "mortalità" dei nostri studenti di ingegneria, che ha reso l'Italia uno dei paesi in cui il bisogno di ingegneri si fa maggiormente sentire.

Ma come fare per operare un cambiamento efficace? Non è facile dare una risposta esauriente. Comunque, oltre ad eliminare per quanto è possibile gli impedimenti di tipo linguistico, quello che bisogna fare senz'altro è rimuovere tutte le difficoltà di carattere mnemonico (1) che spesso assillano i nostri studenti, trasferendo ad essi un'abitudine ormai diffusa anche nell'ambito dei docenti. Infatti, è difficile trovare uno studioso che nel corso di una conferenza non si aiuti nell'esposizione con degli appunti o con lucidi e lavagna luminosa. E non si capisce perchè questa pratica non debba essere consentita ad uno studente, almeno parzialmente, anche in sede di esami. Quello che importa è che gli abbia chiare le linee generali dell'argomento trattato, riesca a capire e apprezzare una dimostrazione, comprendendo i vari passaggi e il ruolo che le varie ipotesi svolgono in essa. In definitiva, riesca a percepire il palpito che una pagina di matematica emana.

E affinché questo avvenga è necessario "tornare a parlare" di matematica con i nostri studenti, cercando di spiegare le ragioni di questa disciplina e le linee di pensiero che essa spesso nasconde, evidenziando i collegamenti con la realtà (se ci sono (2)) o illustrando i motivi di impostazioni che appaiono in contrasto con una certa nostra maniera di percepire la realtà stessa (cfr [M]).

Non vogliamo dilungarci in esempi, anche perchè ciò che interessa è soprattutto smuovere le acque. Ci limitiamo a ricordare qualche situazione, senza pretendere che la nostra risposta sia la più adeguata sul piano metodologico.

Si pensi alla continuità di una funzione reale f definita su di un intervallo chiuso $[a, b]$. Quasi tutti i testi universitari del primo anno dicono che f è continua in $[a, b]$ se essa è continua in ogni punto dell'intervallo in questione.

-
1. Siamo convinti del fatto che sia importante educare la memoria, tuttavia ciò dovrebbe essere attuato con interventi ad hoc.
 2. Si pensi all'algebra Lineare, disciplina perfetta nel suo rigore e nella sua semplicità, che nei corsi universitari del primo anno ha finito, purtroppo, col camuffare la vecchia geometria.

Così, come di frequente avviene in matematica, si esprime in termini “locali” una proprietà di tipo “globale”. Però, purtroppo, quella proprietà locale non sembra tradurre in maniera adeguata una certa visualizzazione intuitiva, che spesso si cerca di dare della nozione di funzione continua in un intervallo chiuso dicendo che essa esprime l’idea mentale che ci si fa di una curva “materiale” tracciata su di un foglio (da sinistra verso destra o da destra verso sinistra) senza sollevare la penna. O meglio, mentre ci si può render conto facilmente che “l’idealizzazione” di una tale curva gode di quella proprietà locale, solo una mente dotata di notevoli doti speculative riesce a capire agevolmente che quella proprietà, estesa a tutti i punti dell’intervallo $[a, b]$, può – ma non sempre, come si vedrà – far sparire come per incanto alcune situazioni “patologiche”.

È facile dare un esempio elementare di tali situazioni. Basta immaginare di associare ad ogni numero reale x compreso tra 0 e 1 un numero reale compreso tra 0 e lo stesso x . Si viene così ad avere una funzione f continua in 0. Infatti essa assume in 0 il valore 0 e, grazie al “teorema dei carabinieri”, ha 0 come limite per x che tende a 0, poichè risulta “imprigionata” tra la funzione nulla e la funzione $y = x$ (queste ultime funzioni hanno chiaramente come limite 0, per x tendente a 0).

Orbene, in una ideale rappresentazione grafica quella funzione apparirebbe, salvo casi particolari, come uno sfarfallio di punti che pur si “addensano” intorno al punto $(0, 0)$. Per cui può risultare difficoltoso afferrare che, nel caso in cui la continuità riguardi tutti i punti dell’intervallo di definizione, quel turbinio di punti si acqueti come per incanto.

Sul piano didattico il problema si potrebbe risolvere considerando dapprima il caso delle funzioni monotone. Ad esempio, se f è una funzione crescente in un intervallo chiuso $[a, b]$, allora è abbastanza intuitivo il fatto che la continuità si esprima attraverso la proprietà che la f assuma in $[a, b]$ tutti i valori compresi tra $f(a)$ ed $f(b)$. Poichè si può provare agevolmente che per funzioni reali monotone quest’ultima proprietà è equivalente a quella della continuità in ogni punto di $[a, b]$, è chiaro che essa può essere un “cavallo di Troia” utile alla penetrazione nelle difese psicologiche dello

studente, onde egli possa “accettare” l'usuale definizione di continuità. Tanto più che molte delle funzioni reali continue definite in un intervallo chiuso $[a, b]$ si possono ripartire in funzioni monotone definite in un numero finito di sottointervalli chiusi di $[a, b]$.

Si potrebbero dare esempi di funzioni continue in $[a, b]$ che non sono monotone in alcun intervallo incluso in $[a, b]$; tuttavia quelli che io conosco richiedono dimostrazioni un pò pesanti.

Situazioni eccezionali più tenui possono essere dovute anche alla presenza in $[a, b]$ di alcuni punti particolari, ad esempio punti in cui la “sfortunata” funzione non sia nè crescente, nè decrescente, non abbia un minimo relativo e non abbia un massimo relativo (“i punti critici”).

Di questi punti critici è bene dare esempi significativi, ma senza ricorrere a costruzioni difficoltose e dense di complicazioni che mettano in ombra il fatto essenziale che si vuol presentare. Spesso una descrizione discorsiva e snella, di cui si riesca ad intravedere la traduzione in “matematiche” (si veda (G), §0, n°2), può risultare più efficace sul piano didattico.

Un caso interessante è quello della funzione così definita nell'intervallo chiuso $[0, 1]$. Si prendano sull'asse delle ascisse i punti di ascissa $1, 1/2, 1/4, \dots, 1/2^n, \dots$; quindi si considerino tutte le circonferenze che hanno rispettivamente come diametro il segmento compreso tra due di quei punti che nella data successione sono “vicini” (cioè, l'uno immediato successivo dell'altro). Dopodichè si consideri di ogni circonferenza, alternativamente al decrescere del diametro, la semicirconferenza i cui punti hanno coordinate non negative oppure quella i cui punti hanno coordinate non positive, iniziando dalla semicirconferenza del primo tipo.

Poichè quelle semicirconferenze esprimono delle funzioni continue che “si saldano” negli estremi degli intervalli chiusi in cui sono definite, allora la funzione che nasce da questa saldatura è continua nell'intervallo semiaperto a sinistra $(0, 1]$; inoltre essa ha chiaramente 0 come limite al tendere a 0 della variabile, per cui può essere univocamente prolungata in una funzione continua f definita su tutto l'intervallo chiuso $[0, 1]$; infatti, basta far corrispondere a 0 il valore 0. Ebbene, 0 è chiaramente per la funzione f uno dei punti critici di cui si parlava.

Una persona che sentisse del disagio a liquidare con disinvoltura la continuità di quelle “semicirconferenze” (che in altri momenti sarà bene evidenziare rigorosamente) potrebbe sostituire ad ogni semicirconferenza il triangolo equilatero costruito sul suo diametro, che sta dalla stessa parte della semicirconferenza (rispetto all’asse delle ascisse). Anzi, la funzione costruita con i triangoli equilateri dà luogo a tanti altri casi significativi, purchè si immagini di ruotare il suo grafico intorno all’origine degli assi cartesiani, avendo l’accortezza di non far perdere ad esso il carattere di grafico di una funzione. E per questo basta che l’angolo di rotazione sia inferiore in valore assoluto a 30 gradi.

Invece chi pensasse che una matematica senza formule sia meno importante (tuttavia noi non siamo d’accordo con questo giudizio) può considerare la ben nota funzione che assume valore $x \sin(1/x)$ quando x è diverso da zero ed assume valore 0 in 0. Anche in tale caso è immediato rendersi conto che 0 è un punto critico per la funzione data.

Richiamiamo l’attenzione su di un’altra proprietà che caratterizza la continuità delle funzioni monotone non decrescenti definite su di un intervallo. I libri che abbiamo consultato non la riportano, forse perchè può essere considerata patrimonio del nostro folklore matematico. Tuttavia ci sembra che didatticamente valga la pena citarla, per un’analogia che c’è con un altro modo di vedere la continuità in una situazione del tutto diversa.

Si pensi all’usuale rappresentazione grafica di una relazione d’ordine su di un insieme finito $\mathcal{S}$ (“diagramma di Hasse”) e ad un percorso “in salita” su tale rappresentazione, che vada da un elemento p ad un elemento q (con $p < q$). È naturale dire che questo percorso è continuo quando esso avviene “senza salti” rispetto ad $\mathcal{S}$. Cioè, due elementi del percorso che risultino l’uno il successivo immediato dell’altro nel percorso sono tali anche sul diagramma. In sintesi: in $\mathcal{S}$ non c’è alcun elemento intermedio rispetto ad essi. Cioè, come si suol dire, quel percorso esprime una catena massimale del sottoinsieme di $\mathcal{S}$ costituito dagli elementi compresi tra p e q .

Ebbene, si verifica subito che condizione necessaria e sufficien-

te perchè una funzione reale non decrescente f definita nell'intervallo chiuso (a, b) sia continua è che il suo grafico sia una catena massimale rispetto all'ordinamento definito in $\mathcal{R} \times \mathcal{R}$ ponendo $(x, y) \leq (z, t)$ se e solo se $x \leq z$ e $y \leq t$; purchè ci si restringa al sottoinsieme di $\mathcal{R} \times \mathcal{R}$ costituito dagli elementi compresi tra $(a, f(a))$ e $(b, f(b))$. Un discorso analogo vale, *mutatis mutandis*, se f è non crescente.

Concludiamo dando un piccolo esempio volto a sottolineare l'opportunità di una presa di coscienza da parte dello studente del ruolo che ipotesi e proprietà varie possono giocare nell'ambito di un teorema. Si pensi alla caratterizzazione (legata alla nozione di catena massimale) delle funzioni reali continue e non decrescenti richiamata poco fa. Se percorriamo una qualsiasi delle sue facili dimostrazioni, ci rendiamo conto agevolmente che la completezza di $\mathcal{R}$ non gioca in essa alcun ruolo, per cui la dimostrazione può essere automaticamente trasportata al caso delle funzioni non decrescenti assumenti valori razionali e definite su intervalli razionali chiusi, avendo avuto cura di trasferire "pari pari" ai numeri razionali tutti i concetti necessari per parlare di limiti e continuità. Però, purtroppo, non si può dire altrettanto della caratterizzazione classica ricordata ancor prima. Infatti, una funzione non decrescente f definita sull'intervallo razionale chiuso $[a, b] \cap \mathcal{Q}$ (dove $\mathcal{Q}$ è l'insieme dei numeri razionali e $a, b \in \mathcal{Q}$) ed assumente tutti i valori razionali compresi tra i due numeri razionali $f(a)$ e $f(b)$ è senz'altro una funzione continua in ogni punto di $[a, b] \cap \mathcal{Q}$. Tuttavia questa affermazione non può essere invertita, potendosi dare molti controesempi. Ci limitiamo a considerare quello che si ottiene dalla funzione che sui numeri reali strettamente minori di $\sqrt{2}$ assume valore 0 e sugli altri numeri reali assume valore 1. Tale funzione è non decrescente, inoltre essa è continua su ogni numero reale diverso da $\sqrt{2}$, per cui è continua su tutti i numeri razionali, dove assume valori razionali. Ne consegue che la sua restrizione f a $[0, 2] \cap \mathcal{Q}$ è una funzione continua e non decrescente, tuttavia f non assume tutti i valori dell'intervallo razionale $[f(0), f(2)] \cap \mathcal{Q} (= [0, 1] \cap \mathcal{Q})$, dal momento che $1/2$ è un valore razionale che non è assunto da essa.

Quanto detto or ora induce a pensare che forse sarebbe più aderente all'idea intuitiva di continuità una definizione restrittiva che non sia di tipo topologico, ma sia legata semplicemente alla nozione di ordinamento totale. In definitiva, dato un insieme totalmente ordinato $(\mathcal{S}, \leq)$ e dati due elementi $a, b \in \mathcal{S}$, con $a < b$, si potrebbe dire che una funzione f definita sull'“intervallo” $[a, b]$ ($= \{x \in \mathcal{S} : a \leq x \leq b\}$) e assumente valori in $\mathcal{S}$ è “monotonamente continua” se essa è monotona ed assume tutti i valori compresi tra $f(a)$ e $f(b)$. Più in generale, f potrebbe dirsi “continua” in $[a, b]$ se $[a, b]$ si può suddividere in un numero finito di intervalli chiusi, che a due a due si intersecano al più in un estremo, sui quali la restrizione di f è monotonamente continua.

Sarebbe interessante vedere se anche con questo tipo di definizione compaiono situazioni paradossali. È chiaro che sulla base della definizione data or ora la funzione definita su $[0, 2] \cap \mathcal{Q}$, che abbiamo visto poco fa, non sarebbe più continua.

Le considerazioni svolte precedentemente fanno capire come uno studio efficace dei numeri reali e delle funzioni reali, essenziale per un'educazione matematica completa, non possa prescindere da un esame approfondito di situazioni riguardanti sottinsiemi significativi dell'insieme dei numeri reali. Io penso che esso (ma non solo esso) sia molto più illuminante e formativo di tanti complicati, assillanti ed estenuanti esercizi che, come forche caudine, aspettano al varco degli esami scritti i nostri poveri studenti, simili alle famigerate espressioni frazionarie a più piani che tutti deprechiamo, e che contribuiscono a dare un'immagine arcigna della matematica. Un'immagine che ognuno di noi dovrebbe sentirsi impegnato a cancellare.

BIBLIOGRAFIA

- [G] R. GODEMENT *Cours d'algèbre*, Hermann, Paris (1973).
- [M] C.F. MANARA *I frattali: capitolo nuovo della geometria*, Mathesis, Sez. di Varese (1990/91).
- [P] G. PRODI *Insegnamento secondario e insegnamento universitario della matematica*, Archimede, n. 4 (1989).

NUMERI LATERALI E DIAGONALI

Silvio Maracchia

Dipartimento di Matematica - Università "La Sapienza" - Roma

1. PREMESSA

I matematici greci divisero i numeri in molti modi, alcuni dei quali ancora usati da noi (numeri pari, dispari, quadrati, primi, triangolari, ecc.), altri noti soltanto a studiosi di teoria dei numeri (numeri perfetti, amici, abbondanti, ecc.), altri ancora caduti in disuso e noti solo a studiosi di storia della matematica (numeri imparipari, parimente pari, poligonal, eteromechi, solidi, ecc.).

A quest'ultima categoria appartengono i cosiddetti *numeri diagonali e laterali* di probabile origine pitagorica ma comunque sicuramente noti ai tempi di Platone, "numeri - scrive Teone di Smirne - che armonizzano le figure" (1).

2. NUMERI DIAGONALI E LATERALI

È proprio Teone di Smirne (I metà del secondo sec. d.C.) che ci consente di conoscere questo risultato della matematica greca (2) per il fatto che Platone aveva considerato come approssimazione del rapporto tra diagonale e lato di un quadrato il valore $7/5$ (3).

Teone scrive infatti l'opera *Esposizione delle conoscenze matematiche utili per la lettura di Platone*, un'opera che, pur giunta gravemente mutilata, ci permette di ottenere varie notizie sulla matematica e sull'astronomia.

-
1. Teone di Smirne, *Esposizione delle conoscenze matematiche utili per la lettura di Platone*, ed. curata da J. Dupuis, Hachette, Paris, 1892, cap. XXXI, p. 72. A questa edizione si farà riferimento per le citazioni successive.
 2. Anche Giamblico (*In Nicom.* Pistelli 91, 3 sgg.) e Proclo (V sec. d.C.) parlano dei numeri diagonali e laterali; ed è quest'ultimo che nel suo commento alla *Repubblica* di Platone, attribuisce ai Pitagorici la loro individuazione affermando, inoltre, che la prova geometrica della loro formazione si trova nel secondo libro degli *Elementi* di Euclide (vedremo in seguito questa "prova" di Euclide). P. Tannery ha mostrato come tali numeri si sarebbero potuti ottenere dalla media armonica studiata appunto dai Pitagorici (per tutte queste notizie e per i vari riferimenti cfr. *De Pythagore à Euclide* di P. H. Michel, Les Belles Lettres, Paris, 1950, 1950, pp. 427-441).
 3. Cfr. Platone, *Repubblica*, 546 c.

Ebbene, i numeri diagonali e laterali vengono trattati, scrive Teone, poiché in grado di approssimare sempre più il rapporto detto tra diagonale e lato di un quadrato.

Bisogna cominciare, scrive ancora Teone, dall'unità sia per i numeri diagonali che per i numeri laterali poiché l'unità è il principio di tutto e quindi è in potenza sia la diagonale e sia il lato. Successivamente, al primo numero diagonale stabilito ($d_1=1$) vanno aggiunti due laterali per ottenere il nuovo diagonale ($d_2 = 1 + 2 \cdot 1$) e al primo numero laterale ($l_1=1$) va aggiunto il primo diagonale ($l_2 = 1 + 1$). E così via, ottenendo pertanto le due successioni:

$$(1) \quad \begin{array}{rcccccc} d_n) & 1 & 3 & 7 & 17 & 41 \dots \\ l_n) & 1 & 2 & 5 & 12 & 29 \dots \end{array}$$

con la legge di formazione:

$$(2) \quad \begin{array}{ll} d_{k+1} = d_k + 2l_k & d_1 = 1 \\ l_{k+1} = l_k + d_k & l_1 = 1 \end{array}$$

Teone non scrive simbolicamente le formule (1) e (2), ma dopo aver calcolato alcune coppie delle successioni, conclude affermando: “*E così di seguito continuando l'addizione*”. Egli osserva inoltre che la “proporzione è alterna” intendendo con questo che il quadrato del numero diagonale è alternativamente maggiore o minore di una unità del doppio del quadrato del corrispondente laterale. Tradotto in simboli possiamo scrivere:

$$(3) \quad d_k^2 = 2 l_k^2 + (-1)^k$$

3. PLATONE E I NUMERI DIAGONALI E LATERALI

Prima di ulteriori osservazioni sui numeri diagonali e laterali, vediamo che cosa ha provocato l'esposizione di Teone e cioè vediamo che cosa dice Platone a tale riguardo.

Ebbene, Platone sta descrivendo nella *Repubblica* il cosiddetto “numero nuziale” (4) che presiede alla generazione umana di cui è il periodo e che definisce in cinque modi diversi.

4. Platone, *Repubblica*, 546 b-d; per varie notizie su tale argomento cfr. *Platone e la matematica nel mondo antico* di A. Frajese, Univers. Studium, Roma, 1963, pp. 145-150.

Le interpretazioni di queste definizioni non sono univoche; noi seguiremo quelle di A. Diés (5) che sembrano le più attendibili e portano, come dev'essere, tutte allo stesso numero e giustificano inoltre la presenza dei numeri diagonali e laterali in Teone.

Il numero in questione è 12.960.000 e uno dei modi per raggiungerlo viene indicato da Platone come "composto parte da fattori uguali parte da fattori disuguali: di cento quadrati delle *diagonali razionali* di cinque, ciascuno diminuito di uno, o di cento quadrati delle diagonali irrazionali diminuiti di due, e di cento cubi di tre" (6).

Si tratta di due metodi assai simili: nel primo si considera $100(7^2 - 1)$ e nel secondo $100[(\sqrt{50})^2 - 2]$, valori da moltiplicare successivamente per "cento cubi di tre", cioè per $100 \cdot 3^3$.

Ebbene, nel primo metodo Platone indica il sette quale "diagonale razionale di cinque", intendendo cioè che, rimanendo nell'ambito delle grandezze razionali, il valore approssimato della diagonale di un quadrato di lato cinque è appunto sette (7). Ma la relazione sette-cinque è proprio una di quelle che si trovano tra due particolari numeri diagonali e laterali corrispondenti ($d_3 = 7$; $l_3 = 5$).

4. DIMOSTRAZIONI E IPOTESI SUI NUMERI DIAGONALI E LATERALI

Possiamo ora dimostrare la proprietà (3) indicata da Teone deducendola dalla definizione (2) mediante il procedimento di induzione.

5. A. Diés, *Le nombre de Platon. Essai d'exégèse et d'histoire* (Acc. des Inscript. et belles lettres t.XIV, Paris, 1936); citazione tratta dall'opera di A. Frajese citata nella nota precedente.

6. Cfr. l'op.cit. di A. Frajese, p. 146 (la sottolineatura è nostra). Due altri metodi portano al calcolo di $(3 \cdot 4 \cdot 5)^4$ espresso in due modi diversi.

7. Come viene mostrato nel *Menone* (82a - 85b), Platone è a conoscenza che, indicando con d ed l rispettivamente diagonale e lato di uno stesso quadrato (o le loro misure), si ha $d^2 = 2 l^2$, per cui la (vera) lunghezza della diagonale del quadrato di lato 5 risulta da $d^2 = 2 \cdot 5^2 = 50$ [egli però non può scrivere, come abbiamo fatto noi $(\sqrt{50})^2 = 50$].

Vedremo anche come si possono ottenere i numeri diagonali e laterali attraverso l'algoritmo delle frazioni continue. Dopo questo vedremo che la (3) non è che un caso particolare di una famosa equazione di analisi indeterminata che ebbe una notevole importanza nello sviluppo della teoria dei numeri.

Metodo di induzione

La proprietà (3) è vera per $k = 1$, infatti, essendo $d_1 = l_1 = 1$, si ha:

$$d_1^2 = 2l_1^2 - 1 = 2l_1^2 + (-1)^1$$

Se la proprietà (3) si suppone vera per $k = n$ e cioè se si suppone che

$$d_n^2 = 2l_n^2 + (-1)^n$$

oppure

$$d_n^2 - 2l_n^2 = (-1)^n$$

essa si può dimostrare per $k = n+1$ e cioè che

$$d_{n+1}^2 = 2l_{n+1}^2 + (-1)^{n+1}$$

Infatti, per le (2) si ha:

$$d_{n+1}^2 = (d_n + 2l_n)^2 = d_n^2 + 4d_n l_n + 4l_n^2$$

e

$$2l_{n+1}^2 = 2(d_n + l_n)^2 = 2d_n^2 + 4d_n l_n + 2l_n^2$$

pertanto

$$d_{n+1}^2 = 2l_{n+1}^2 - d_n^2 + 2l_n^2$$

e quindi

$$d_{n+1}^2 = 2l_{n+1}^2 - (d_n^2 - 2l_n^2) = 2l_{n+1}^2 - (-1)^n = 2l_{n+1}^2 + (-1)^{n+1}$$

Notiamo che le (2) consentono di osservare che le due successioni sono entrambe divergenti ($d_{k+1} > d_k$ e $l_{k+1} > l_k$ con d_k e l_k numeri naturali), per cui dalla (3) si ottiene:

$$\left(\frac{d_k}{l_k}\right)^2 = 2 + \frac{(-1)^k}{l_k^2}$$

e quindi

$$\lim_{k \rightarrow +\infty} \left(\frac{d_k}{l_k}\right)^2 = 2 \quad \text{e} \quad \lim_{k \rightarrow +\infty} \frac{d_k}{l_k} = \sqrt{2}$$

Anche senza la teoria dei limiti, ai matematici greci era possibile verificare che i vari rapporti d_k/l_k riuscivano ad approssimare sempre meglio il rapporto tra diagonale e lato del quadrato, poiché

è sufficiente verificare che il rapporto dei rispettivi quadrati si avvicina sempre più a due.

Ad esempio (ritrovando l'alternanza delle approssimazioni):

$$\begin{array}{lll} \frac{3^2}{2^2} = \frac{9}{4} & \text{e si ha} & \frac{9}{4} - 2 = \frac{1}{4} \\ \frac{7^2}{5^2} = \frac{49}{25} & \text{e si ha} & 2 - \frac{49}{25} = \frac{1}{25} \\ \frac{17^2}{12^2} = \frac{289}{144} & \text{e si ha} & \frac{289}{144} - 2 = \frac{1}{144} \end{array}$$

e così via (8).

Dalla teoria delle frazioni continue sappiamo che un radicale quadratico $a \pm \sqrt{b}$ dà origine ad una frazione continua illimitata periodica, semplice o mista. Nel caso di $\sqrt{2}$ si ha:

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \dots}}}$$

Indicate con $R_1 = 1$; $R_2 = 1 + 1/2$; $R_3 = 1 + \frac{1}{2 + \frac{1}{2}}$ ecc. le

“ridotte” di tale sviluppo e con P_1, Q_1 ; P_2, Q_2 ; P_3, Q_3 , ecc. i corrispondenti numeratori e denominatori delle ridotte, si può notare che essi sono, nell'ordine, proprio i numeri diagonali e laterali:

$$\begin{array}{lll} R_1 = 1 & P_1 = 1 & Q_1 = 1 \\ R_2 = 1 + \frac{1}{2} = \frac{3}{2} & P_2 = 3 & Q_2 = 2 \\ R_3 = 1 + \frac{1}{1 + \frac{1}{2}} = \frac{7}{5} & P_3 = 7 & Q_3 = 5 \quad \text{ecc.} \end{array}$$

8. Notiamo che il limite calcolato risulterebbe sempre $\sqrt{2}$ anche se anziché ± 1 si avesse nella (3) una qualsiasi altra costante.

D'altra parte, la legge di formazione dei numeratori e denominatori delle ridotte del generico sviluppo

$$a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4 + \dots}}}$$

è la seguente:

$$(4) \quad \begin{cases} P_n = a_n P_{n-1} + P_{n-2} \\ Q_n = a_n Q_{n-1} + Q_{n-2} \end{cases} \quad n = 3, 4, 5, \dots$$

ed essa consente, nel caso particolare di $\sqrt{2}$, di calcolare ugualmente i numeri diagonali e laterali tenendo conto che in questo caso $P_1 = 1$, $Q_1 = 1$; $P_2 = 3$, $Q_2 = 2$ e che $a_n = 2$ (con $n > 1$).

Infatti:

$$\begin{cases} P_3 = a_3 P_2 + P_1 = 2 \cdot 3 + 1 = 7 \\ Q_3 = a_3 Q_2 + Q_1 = 2 \cdot 2 + 1 = 5 \\ P_4 = a_4 P_3 + P_2 = 2 \cdot 7 + 3 = 17 \\ Q_4 = a_4 Q_3 + Q_2 = 2 \cdot 5 + 2 = 12 \\ P_5 = a_5 P_4 + P_3 = 2 \cdot 17 + 7 = 41 \\ Q_5 = a_5 Q_4 + Q_3 = 2 \cdot 12 + 5 = 29 \end{cases} \quad \text{ecc.}$$

Il fatto che i numeri diagonali e laterali risultano numeratori e denominatori delle ridotte dello sviluppo in frazioni continue di $\sqrt{2}$, ha fatto avanzare l'ipotesi che i matematici greci conoscessero questo sviluppo. D'altra parte l'algoritmo euclideo, inerente al calcolo del massimo comun divisore di due numeri, avrebbe potuto portare allo sviluppo in frazione continua (limitata) di un numero razionale (9) (se poi due grandezze sono tra loro incommensurabili il procedimento a loro applicato non avrebbe avuto termine (10)).

9. Per queste considerazioni rimandiamo, ad esempio, alle *Frazioni continue* di C. D. Olds, Zanichelli, Bologna 1968, pp. 24-25.

10. Cfr. P.H. Michel, *De Pythagore à Euclide*, op. cit., pp. 432-433; cfr. anche A. Frajese, *Attraverso la storia della Matematica*, Le Monnier, Firenze, 1969, p. 247 passim.

Infine, ricordano i sostenitori di questa ipotesi, anche le approssimazioni di $\sqrt[3]{3}$ operate da Archimede sono alcune ridotte del suo sviluppo in frazione continua.

Notiamo una volta per tutte che i matematici greci non avevano cognizione dei numeri irrazionali, pertanto, anziché parlare di $\sqrt{2}$ o $\sqrt{3}$ e così via, essi consideravano i rapporti tra grandezze che risultano uguali a tali valori. Oppure consideravano i quadrati delle grandezze in questione i cui rapporti erano dunque uguali a 2 o a 3 e così via. Per inciso notiamo che fu questa la risposta alla scoperta di grandezze incommensurabili.

C'è da osservare però che solo col senno del poi possiamo dedurre dall'algoritmo euclideo il procedimento di frazione continua visto che non abbiamo alcuna indicazione esplicita. Allorché le due grandezze prese in esame sono poi tra loro incommensurabili, il loro rapporto non può essere un numero (11) poiché i matematici greci, non avendo, come si è detto, il concetto di numero irrazionale, non potevano che considerare numeri interi o razionali con i quali è possibile rappresentare solo grandezze tra loro commensurabili. Anzi, la possibilità di un procedimento infinito, è usato da Euclide per stabilire l'incommensurabilità di due grandezze, ma la dimostrazione (*Elementi* X, 2) è del tutto teorica e non prende in considerazione le varie misure delle successive grandezze sempre più piccole che vi si trovano.

Riguardo alle ridotte di $\sqrt{2}$ abbiamo già visto, inoltre, che vengono calcolate in modo differente da quello che verrebbe dalla consapevolezza del suo sviluppo in frazione continua (12).

Ma l'argomento a nostro parere più forte contro l'ipotesi che i matematici greci potessero conoscere questo sviluppo, è proprio quello che coinvolge le ridotte considerate da Archimede per approssimare $\sqrt[3]{3}$.

11. Si veda a questo proposito un significativo brano della *Metafisica* di Aristotele (1021 a, 3-7) il cui esame si trova in *Aristotele e l'incommensurabilità* di S. Maracchia, *Archive for History of Exact Sciences*, 21, 1980, pp. 213 sgg.

12. C'è da dire però che, fornendo le stesse successioni di numeri, la (2) e la (4) risultano sostanzialmente equivalenti. È possibile infatti da una ottenere l'altra.

Infatti, come è noto, e come osserva anche Teone a proposito dei numeri diagonali e laterali, le ridotte individuano alternativamente valori per difetto e per eccesso del rapporto tra diagonale e lato di un quadrato (13). Dovendo quindi calcolare un valore per difetto e uno per eccesso di $\sqrt{3}$, Archimede, se fosse stato a conoscenza del suo sviluppo in frazione continua, avrebbe considerato due ridotte successive e non, come fa, due ridotte tra loro non consecutive (14).

Ritorniamo però alla formula (3) $d_k^2 = 2l_k^2 + (-1)^k$ descritta da Teone allorché osserva che, poiché $d^2 = 2l^2$ (essendo d ed l diagonale e lato di uno stesso quadrato), con i numeri laterali e diagonali si ha un'alternanza:

$$\text{a volte } d_k^2 > l_k^2 \text{ essendo } d_k^2 = 2l_k^2 + 1$$

$$\text{a volte } d_k^2 < l_k^2 \text{ essendo } d_k^2 = 2l_k^2 - 1.$$

Questa osservazione potrebbe dare la chiave della formazione di tali numeri: poiché $d^2 = 2l^2$ ma non è possibile trovare due numeri interi soddisfacenti tale relazione (15), si sono cercati numeri interi che si avvicinassero il più possibile a questa relazione, cioè numeri che risolvessero, in definitiva, equazioni del tipo:

$$(5) \quad x^2 = 2y^2 + 1 \quad \quad x^2 = 2y^2 - 1$$

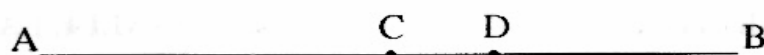
e da ciò la determinazione per tentativi di alcune coppie di numeri (diagonali e laterali) risolvendo tali equazioni e, successivamente, la determinazione della legge di formazione (2) che porta appunto ad approssimazioni sempre migliori di $\sqrt{2}$.

13. Giustificheremo tra poco questa osservazione.

14. Archimede considera $\frac{265}{153} < \sqrt{3} < \frac{1351}{780}$ ma il valore per difetto è la nona ridotta e quello per eccesso la dodicesima.

15. Questo è noto alla matematica greca poiché è nota l'incommensurabilità delle due grandezze.

Si trovano negli *Elementi* di Euclide, come ricorda Proclo, due proposizioni (II, 9 e II, 10) nelle quali si riesce a dimostrare geometricamente che per i due segmenti AD e AC della figura, con C punto medio di un segmento AB e



D un punto qualsiasi di CB (16), per i quali evidentemente:

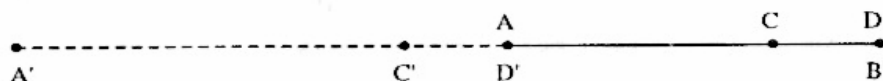
$$(6) \quad \begin{cases} AD = 2CD + DB \\ AC = CD + DB \end{cases}$$

[relazioni simili a quelle indicate dalle (2)], si ha:

$$(7) \quad AD^2 - 2AC^2 = 2CD^2 - DB^2$$

La (7) mostra che, a prescindere dal segno, la differenza $2CD^2 - DB^2$ rimane invariata se si costruiscono i due segmenti corrispondenti AC e AD mediante le (6). Si può dunque proseguire costruendo altri due segmenti A'D' e A'C' a partire da AD e AC che si trovino nella stessa relazione indicata dalle (6).

Ebbene, con $C'D' = AC$ e $A'C' = C'B'$ ($D' = A$ e $B' = D$) si ha infatti:



$$(6') \quad \begin{cases} A'D' = 2C'D' + D'B' = 2AC + AD \\ A'C' = C'D' + D'B' = AC + AD \end{cases}$$

per cui anche in questo caso, con la stessa dimostrazione geometrica di Euclide, si ottiene:

$$(7') \quad A'D'^2 - 2A'C'^2 = 2AC^2 - AD^2$$

e così via per nuovi punti A'', D'', C'', ecc.

In tal modo, a partire da $CD = DB = 1$, si ottengono coppie di segmenti AD, AC; A'D', A'C', ecc. le cui lunghezze sono pertanto quelle dei numeri diagonali e laterali che conosciamo.

16. Questa è la proposizione II, 9; nella proposizione successiva il punto D è preso sul prolungamento di AB dalla parte di B.

In ogni caso, come è stato già osservato, il rapporto dei quadrati delle coppie dei segmenti ottenuti (AD^2 / AC^2 , $A'D'^2 / A'C'^2$, $A''D''^2 / A''C''^2$, ecc.) tenderebbe a due.

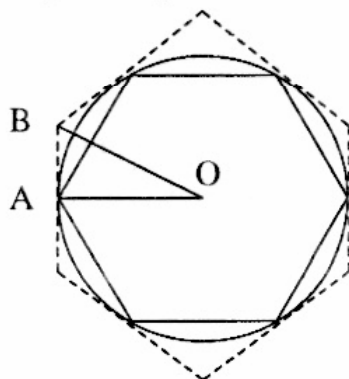
5. PRIMA ESTENSIONE DEI NUMERI DIAGONALI E LATERALI

Analogamente a quanto è stato fatto per l'approssimazione di $\sqrt{2}$, allorché il rapporto tra due grandezze m ed n è, come scriveremo noi, $\sqrt{3}$ cioè se si ha $m^2 = 3n^2$, è possibile seguire una strada simile a quella vista.

Osserviamo intanto che un tale rapporto venne ottenuto da Archimede nell'opera *Misura del cerchio* allorché, considerando poligoni regolari inscritti e circoscritti al cerchio, ottenne la famosa relazione:

$$3 + \frac{10}{71} < \pi < 3 + \frac{10}{70} \quad (\text{cioè } \pi = 3,14 \dots)$$

Partendo dall'esagono, egli si ferma a poligoni di 96 lati. Ebbene,



essendo AB metà del lato dell'esagono circoscritto si ha: $OA/AB = \sqrt{3}$.

Non seguiremo il procedimento di Archimede, notiamo solo che, essendo $\angle AOB = 30^\circ$ segue $OA/AB = \cotg 30^\circ = \sqrt{3}$.

Come sappiamo, Archimede usa poi le approssimazioni notevoli:

$$\frac{265}{153} < \sqrt{3} < \frac{1351}{780} \quad (17)$$

17. Per il procedimento di Archimede e per le approssimazioni indicate cfr. A. Frajese, *Attraverso la storia della matematica* op.cit., pp. 285 sgg.

Notiamo, però, che anche nel *Teeteto* di Platone si trova la circostanza che il lato del quadrato di area 3 (e quindi, scriveremmo noi, il lato di misura $\sqrt{3}$) non è commensurabile con il lato del quadrato di lato 1 e cioè non il segmento lungo 1(18).

Ma, tornando all'analogia con l'approssimazione di $\sqrt{2}$, non essendo possibile risolvere l'equazione $m^2 = 3n^2$ con numeri interi, si saranno cercate coppie di numeri che si avvicinassero ad essa col risolvere equazioni del tipo:

$$(8) \quad x^2 = 3y^2 \pm 1 \quad x^2 = 3y^2 \pm 2$$

analoghe alla (3).

Anche in questo caso è possibile trovare facilmente per tentativi coppie di numeri risolvienti le (8) o alcune di esse e risalire poi da tali soluzioni a delle leggi di formazione del tipo delle (2).

Si ottiene pertanto:

$$(9) \quad \begin{aligned} m_{k+1} &= 2m_k + 3n_k \\ n_{k+1} &= m_k + 2n_k \end{aligned}$$

che forniscono, a partire dalla soluzione $m_1 = 1, n_1 = 1$ dell'equazione $x^2 = 3y^2 - 2$, le successioni

$$(10) \quad \begin{array}{rcccl} m_n & 1 & 5 & 19 & 71 & \dots \\ n_n & 1 & 3 & 11 & 41 & \dots \end{array}$$

della stessa equazione, e, a partire dalla soluzione $m_1 = 2, n_1 = 1$ dell'equazione $x^2 = 3y^2 + 1$, porta alle successioni

$$(10') \quad \begin{array}{rcccl} m_n & 2 & 7 & 26 & 97 & \dots \\ n_n & 1 & 4 & 15 & 56 & \dots \end{array}$$

6. EQUAZIONE DI PELL-EULERO

Le equazioni

$$\begin{aligned} d_k^2 &= 21_k^2 + (-1)^k \\ x^2 &= 3y^2 \pm 1 \end{aligned}$$

sono casi particolari dell'equazione

$$x^2 = Ny^2 \pm 1 \quad (N \text{ numero intero non quadrato})$$

che consentono di approssimare la $\sqrt{N}$.

18. V. Platone, *Teeteto*, 147c sgg.

Ebbene, nel caso ancora più particolare:

$$(11) \quad x^2 = Ny^2 + 1 \text{ (oppure } x^2 - Ny^2 = 1)$$

essa si trova nel famoso *Problema dei buoi* enunciato da Archimede con $N = 4.729.494$. La più piccola soluzione del problema di Archimede porta comunque a numeri di oltre 200.000 cifre (19).

Questa equazione (11), da risolversi con numeri interi non nulli, prende il nome improprio di “*equazione di Pell-Eulero*” e fu studiata per secoli dai più grandi matematici.

Nella presunzione dell’esistenza di una soluzione, essa fu risolta da matematici arabi. Fu però Lagrange per primo a dimostrare, per ogni valore di N , l’esistenza di una soluzione dalla quale ricavarne poi infinite.

Tale operazione si trova risolta anche da Dirichlet con procedimenti che coinvolgono l’algebra astratta (20) e per essa furono scritti numerosi lavori e il libro *Geschichte der Gleichung $t^2 - Du^2 = 1$* (Storia dell’equazione $t^2 - Du^2 = 1$) da H. Konen (21).

19. Si veda ad esempio *Le scienze esatte nell’antica Grecia* di G. Loria, Hoepli, Milano, 1914, pp. 932 sgg. Si vedano anche il libro di H. Konen che verrà citato compiutamente più avanti e quanto ne scrive T. Heath in *Diophantus of Alexandria* (Cambridge, Univ. Press 1910) Section II, pp. 227 sgg.

20. V. ad esempio *Algèbre* di P. Dubreil, Gauthier-Villars, Paris 1946, pp. 202 sgg.

21. Leipzig, 1901. Nel libro vengono descritti i contributi, oltre che dei matematici arabi, di Fermat, Frenicle, Wallis, Brouncker, Eulero, Lagrange, Gauss, Dirichlet e Jacobi.

MATEMATICA E FIGURINE*

Antonio Maturo

*Dipartimento di Scienze, Storia dell'Architettura e Restauro
Università "G. D'Annunzio", Viale Pindaro 42, Pescara*

1. IL PROBLEMA DELLE FIGURINE

Alcuni anni fa mio figlio Alessandro decise di fare la collezione delle figurine dei calciatori.

Si trattava di riempire un album di 600 figurine ed ogni figurina costava L.10 (oggi costano £.50).

Fui contento della sua decisione perchè pensavo di aver trovato una maniera per impegnarlo e anche per farlo crescere culturalmente (aveva tre anni e sei mesi).

Stimavo di comprargli 20 figurine al giorno, spendendo così £. 200 (equivalenti a £. 1.000 attuali) e mi sembrava che la soluzione fosse abbastanza economica. Mi insospettiva un poco il fatto che l'album, molto bello e ricco di notizie, fosse distribuito gratuitamente, ma non diedi molto peso a tale circostanza.

Dopo tre o quattro giorni, però, la situazione precipitò: mio figlio, preso dall'ansia di completare l'album, voleva le 20 figurine (e possibilmente di più) ogni volta che passavamo davanti ad un giornalaio.

Cominciarono inoltre rapidamente a comparire i dopplioni, i triploni e l'album non si riempiva mai.

Mi chiesi allora preoccupato: quanto si spende a riempire un album di figurine?

Alessandro non aveva buttato nessuna figurina multipla, nè aveva amici per fare scambi, per cui cominciai a raccogliere dei dati.

Su 530 figurine acquistate ne erano state messe, sull'album, 339, mentre 191 erano almeno dopplioni.

* Lavoro presentato al Convegno Nazionale *Mathesis* 1989,
Gioia del Colle, 29-30 aprile

Cinque giorni dopo notai che, su 790 figurine acquistate ne erano state attaccate, sull'album, 427, mentre 363 erano almeno doppioni. In cinque giorni erano state acquistate 260 figurine (per una spesa equivalente a £.13.000 attuali) e di queste sull'album ne erano state attaccate appena 88.

Alcuni giorni dopo il totale delle figurine acquistate era diventato 930 mentre quelle attaccate all'album erano 471.

A questo punto pensai che, come professore di Matematica, fosse doveroso trovare una formula per stabilire la relazione fra figurine acquistate e figurine messe nell'album e, in definitiva, per prevedere la spesa necessaria per riempire l'album.

2. RISOLUZIONE APPROSSIMATA CON LE EQUAZIONI DIFFERENZIALI

Indichiamo con x il numero di figurine acquistate e con y quelle delle figurine messe nell'album.

La probabilità di attaccare una ulteriore figurina acquistata sull'album è, allora,

$$(2.1) \quad p(y) = (600 - y) / 600.$$

Facciamo le seguenti approssimazioni:

(1) trattiamo y e x come variabili continue e non discrete come effettivamente sono. Ammettiamo che si possa comprare una quantità infinitesima dx di figurine e metterne, in corrispondenza, una quantità dy ;

(2) interpretiamo le probabilità come frequenze.

In base alle (1) e (2) possiamo allora scrivere

$$dy = \frac{600 - y}{600} dx; \quad \frac{600}{600 - y} dy = dx$$

Integrando con la condizione iniziale che per $x = 0$ (nessuna figurina acquistata) sia $y = 0$ (nessuna figurina messa) si ottiene

$$(2.2) \quad x = 600 \log \frac{600}{600 - y}.$$

Confrontiamo i dati teorici, ottenuti dalla formula (2.2), con quelli osservati. Si ha la tabella

y	x osservato	x teorico
339	530	499
427	790	746
471	930	922

Tabella 1

Il grado di vicinanza fra valori teorici e valori osservati era veramente eccezionale, tenuto conto del fatto che le approssimazioni effettuate erano tali da far inorridire sia un "analista" che un "probabilista".

Ciò, tuttavia, non era affatto consolante per il mio portafoglio, poichè dalla (2.2) segue

$$(2.3) \quad \lim_{y \rightarrow 600} x = +\infty$$

Leggendo attentamente l'album, però, ho scoperto che la casa editrice dell'album era disposta a fornire, al prezzo di £. 25 l'una, fino a 20 figurine (per una spesa equivalente a £. 2.500 attuali, un vero affare per me!).

In base alla (2.2) le figurine da acquistare per attaccarne 580 sull'album sono, approssimativamente, $600 \log 30 = 2.041$ (per una spesa equivalente a £. 102.050 attuali).

Dopo aver fatto tutti questi calcoli (nel frattempo avevo acquistato altre 200 figurine, quasi tutte doppioni) ho spiegato a mio figlio che non era più il caso di proseguire la collezione poichè a quel punto si dovevano comprare quasi 10 figurine per metterne una e quindi ogni figurina messa costava in media una somma pari alle attuali £. 500.

Alessandro capì perfettamente i miei calcoli e non mi chiese più soldi per le figurine. Però i giorni successivi lo vidi sempre con le tasche piene di doppioni acquistati dai nonni.

3. RISOLUZIONE DEL PROBLEMA CON IL CALCOLO DELLE PROBABILITÀ

Supponiamo di aver messo y figurine.

La probabilità di attaccare una ulteriore figurina acquistata è allora data dalla (2.1).

Indichiamo con $N(y)$ il numero di figurine da acquistare per metterne una nuova dopo che ne sono state attaccate y .

La variabile casuale $N(y)$ può assumere i valori $1, 2, \dots, n, \dots$ con probabilità $p(y), (1 - p(y)) p(y), \dots, (1 - p(y))^{n-1} p(y)$,

$$\text{media } m(y) = \frac{1}{p(y)} \quad \text{e varianza } V(y) = \frac{1 - p(y)}{[p(y)]^2}.$$

Il numero di figurine da acquistare per poterne mettere y è allora:

$$X = N(0) + N(1) + \dots + N(y - 1)$$

Il valore medio di X è dato dalla formula:

$$(3.1) \quad m(X) = \sum_{i=0}^{y-1} m(i) = 600 \sum_{i=0}^{y-1} \frac{1}{600 - i}$$

e, poichè le $N(i)$ sono indipendenti, la varianza di X è uguale a

$$(3.2) \quad V(X) = \sum_{i=0}^{y-1} V(i) = 600 \sum_{i=0}^{y-1} \frac{i}{(600 - i)^2}$$

È noto che (cfr. ad es. [3], [6])

$$(3.3) \quad \sum_{i=1}^n \frac{1}{i} = \log n + \log \left(1 + \frac{1}{n}\right) + \gamma - \frac{\Theta_{n+1}}{n+1}$$

dove Θ_{n+1} è un numero appartenente all'intervallo $(0, 1)$, la successione di termine generale $\Theta_{n+1} / (n+1)$ è strettamente decrescente e γ è la costante di Eulero-Mascheroni, uguale a 0,5772155...

Dalla (3.1) si ottiene:

$$m(X) = 600 \left[\sum_{j=1}^{600} \frac{1}{j} - \sum_{j=1}^{600-y} \frac{1}{j} \right]$$

e quindi, per la (3.3),

$$m(X) = 600 \left[\log 600 - \log(600 - y) + \log \left(1 + \frac{1}{600} \right) - \log \left(1 + \frac{1}{600 - y} \right) + \right. \\ \left. - \Theta_{601} / 601 + \Theta_{600-y+1} / (600 - y + 1) \right]$$

ossia, ponendo

$$(3.4) \quad \alpha = \left[\log \left(1 + \frac{1}{600} \right) - \log \left(1 + \frac{1}{600 - y} \right) \right] + \left[\frac{\Theta_{600-y+1}}{600 - y + 1} - \frac{\Theta_{601}}{601} \right]$$

$$(3.5) \quad m(X) = 600 \left(\log \frac{600}{600 - y} + \alpha \right)$$

Con facili calcoli si vede subito che, per $y \leq 580$ è certamente $|\alpha| < 0,048$ e, per $y \leq 500$, risulta $|\alpha| < 0,002$.

Quasi sicuramente, poi, α è molto più vicino a zero da quanto appare dalle precedenti disuguaglianze, poichè, nella (3.4), il termine fra parentesi quadre a sinistra, negativo, e quello fra parentesi quadre a destra, positivo, sono dello stesso ordine di grandezza.

La (3.5) permette di chiarire il significato della (2.2).

Infatti confrontando le due formule segue che

$$m(X) = x + 600 \alpha$$

Possiamo osservare che 600α è minore di 29 per $y \leq 580$ ed è minore di 2 per $y \leq 500$. Ciò implica che si può assumere, con buona approssimazione, $m(X) \approx x$.

Per poter valutare la variabilità di X , diamo una maggiorazione della sua varianza.

Poichè:

$$\frac{i}{(600-i)^2} = \frac{600}{(600-i)^2} - \frac{1}{600-i}$$

risulta, per la (3.2),

$$V(X) = 600^2 \sum_{j=600-y+1}^{600} \frac{1}{j^2} - m(X)$$

e quindi

$$V(X) < 600^2 \sum_{j=600-y+1}^{600} \frac{1}{j(j-1)} - m(X) = 600^2 \left(\frac{1}{600-y} - \frac{1}{600} \right) - m(X)$$

Si ha dunque

$$(3.6) \quad V(X) < \frac{600^2}{600-y} - 600 \left(1 + \log \frac{600}{600-y} \right) - 600 \alpha .$$

In particolare, per $y \leq 580$, essendo $V(X)$ funzione crescente di y , si ha

$$V(X) < 600^2/20 - 600(1 + \log 30) + 600 \mid \alpha \mid$$

ossia, tenuto conto delle limitazioni su α ,

$$V(X) < 18000 - 2641 + 29 = 15388.$$

Lo scarto quadratico medio, in corrispondenza, soddisfa alla condizione

$$\sigma < 125$$

Per $y \leq 500$, invece, si ha, per la (3.6),

$$V(X) < 600^2/100 - 600 (1 + \log 6) + 600 \alpha$$

da cui

$$V(X) < 18000 - 1675 + 2 = 1927; \sigma < 44.$$

Assumendo l'ipotesi che la X sia distribuita normalmente, per $y \leq 500$ la probabilità dell'evento $\{ |X - m(X)| < 44 \}$ è maggiore di 0,68 ; quella dell'evento $\{ |X - m(X)| < 88 \}$ è maggiore di 0,95.

Ciò spiega i risultati della tabella 1.

Al crescere di y gli intervalli di confidenza al 68% e al 95% si allargano e per $y = 580$, massimo valore di y , si ottiene, sempre nell'ipotesi che X sia normale,

$$\text{prob} \{ |X - m(X)| < 125 \} > 0,68$$

e

$$\text{prob} \{ |X - m(X)| < 250 \} > 0,95$$

BIBLIOGRAFIA

- [1] G.CALOT. *Cours de calcul des probabilités*. Dunod, Paris 1967
- [2] G.DALL'AGLIO. *Calcolo delle probabilità*. Zanichelli, Bologna. 1987
- [3] A.GHIZZETTI. *Complementi ed esercizi di Analisi Matematica vol. I*. Libreria Veschi, Roma. 1966
- [4] A.MATURO. *Numeri pseudocasuali*. Montefeltro Edizioni Urbino 1983
- [5] R.SCOZZAFAVA. *Introduzione alla Probabilità e alla Statistica*. Libreria Veschi, Roma. 1984
- [6] G.ZWIRNER. *Esercizi e complementi di Analisi Matematica vol. I*. Cedam, Padova. 1967

TEORIA DELLA DIVISIBILITÀ E ALGORITMO EUCLIDEO NELL'INSEGNAMENTO NELLE SCUOLE SUPERIORI

Antonio Maturo* - Fabiola Paccapelo** - Loredana Renzullo*

*Dipartimento di Scienza, Storia dell'Architettura e Restauro
Università "G. D'Annunzio" viale Pindaro, 42 - Pescara

**Liceo Scientifico "L. Da Vinci" via Collemarino - Pescara

Sommario. Si presenta una proposta per introdurre la teoria della divisibilità nelle Scuole Medie Superiori. Nel primo paragrafo si richiama la divisibilità in $\mathcal{N}$ evidenziando le proprietà da generalizzare ed il concetto di norma. Nel secondo paragrafo si danno le proprietà che riteniamo fondamentali per presentare nelle scuole la divisibilità in un anello integro, con particolare riferimento a quelli normati ed euclidei. Negli esempi si mettono bene in luce le differenze fra elementi primi ed elementi irriducibili e quindi l'importanza dell'algoritmo euclideo.

1. DIVISIBILITÀ IN $\mathcal{N}$

Sia $\mathcal{N} = \{0, 1, 2, \dots\}$ l'insieme dei naturali.

Definizione 1.1 (Divisibilità)

Siano $a, b \in \mathcal{N}$. Diciamo che a divide b oppure che a è un divisore di b e scriviamo $a|b$ se:

$$\exists c \in \mathcal{N}: b = ca.$$

Poniamo $\forall a, b \in \mathcal{N}$,

$$a \oplus b \Leftrightarrow \begin{cases} b = 0 \\ b \neq 0, a \neq 0, a \leq b \end{cases}$$

Proposizione 1.1

La relazione $\oplus$ in $\mathcal{N}$ è di ordine totale con 1 elemento minimo e 0 elemento massimo.

Proposizione 1.2

La relazione $|$ in $\mathcal{N}$ è di ordine parziale, con 1 elemento minimo e 0 elemento massimo, ossia:

$$1|b \text{ e } b|0 \quad \forall b \in \mathcal{N}$$

Proposizione 1.3

$a|b \Rightarrow a \subseteq b$, ossia $|$ è una restrizione di $\subseteq$.

Proposizione 1.4

Il numero 1 ha un solo divisore; ogni $a \neq 1$ ha almeno 2 divisori: 1 ed a .

Il numero 0 ha infiniti divisori; ogni $a \neq 0$ ha un numero finito di divisori, non superiori ad a .

Definizione 1.2

Sia $p \in \mathcal{N} - \{0, 1\}$. Diciamo che p è *irriducibile* o *indecomponibile* o *primo* se:

$$a|p \Rightarrow (a = 1 \text{ oppure } a = p)$$

Definizione 1.3

Siano $a, b \in \mathcal{N}$. Diciamo che $d \in \mathcal{N}$ è un *massimo comune divisore* tra a e b se:

$$(D1) \quad d|a, d|b$$

$$(D2) \quad (c|a, c|b) \Rightarrow c|d$$

Proposizione 1.5

$\forall a, b \in \mathcal{N}$ esiste al più un massimo comun divisore di a e b indicato con $D(a, b)$.

Dim.: d e d' massimi comuni divisori di a e $b \Rightarrow (d|d' \text{ e } d'|d) \Rightarrow (d \leq d', d' \leq d) \Rightarrow d = d'$

Proposizione 1.6

$$\forall a \in \mathcal{N}, D(a, 0) = a, D(a, 1) = 1$$

Proposizione 1.7

$$\left. \begin{array}{l} \exists D(b, r) \\ a = bx + r \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists D(a, b) \\ D(a, b) = D(b, r) \end{array} \right.$$

Dim.: $(d = D(b, r), a = bx + r) \Rightarrow (d|b, d|r) \Rightarrow (d|a, d|b);$

$$(c|a, c|b, r = a - bx) \Rightarrow c|r, c|b \Rightarrow c|d.$$

Teorema 1.1

Per ogni $a, b \in \mathcal{N}$, $b \neq 0$ esiste una unica coppia ordinata (q, r) di elementi di $\mathcal{N}$, detti rispettivamente *quoziente* e *resto* della divisione di a per b e indicati con $\text{quot}(a, b)$, $\text{rest}(a, b)$ tali che:

$$(DR1) \quad a = bq + r$$

$$(DR2) \quad 0 \leq r < b$$

Proposizione 1.8

Per ogni $a, b \in \mathcal{N}$, $b \neq 0$, poniamo $r = \text{rest}(a, b)$.

Si ha:

$$d = D(b, r) \Rightarrow d = D(a, b)$$

$$r = 0 \Rightarrow D(a, b) = b$$

Dim.: segue dalle prop. 1.6 e 1.7

Teorema 1.2 (Algoritmo euclideo)

Per ogni $a, b \in \mathcal{N}$, $b \neq 0$, esiste una successione finita $\{r_1, r_2, \dots, r_n\}$, $n \geq 2$ di elementi di $\mathcal{N}$ tali che:

$$\text{posto} \quad a = r_0, \quad b = r_1,$$

$$(AE1) \quad r_i > 0 \quad \forall i < n \quad \text{e} \quad r_n = 0$$

$$(AE2) \quad r_i = \text{rest}(r_{i-2}, r_{i-1})$$

$$(AE3) \quad D(a, b) = r_{n-1}$$

Corollario del teorema 1.2

Per ogni $a, b \in \mathcal{N}$ esiste ed è unico il massimo comun divisore di a e b .

Teorema 1.3

Per ogni $a, b \in \mathcal{N}$ e $\forall m \in \mathcal{N} - \{0\}$, $a|b \Leftrightarrow am|bm$

$$d = D(a, b) \Leftrightarrow md = D(ma, mb)$$

Teorema 1.4 (Combinazione lineare)

Per ogni $a, b \in \mathcal{N}$ esistono h e $k \in \mathbb{Z}$ tali che, posto $d = D(a, b)$,

$$d = ah + bk.$$

Inoltre

$$d = \min \{ |ah + bk|, h \in \mathbb{Z}, k \in \mathbb{Z} \}.$$

Definizione 1.4

Per ogni $a, b \in \mathcal{N}$, si dice che a e b sono *primi fra loro* se $D(a, b) = 1$.

Corollario del Teorema 1.4

$$a \text{ e } b \text{ primi fra loro} \Leftrightarrow \exists h, k \in \mathbb{Z} : 1 = ah + bk.$$

Teorema 1.5 (Fondamentale della teoria della divisibilità).

$$\text{Per ogni } a, b \in \mathcal{N} \ (k|ab, D(k, a) = 1) \Rightarrow k|b$$

$$\text{Dim. } (D(k, a) = 1) \Rightarrow (D(kb, ab) = b) \Rightarrow (k|kb, k|ab \Rightarrow k|b)$$

Teorema 1.6

$$\text{Per ogni } a \in \mathcal{N}_1 \ p \text{ primo} \Rightarrow (p \nmid a \Leftrightarrow D(p, a) = 1)$$

$$\text{Dim. } (p \text{ primo}, p \nmid a) \Leftrightarrow (p \text{ primo e solo } 1 \text{ è tale che } 1|p \text{ e } 1|a)$$

Teorema 1.7

Per ogni $a, b \in \mathcal{N}$

$$p \text{ primo} \Leftrightarrow p|a \cdot b \Leftrightarrow p|a \text{ o } p|b$$

$$\text{Dim.: } (\Rightarrow) (p \text{ primo}, p|ab, p \nmid a) \Rightarrow (p|ab, D(p, a) = 1) \Rightarrow p|b$$

$$(\Leftarrow) (p \text{ non primo}) \Rightarrow (\exists r, s, 1 < r, s < p : p|r \cdot s, p \nmid r, p \nmid s)$$

Teorema 1.8 (Fondamentale dell'aritmetica)

Ogni $a \in \mathcal{N} - \{0, 1\}$ si può esprimere in un unico modo come prodotto $a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_n^{\alpha_n}$ con i p_i fattori primi tali che $p_1 < p_2 < \dots < p_n$ e $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathcal{N} - \{0\}$

2. DIVISIBILITÀ IN UN DOMINIO DI INTEGRITÀ

Sia $\mathcal{A}$ un anello commutativo unitario.

Definizione 2.1 (Divisibilità)

Siano $a, b \in \mathcal{A}$.

Diciamo che a *divide* b oppure che a è un divisore di b e scriviamo $a|b$ se $\exists c \in \mathcal{A} : b = c \cdot a$.

Teorema 2.1

La relazione $|$ è un preordine in $\mathcal{A}$, ossia è riflessiva e transitiva.

Se $b \in \mathcal{A}$ ed x è un elemento invertibile di $\mathcal{A}$ risulta $x|b, b|0$.

Definizione 2.2

Si dice che a e $b \in \mathcal{A} - \{0\}$ sono *divisori dello zero* “complementari fra loro” se $a \cdot b = 0$.

Un anello privo di divisori dello zero si dice *dominio di integrità* o *anello integro*.

Definizione 2.3

Due elementi a e $b \in \mathcal{A}$ si dicono *associati* se $a|b$ e $b|a$.

Teorema 2.2

Sia $\mathcal{A}$ un dominio di integrità. Allora

$(a \text{ associato a } b) \Leftrightarrow (\exists c \in \mathcal{A}, c \text{ invertibile} : a = c \cdot b)$

Teorema 2.3

La relazione in $\mathcal{A}$, $a \sim b$ se e solo se a è associato a b è una relazione di equivalenza. Essa determina una partizione in classi di associati.

Se $\mathcal{A}$ è un dominio di integrità, detta $[a]$ la classe degli associati ad $a \in \mathcal{A}$, risulta:

$$[a] = \{x \in \mathcal{A} : \exists c \text{ invertibile } \in \mathcal{A} \text{ tale che } x = c \cdot a\}.$$

Teorema 2.4

Sia $\mathcal{D}(a)$ l'insieme dei divisori di $a \in \mathcal{A}$. Allora:

$$a \sim b \Leftrightarrow \mathcal{D}(a) = \mathcal{D}(b).$$

Teorema 2.5

Ogni elemento invertibile $a \in \mathcal{A}$ è tale che $\mathcal{D}(a) = [1]$.

Ogni elemento non invertibile b di un dominio di integrità è tale che:

$$\mathcal{D}(b) \supseteq [1] \cup [b] \quad \text{con} \quad [1] \cap [b] \neq \emptyset$$

Definizione 2.4

Si dice *norma* in $\mathcal{A}$ ogni funzione $v : \mathcal{A} \rightarrow \mathcal{N}$ tale che

$$(N1) \quad v(a) = 0 \Leftrightarrow a = 0$$

$$(N2) \quad v(a \cdot b) = v(a) \cdot v(b).$$

Teorema 2.6

Esiste una norma in $\mathcal{A}$ se e solo se $\mathcal{A}$ è un dominio di integrità.

Dim. Se $a \cdot b = 0$, $a \neq 0$, $b \neq 0$ e valgono (N1) ed (N2) si ha:

$$v(a \cdot b) = v(0) = 0 \text{ e } v(a \cdot b) = v(a) \cdot v(b) \neq 0. \quad \text{Assurdo.}$$

La $v(0) = 0$, $v(a) = 1$ per $a \neq 0$ è una norma se $\mathcal{A}$ è un dominio d'integrità.

Supponiamo, da ora in poi, che $\mathcal{A}$ sia un dominio di integrità e che v sia una norma in $\mathcal{A}$.

Teorema 2.7

$$a|b \Rightarrow v(a)|v(b) \quad \text{in } \mathcal{K}$$

$$\text{Dim.:} \quad b = ca \Rightarrow v(b) = v(c) \cdot v(a)$$

Corollario al teorema 2.7

$$a \sim b \Rightarrow v(a) = v(b)$$

$$a \text{ invertibile} \Rightarrow v(a) = 1$$

$$\text{Dim.:} \quad a \sim b \Rightarrow v(a)|v(b) \text{ e } v(b)|v(a)$$

$$v(1) = v(1 \cdot 1) = v(1) \cdot v(1) \Rightarrow v(1) = 1$$

Definizione 2.5

Sia $p \in \mathcal{A} - ([0] \cup [1])$. Diciamo che p è *indecomponibile* o *irriducibile* se:

$$a|p \Rightarrow (a \in [1] \text{ oppure } a \in [p])$$

Diciamo che $p \in \mathcal{A} - \{[0] \cup [1]\}$ è *primo* se, $\forall a, b \in \mathcal{A}$

$$p|a \cdot b \Leftrightarrow p|a \text{ o } p|b.$$

Teorema 2.8

Se p è primo in $\mathcal{A}$ allora è anche irriducibile.

Dim.: Sia p primo e sia $a \in \mathcal{A}$ tale che $a|p$.

Allora $\exists k \in \mathcal{A} \mid p = k \cdot a$.

Essendo p primo $p|a$ o $p|k$.

Poiché $k|p$ e $a|p$ uno fra k e a è un associato di p e l'altro è un elemento invertibile. Segue che $a \in [1]$ o $a \in [p]$ e quindi che p è irriducibile.

Definizione 2.6

Siano $a, b \in \mathcal{A}$. Diciamo che $d \in \mathcal{A}$ è un *massimo comun divisore* tra a e b se

$$(D1) \quad d|a, d|b$$

$$(D2) \quad (c|a, c|b) \Rightarrow c|d.$$

Teorema 2.9

Per ogni $a, b \in \mathcal{A}$ se d è un massimo comun divisore tra a e b sono tali tutti e soli gli elementi di $[d]$.

Sia $\mathcal{D}(a, b)$ eventualmente vuoto, l'insieme dei massimi comuni divisori di a e b .

Teorema 2.10

Per ogni $a \in \mathcal{A}$ e per ogni c invertibile $\in \mathcal{A}$

$$D(a, 0) = [a], \quad D(a, c) = [1].$$

Teorema 2.11

$$\left. \begin{array}{l} D(b, r) \neq \emptyset \\ a = bx + r \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} D(a, b) \neq \emptyset \\ D(a, b) = D(b, r) \end{array} \right.$$

Definizione 2.7

La norma v in $\mathcal{A}$ si dice *euclidea* se soddisfa la

$$(N3) \quad \forall a, b \in \mathcal{A}, \quad b \neq 0, \quad \exists q, r \in \mathcal{A} \quad \text{tali che:}$$

$$(DR1) \quad a = bq + r$$

$$(DR2) \quad 0 \leq v(r) < v(b).$$

Se v è euclidea, $\mathcal{A}$ si dice *dominio di integrità euclideo*.

Teorema 2.12 (Algoritmo euclideo)

Per ogni $a, b \in \mathcal{A}$, $b \neq 0$, se v è una norma euclidea esiste una successione finita $\{r_1, r_2, \dots, r_n\}$, $n \geq 2$, di elementi di $\mathcal{A}$ tali che posto $a = r_0$, $b = r_1$:

$$(AE1) \quad v(r_i) > 0 \quad \forall i < n \quad \text{e} \quad v(r_n) = 0$$

$$(AE2) \quad \exists q_i \in \mathcal{A} : r_{i-2} = q_i r_{i-1} + r_i, \quad v(r_i) < v(r_{i-1})$$

$$(AE3) \quad D(a, b) = [r_{n-1}].$$

Definizione 2.8

Sia $\mathcal{A}$ un anello.

Per ogni $a \in \mathcal{A} - ([0] \cup [1])$, si dice *fattorizzazione* di a ogni successione finita $\{a_0, a_1, a_2, \dots, a_n\}$ di elementi di $\mathcal{A}$ con $n \geq 1$ tali che:

$$(F_1) \quad a_0 \in [1]$$

$$(F_2) \quad a_i \text{ è primo, } \forall i \in \{1, 2, \dots, n\}$$

$$(F_3) \quad a = a_0 a_1 a_2 \dots a_n$$

Definizione 2.9

Per ogni $a \in \mathcal{A} - ([0] \cup [1])$, due fattorizzazioni

$$\mathcal{S}_1 = \{a_0 a_1 a_2 \dots a_n\} \quad \text{e} \quad \mathcal{S}_2 = \{b_0 b_1 b_2 \dots b_m\}$$

si dicono *equivalenti* se esiste una biiezione φ di $\mathcal{S}_1$ in $\mathcal{S}_2$ tale che:

$$[a_i] = [\varphi(a_i)], \quad \forall i \in \{0, 1, 2, \dots, n\}.$$

(In particolare $n = m$).

Teorema 2.13 [cfr. [1]].

Sia $\mathcal{A}$ un anello euclideo. Per ogni $a \in \mathcal{A} - ([0] \cup [1])$, esiste almeno una fattorizzazione di a e se $\mathcal{S}_1$ ed $\mathcal{S}_2$ sono due fattorizzazioni di a , allora $\mathcal{S}_1$ è equivalente ad $\mathcal{S}_2$.

Teorema 2.14

Sia $\mathcal{A}$ un anello euclideo. Un elemento di $\mathcal{A}$ è irriducibile se e solo se è primo.

3. ESEMPI ED ESERCIZI**3.1 Caso con due fattorizzazioni non equivalenti**

Si consideri il dominio di integrità:

$$H(3) = \{z \in \mathbf{C} : z = a + ib\sqrt{3}, \quad a \in \mathcal{A}, b \in \mathcal{A}\}$$

risulta

$$4 = (1 + i\sqrt{3})(1 - i\sqrt{3}) \quad \text{e} \quad 4 = 2 \cdot 2$$

Le due fattorizzazioni di 4 *non sono equivalenti*.

Gli elementi:

$$2, \quad 1 + i\sqrt{3}, \quad 1 - i\sqrt{3} \quad \text{sono irriducibili in } H(3).$$

Essi, però, non sono primi. Infatti:

$$2 \mid (1 + i\sqrt{3})(1 - i\sqrt{3}) \quad \text{ma} \quad 2 \nmid (1 + i\sqrt{3}) \quad \text{e} \quad 2 \nmid (1 - i\sqrt{3}).$$

3.2 Esempi di domini di integrità dotati di norma

Per ogni $c \in \mathcal{N} - \{0\}$, consideriamo l'insieme $H(c)$ dei numeri complessi:

$$z = a + ib\sqrt{c} \quad \text{con } a, b \in \mathcal{N}.$$

$H(c)$, rispetto alle usuali operazioni di addizione e moltiplicazione in $\mathcal{C}$ forma un *dominio d'integrità*.

La *norma* di $z = a + ib\sqrt{c}$ è il numero:

$$N(z) = z \cdot z', \quad \text{con } z' = a - ib\sqrt{c} \quad (\text{coniugato di } z)$$

Precisamente:

$$N(z) = a^2 + b^2c.$$

Si ha che

$$N(z) \in \mathcal{N} \quad \text{e inoltre}$$

- 1) $N(z) = 0 \Leftrightarrow z = 0$
- 2) $\forall z_1, z_2 \in \mathcal{A}; \quad N(z_1 \cdot z_2) = N(z_1) \cdot N(z_2)$

3.3 Divisione con il resto in $H(c)$

(a) Teoria. **Teorema [cfr. [1]]**

Siano u e $v \in H(c)$, con $v \neq 0$. Esistono, in $H(c)$ due elementi q ed r , detti *quoziente e resto della divisione* di u per v , tali che:

$$[D_1] \quad u = vq + r$$

$$[D_2] \quad N(r) \leq (1 + c) N(v) \mid 4.$$

Per $c \leq 2$, $N(r) < N(v)$ per cui $H(c)$ è un anello euclideo.

Per $c = 3$ dalla $[D_2]$ segue $N(r) \leq N(v)$.

(b) Come si calcolano q , r , $N(r)$

Si ha che $u = vq + r \Rightarrow uv' = vv'q + v'r$ e $vv' = N(v) \in \mathcal{A}$. Allora si divide uv' per $N(v)$ come in $\mathcal{Z}$, considerando separatamente la parte reale e quella immaginaria e ottenendo q e $v'r$.

Poniamo $r^* = v'r.$

Essendo $N(r^*) = N(v'r) = N(v') \cdot N(r),$

si ottiene $N(r) = N(r^*) \mid N(v').$

Inoltre da $r^* = v'r,$

si ottiene $vr^* = vv'r, \quad vr^* = N(v)r,$

$$r = \frac{1}{N(v)} vr^*.$$

(c) Esempio

Dividiamo $u = 7 + i8\sqrt{2}$ per $v = 4 + i7\sqrt{2}.$

Si ha: $uv' = 140 - i17\sqrt{2}, \quad N(v) = 114$

Sia $q = \text{quot}(uv', 114)$, $r^* = \text{rest}(uv', 114)$.

Si ottiene $q = 1$, $r^* = 26 - i17\sqrt{2}$.

Segue $N(r^*) = 1254$, $N(r) = 1254 : 114 = 11$,

$$\begin{aligned} r &= \frac{v}{N(v)} \quad r^* = \frac{1}{114} (4 + i7\sqrt{2})(26 - i17\sqrt{2}) = \\ &= \frac{1}{114} (342 + i114\sqrt{2}) = 3 + i\sqrt{2}. \end{aligned}$$

Notiamo che $N(r) = 11 \ll N(v) = 114$

(d) Un altro esempio

Dividiamo 2 per $1 + i\sqrt{3}$.

Si ha $u = 2$, $v = 1 + i\sqrt{3}$

$$N(v) = 4, \quad uv' = 2(1 - i\sqrt{3}) = 2 - 2i\sqrt{3},$$

$$q = -i\sqrt{3}, \quad r^* = 2 + 2i\sqrt{3}, \quad N(r^*) = 16, \quad \text{da cui}$$

$$N(r) = N(r^*) | N(v) = 4, \quad \text{ossia} \quad N(r) = N(v).$$

3.4 Un esempio in cui $\mathcal{D}(a, b) = \emptyset$

In $H(3)$ sia

$$a = 4, \quad b = 2(1 + i\sqrt{3})$$

Risulta che $D(a, b) \neq \emptyset$.

Infatti se $d \in \mathcal{D}(a, b)$ si ha $(z|a, z|b) \Rightarrow z|d$.

Allora $1 + i\sqrt{3}$ e 2 dividono d.

I divisori di b sono $1, 2, 1 + i\sqrt{3}, b$ e poiché:

$$2 \nmid (1 + i\sqrt{3}) \text{ e } (1 + i\sqrt{3}) \nmid 2,$$

deve essere $d = b$.

Allora: $b = 2(1 + i\sqrt{3}) \mid 4$,

ossia $1 + i\sqrt{3} \mid 2$, assurdo.

BIBLIOGRAFIA

- [1] A. MATURO – B. RIZZI, *Sulla decomposizione di un numero primo in somma di tre quadrati* Periodico di Matematiche n° 3 – 1990.
- [2] A. MATURO – B. RIZZI, *Sulla decomposizione di un numero naturale in somma di quadrati* Periodico di Matematica n° 4 – 1990
- [3] A. WEIL – *Teoria dei numeri* Einaudi 1993

CALCOLO DI AREE E VOLUMI PER MEZZO DI SUCCESSIONI DI NUMERI A CASO*

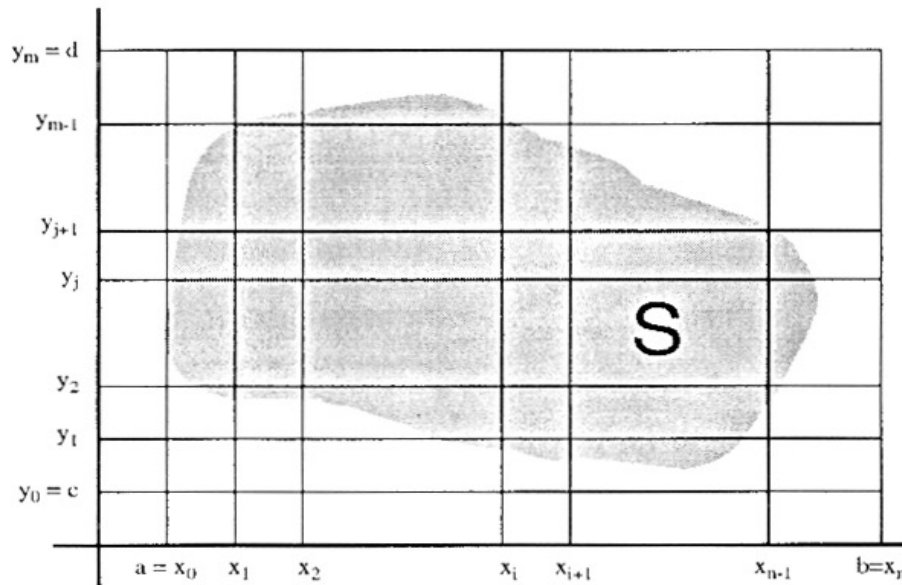
Giuseppina Varone
Istituto Mecenate, via Marco Polo - Pescara

1. AREA DI UN INSIEME PIANO: IL PROCEDIMENTO DI PEANO-JORDAN

Sia S un insieme limitato del piano. Allora esiste un intervallo $I = [a, b] \times [c, d]$ che lo contiene.

Siano $x_0 = a < x_1 < x_2 \dots < x_{n-1} < x_n = b$ punti dell'intervallo $[a, b]$ e siano $y_0 = c < y_1 < y_2 \dots < y_{m-1} < y_m = d$ punti dell'intervallo $[c, d]$.

Tracciando dai punti x_i le parallele all'asse y e dai punti y_i le parallele all'asse x , l'intervallo I viene suddiviso in $m \cdot n$ intervalli $I_{ij} = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$.



* Conferenza tenuta al Convegno Interregionale "Probabilità, Statistica e Crittografia" - Isernia, 1992

L'insieme degli intervallini I_{ij} si dice partizione dell'intervallo I . Indichiamolo con Π . La somma delle misure degli intervallini I_{ij} contenuti in S è un valore per difetto dell'area di S e si indica con $S_1(\Pi)$. La somma delle misure degli intervallini I_{ij} che hanno intersezione non vuota con S è un valore per eccesso dell'area di S e si indica con $S_2(\Pi)$.

Indichiamo con Σ_1 l'insieme delle somme $S_1(\Pi)$ e con Σ_2 l'insieme delle somme $S_2(\Pi)$ al variare di Π nell'insieme delle partizioni dell'intervallo I .

I due insiemi Σ_1 e Σ_2 sono separati ossia si dimostra che dette Π_1 e Π_2 due partizioni di I risulta sempre $S_1(\Pi_1) \leq S_2(\Pi_2)$.

Allora detti σ_1 e σ_2 rispettivamente l'estremo superiore di Σ_1 e l'estremo inferiore di Σ_2 risulta $\sigma_1 \leq \sigma_2$. I numeri σ_1 e σ_2 si dicono rispettivamente misura interna (m_i) e misura esterna (m_e) dell'insieme S . Se risulta $\sigma_1 = \sigma_2$ allora l'insieme S si dice misurabile secondo Peano-Jordan e il numero $\sigma_1 = \sigma_2$ si dice misura dell'insieme S secondo Peano-Jordan e si indica con $m(S)$.

2. AREA DI UN INSIEME PIANO PER MEZZO DI NUMERI A CASO

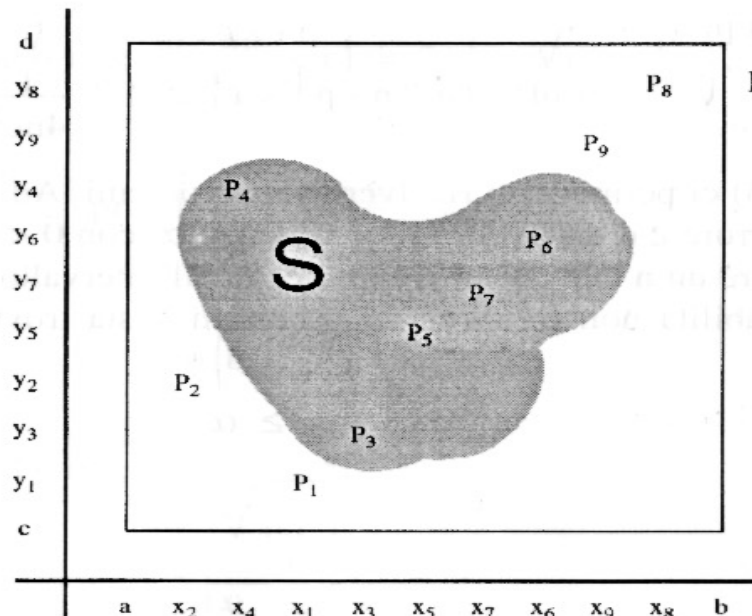
Sia S un insieme limitato del piano misurabile secondo Peano-Jordan e sia $I = [a, b] \times [c, d]$ un intervallo contenente S .

Generiamo n numeri a caso nell'intervallo $[a, b]$ ed n numeri a caso nell'intervallo $[c, d]$. Siano $x_1, x_2, x_3, \dots, x_n$ i numeri appartenenti all'intervallo $[a, b]$ e $y_1, y_2, y_3, \dots, y_n$ i numeri appartenenti all'intervallo $[c, d]$.

I punti $P_i = (x_i, y_i)$ possono appartenere o no all'insieme S .

Sia m il numero di tali punti appartenenti ad S .

Per n abbastanza grande il rapporto tra l'area di S e l'area di I è approssimativamente dato dal numero m/n e generalmente l'approssimazione migliora al crescere di n .



Indichiamo con p la probabilità che un punto preso a caso nell'intervallo I cada nell'insieme S . Allora

$$p = \frac{\text{Area } S}{\text{Area } I}$$

Per la legge di Bernoulli il numero m di volte che, in n prove, un punto p preso a caso in I cade in S è una variabile casuale con media np e varianza npq , con $q = (1 - p)$.

La frequenza m/n è una variabile casuale con media p e varianza pq/n .

Dal Teorema di Bienaymé-Tchebyceff si ha che se X è una variabile casuale con media μ e varianza σ^2 allora

$$(2.1) \quad \text{prob} \{ |X - \mu| < \varepsilon \} \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

Segue che

$$(2.2) \quad \text{prob} \{ |m/n - p| < \varepsilon \} \geq 1 - \frac{pq}{n\varepsilon^2}.$$

La funzione $y = pq = p(1 - p)$ ha il massimo per $p = 1/2$ ed in tale punto assume il valore $1/4$. Di conseguenza risulta che, qua-

lunque sia p ,

$$(2.3) \quad \text{prob} \left\{ \left| m/n - p \right| < \varepsilon \right\} \geq 1 - \frac{1}{4n\varepsilon^2}.$$

La (2.3) ci permette di risolvere vari problemi. Ad esempio, fissato un errore ε e un livello di confidenza α , con $0 < \alpha < 1$ si può determinare un n tale che m/n appartenga all'intervallo $(p - \varepsilon, p + \varepsilon)$ con probabilità non inferiore ad α . Infatti basta trovare un n tale che:

$$1 - \frac{1}{4n\varepsilon^2} \geq \alpha$$

ossia

$$(2.4) \quad n \geq \frac{1}{4\varepsilon^2 (1 - \alpha)}.$$

Ad esempio, per $\alpha = 90\%$ ed $\varepsilon = 0,1$ si ottiene $n \geq 250$.

Le valutazioni ottenute in base alle (2.3) e (2.4) sono in generale pessimistiche, nel senso che si ottiene un errore minore di ε al livello di confidenza α per valori di n in genere molto minori al secondo membro della (2.4).

Una valutazione più precisa si ottiene tenendo conto che m è una variabile binomiale con media $n \cdot p$ e varianza $n \cdot p \cdot q$, per cui per n abbastanza grande la variabile casuale ha distribuzione normale centrata ridotta.

Indicando con $\phi(x)$ la funzione di ripartizione della variabile casuale normale centrata ridotta risulta allora per ogni $k > 0$

$$\text{prob} \left(-k < \frac{m - np}{\sqrt{npq}} < k \right) = \phi(k) - \phi(-k) = 2\phi(k) - 1.$$

Segue

$$\text{prob} \left(-k < \frac{m - np}{n} \cdot \frac{n}{\sqrt{npq}} < k \right) = 2\phi(k) - 1$$

Posto $k \cdot \sqrt{pq/n} = \varepsilon$ si ha

$$\text{prob} \left(-\varepsilon < \frac{m - np}{n} < \varepsilon \right) = 2\phi \left(\varepsilon \cdot \sqrt{n/pq} \right) - 1.$$

Ponendo

$$\text{prob} \left(-\varepsilon < \frac{m - np}{n} < \varepsilon \right) \geq \alpha,$$

Si ottiene

$$2\phi \left(\varepsilon \sqrt{\frac{n}{pq}} \right) - 1 \geq \alpha$$

da cui

$$\phi \left(\varepsilon \sqrt{\frac{n}{pq}} \right) \geq \frac{1 + \alpha}{2},$$

$$\varepsilon \sqrt{\frac{n}{pq}} \geq \phi^{-1} \left(\frac{1 + \alpha}{2} \right)$$

$$(2.5) \quad n \geq \frac{pq}{\varepsilon^2} \left[\phi^{-1} \left(\frac{1 + \alpha}{2} \right) \right]^2.$$

Essendo $pq \leq 1/4$ la (2.5) è verificata per ogni valore di p per

$$(2.6) \quad n \geq \frac{1}{4\varepsilon^2} \left[\phi^{-1} \left(\frac{1 + \alpha}{2} \right) \right]^2.$$

Ad esempio per $\alpha = 0,90$ ed $\varepsilon = 0,1$ si ottiene

$$\phi^{-1} \left(\frac{1 + \alpha}{2} \right) = \phi^{-1} (0,95) = 1,65$$

e quindi dalla (2.6) si ottiene:

$$n \geq 25 \cdot [1,65]^2 = 68.$$

3. METODI PER GENERARE NUMERI A CASO

Uno dei metodi più utilizzati per generare numeri a caso in un intervallo $[a, b]$ è il seguente:

- (1) si fissa un intero positivo m “abbastanza grande” (nelle applicazioni è opportuno che sia almeno dell’ordine di 10^8) e si fissano due numeri interi opportuni h e k appartenenti all’insieme $\{0, 1, 2, \dots, m - 1\}$

- (2) si considera la formula

$$(3.1) \quad x_{r+1} = (hx_r + k) \bmod m,$$

che, assegnato x_r appartenente all’insieme $\{0, 1, 2, \dots, m - 1\}$, permette di calcolare x_{r+1} come resto della divisione di $hx_r + k$ per m .

- (3) a partire da un numero assegnato x_0 , detto seme, si calcolano n numeri $x_1, x_2, \dots, x_{n-1}, x_n$.

- (4) si divide ogni x_i per m ottenendo dei numeri $y_i = x_i / m$, che vengono considerati come numeri a caso nell’intervallo $[0, 1]$;

- (5) dalla formula

$$(3.2) \quad z_i = a + y_i (b - a)$$

si ottengono numeri a caso appartenenti all’intervallo $[a, b]$.

Le proprietà matematiche e statistiche delle successioni ottenute a partire dal generatore (3.1) variano notevolmente al variare di h , k ed x_0 .

Per vari motivi, le scelte più opportune di m sono $m = 10^r$, $m = 2^r$ ed m numero primo.

Si dimostra che le successioni ottenute sono periodiche con periodo δ non superiore ad m ed il massimo periodo possibile $\delta = m$ si ottiene se e solo se sono soddisfatte le seguenti condizioni

- (a) M.C.D. $(h, m) = 1$
- (b) p primo divisore di $m \Rightarrow h \bmod p \neq 1$
- (c) 4 divisore di $m \Rightarrow h \bmod 4 \neq 1$
- (d) M.C.D. $(k, m) = 1$.

Spesso è conveniente porre $k = 0$ ottenendo il generatore congruenziale moltiplicativo

$$(3.3) \quad x_{r+1} = (hx_r) \bmod m.$$

In questo modo si ha il periodo massimo per m numero primo e sotto le condizioni

(a) h ed x_0 sono entrambi primi con m ,

(b) h è una radice primitiva mod m , ossia $h^r \bmod m \neq 1$

per ogni $r \in \{1, 2, \dots, m-2\}$.

ESEMPIO

Consideriamo la (3.3) con m numero primo, ad esempio $m = 41$. Per il Teorema di Legendre (cfr.[2] pag. 72) si ha

$$\left(\begin{array}{c} h \text{ radice primitiva} \\ \bmod 41 \end{array} \right) \Leftrightarrow \left(\begin{array}{c} h^{20} \not\equiv 1 \bmod 41 \\ h^8 \not\equiv 1 \bmod 41 \end{array} \right).$$

I numeri h tali che h è primo con m e $h^{20} \not\equiv 1 \bmod 41$ sono i numeri che non sono quadrati mod 41 ossia tali che non esiste nessun intero positivo tale che $b^2 = h \bmod 41$.

Poiché, rispetto al modulo 41,

$1^2 = 1, 2^2 = 4, 3^2 = 9, 4^2 = 16, 5^2 = 25, 6^2 = 36, 7^2 = 8, 8^2 = 23, 9^2 = 40, 10^2 = 18, 11^2 = 39, 12^2 = 21, 13^2 = 5, 14^2 = 32, 15^2 = 20, 16^2 = 10, 17^2 = 2, 18^2 = 37, 19^2 = 33, 20^2 = 31$ e, per ogni intero positivo s tale che $20 < s \leq 40$ risulta $s^2 = (41-s)^2$, segue che $h^{20} \not\equiv 1 \bmod 41$ per h uguale ad uno dei seguenti numeri:

3 6 7 11 12 13 14 15 17 19 22 24 26 27 28 29 30 34 35 38.

Di questi numeri soddisfano l'uguaglianza $h^8 = 1 \bmod 41$ 3, 14, 27, 38.

I rimanenti 16 numeri sono radici primitive modulo 41 in accordo con la teoria che dice che il numero di radici primitive modulo 41 è $\phi(40) = 40 \cdot 1/2 \cdot 4/5 = 16$, dove con $\phi(m)$ si denota l'indicatore di Gauss.

Consideriamo, per fissare le idee, il caso $h = 7$.

Allora si ha il generatore congruenziale di massimo periodo, 40,

$$(3.4) \quad x_{r+1} = 7x_r \bmod 41.$$

Consideriamo ad esempio il problema di calcolare l'area del triangolo di vertici $(0, 0)$, $(1, 1)$, $(0, 2)$ che è un sottoinsieme dell'intervallo $[0, 2] \times [0, 1]$.

Ponendo $x_0 = 1$, si ottengono, successivamente, i seguenti numeri

1, 7, 8, 15, 23, 38, 20, 17, 37, 13, 9, 22, 31.

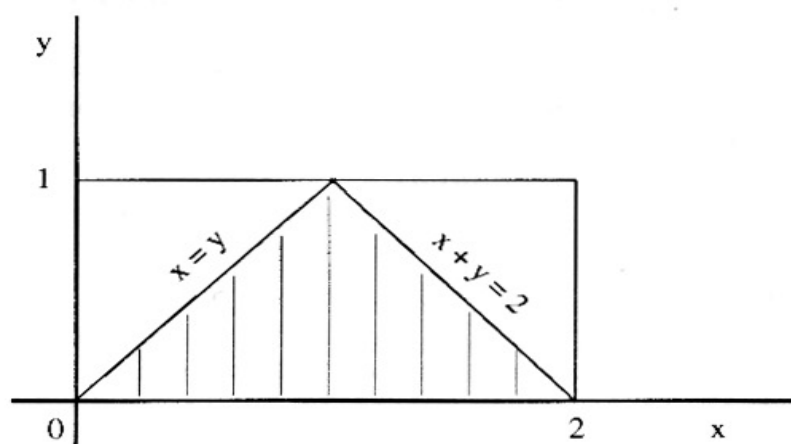
Dividendo per 41, si ottiene

0,024; 0,171; 0,195; 0,366; 0,561; 0,927; 0,458; 0,414; 0,902; 0,317; 0,220; 0,537; 0,756

Considerando, nell'intervallo I i punti

$$P_i = \left(2 \frac{x_i}{41}, \frac{x_{i+1}}{41} \right) \text{ si ottengono}$$

$P_0 = (0,048; 0,171)$, $P_1 = (0,342; 0,195)$, $P_2 = (0,390; 0,366)$,
 $P_3 = (0,732; 0,561)$, $P_4 = (1,122; 0,927)$, $P_5 = (1,854; 0,488)$,
 $P_6 = (0,976; 0,414)$, $P_7 = (0,828; 0,902)$, $P_8 = (1,804; 0,317)$,
 $P_9 = (0,634; 0,220)$, $P_{10} = (0,440; 0,537)$, $P_{11} = (1,074; 0,756)$.



Si vede che sono interni al triangolo i punti $P_1, P_2, P_3, P_6, P_9, P_{11}$, mentre sono esterni al triangolo i punti $P_0, P_4, P_7, P_8, P_{10}$.

Il rapporto $\frac{m}{n} = \frac{\text{punti interni al triangolo}}{\text{numero totale di punti}}$ è esattamente uguale

ad $1/2$, rapporto fra l'area del triangolo e l'area del rettangolo.

BIBLIOGRAFIA

- [1] G. DALL'AGLIO, *Calcolo delle probabilità* (1987), Zanichelli, Bologna.
- [2] G. DI BIASE e A. MATURO, *Sulla verifica di casualità di successioni numeriche in $[0, 1]$ da un punto di vista bayesiano: esempi e considerazioni critiche* (1993), Atti del Convegno Nazionale "Matematica Moderna e Insegnamento: se ne può riparlare?", Ed. Luciani, Roma, pp. 249-256
- [3] A. MATURO, *Numeri Pseudocasuali* (1989), Libreria dell'Università, Pescara.
- [4] R. SCOZZAFAVA, *Introduzione alla Probabilità e alla Statistica* (1984), Veschi Roma.
- [5] G. VARONE, *Generazione di variabili pseudo-probabilistiche per le tecniche di simulazione*, Tesi di Laurea. (1991).

LINGUAGGIO NATURALE, LINGUAGGIO MATEMATICO

Miryam Benvenuto
Liceo Classico "V. Lanza" - Foggia

Una programmazione seria dovrebbe consistere in un lavoro attivo di progettazione che, basandosi sulla specificità culturale e formativa di ciascuna disciplina in quanto particolare forma di conoscenza ed esperienza, consideri la definizione degli obiettivi, la consapevolezza dei metodi di insegnamento/apprendimento, la considerazione dell'allievo nelle sue preliminari abilità, conoscenze e motivazioni, in una progettazione che coinvolga tutte le discipline, per poi andare a definire, nell'ambito di ogni singola disciplina, una scelta di contenuti ed una selezione di materiali e strumenti adeguati.

In questo contesto, la **interdisciplinarietà** non è più una moda, ma diventa una reale esigenza, intesa non solo come sviluppo parallelo di contenuti nei vari ambiti disciplinari, ma come reale confronto fra docenti di diverse discipline sui **metodi** e sul **linguaggio**.

Soffermiamoci appunto sul linguaggio. A volte, diamo per scontato il significato di alcune parole, per noi abbastanza ovvie, non altrettanto, però, per gli studenti. Di qui la necessità di "unificarsi", in qualche modo, sul linguaggio e di partire da quelle che sono le conoscenze pregresse dell'alunno.

Consideriamo, per esempio, il linguaggio delle espressioni. Gli studenti che arrivano dalla scuola media sono spesso molto abili nel risolvere le espressioni con le frazioni, ma, altrettanto spesso, non hanno idea del concetto di frazione e non sanno usare le calcolatrici, soprattutto perché hanno una scarsa dimestichezza con i numeri decimali. Come si giustifica allora la precedente abilità? Il più delle volte non è altro che la consuetudine a ripetere un meccanismo di calcolo: basta introdurre dei numeri decimali, dare un risultato anomalo, perché lo studente cada in crisi.

Proprio partendo dalle espressioni, dopo aver recuperato il concetto di frazione, si può iniziare a parlare di sintassi di una espressione e di come il mancato rispetto di questa sintassi porti ad un cattivo uso della calcolatrice.

La risoluzione di una scheda del tipo seguente:

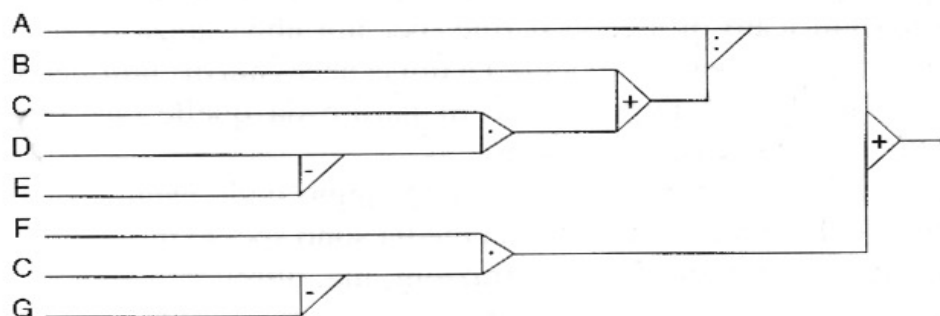
SCHEDA: OPERATORI PER IL CALCOLO

Strumenti: libro di matematica [1]

Rivedere: operatori per il calcolo; schemi di operatori per espressioni.

Esercizi tipo:

- 1) Disegna lo schema di operatori della seguente espressione:
 $9: (6 + 3(4 - 2)) + 1 (3 - 1 / 5)$
- 2) Sostituisci i numeri precedenti con dei numeri a piacere, ripetendo lo stesso numero quando necessario (cfr. il 3 dell'espressione precedente). Qual è lo schema di operatori corrispondente?
- 3) Con ragionamento analogo, sostituisci i numeri con delle lettere. Qual è lo schema di operatori?



conduce inevitabilmente alla considerazione che la struttura di una espressione non dipende dai particolari numeri o lettere presenti, ma dagli operatori e dalle parentesi.

Questa però, non è solo una peculiarità della matematica, come emerge dalla seguente:

**SCHEDA DI VERIFICA INTERDISCIPLINARE: COSÌ È...
SE VI PARE (1).**

Di ogni frase che segue, si possono dare due interpretazioni.
Costruisci l'albero di generazione per ognuna delle interpretazioni date.

ITALIANO: inserisci per ogni frase una diversa punteggiatura, in modo da avere due significati diversi.

- a) Non faccio nulla per guadagnare vivo alla giornata.
- b) Paolo non torna più domani tornerà Andrea

LATINO

- a) Vir sapiens arbitratur minime utilem futurarum rerum scientiam (2).
- b) Ferunt olim immanes beluas in hac regione fuisse (3).

GRECO

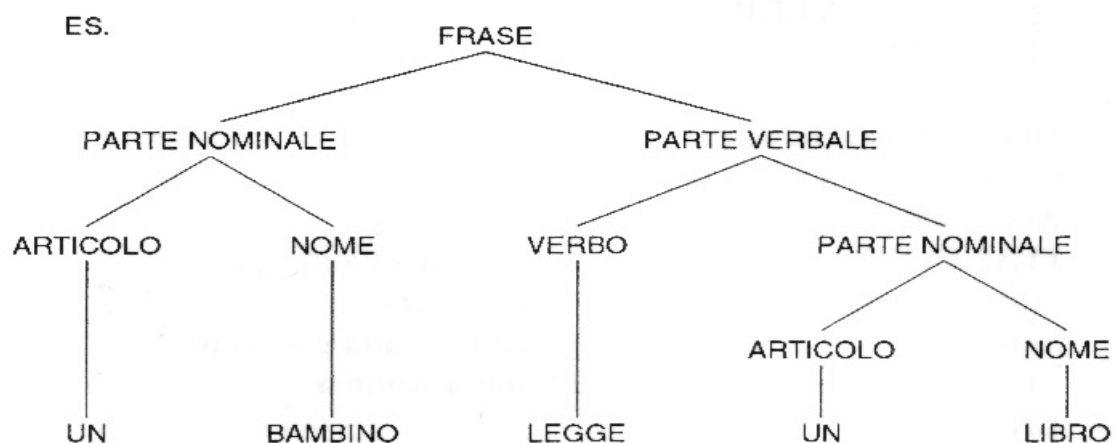
- a) Ἡ γλῶσσα πολλῶν χαχῶν αἰτία ἐδότην.
- b) Νόμιξε τοὺς ἀδελφοὺς τοὺς ἀληθινοὺς φίλους

MATEMATICA: inserisci per ogni frase delle parentesi tonde, in modo da avere due espressioni sintatticamente corrette.

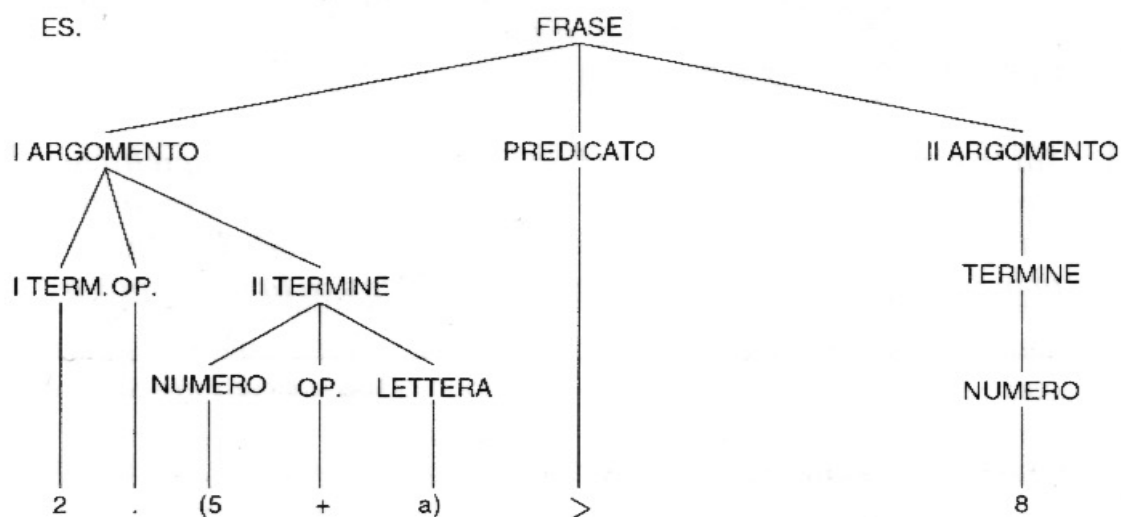
- a) $3 + 14 \times 2 - 14 + 5 - 25 - 1 + 4 \times 2$
- b) $5 \times 3 + 8 - 6 \times 2 + 7 + 4 \times 2$

Sai trovare delle analogie tra l'ambiguità in italiano e quella in matematica? e fra il latino e il greco? Quali sono, secondo te, i rimedi per evitare le ambiguità? Sarebbe possibile l'ambiguità in un linguaggio di programmazione? Motiva la tua risposta

Se si considera, infatti, un albero di generazione per ogni frase, si può notare che l'albero è diverso in corrispondenza di ogni traduzione:



La stessa cosa succede per gli alberi di generazione di "frasi matematiche":



È giusto, quindi, parlare di grammatica e di sintassi anche per il linguaggio matematico:

SCHEDA: GRAMMATICHE E LINGUAGGI

Strumenti: libro di matematica.

Rivedere: grammatiche e linguaggi; grammatiche e diagrammi sintattici.

Esercizio tipo. Analizza le frasi [2]:

"I pirotti carulizzano elasticamente"

" $X \subset Y$ e $Y \subset Z$. Dunque $X \subset Z$ "

Sono corrette dal punto di vista grammaticale? E dal punto di vista sintattico? Se no, prova a sostituire le stringhe e gli operatori usati in modo da avere più informazioni. Fai il diagramma sintattico delle due frasi da te create. (6)

Nella scheda che segue, invece, emerge il problema dell'apprendimento di un altro linguaggio con cui i ragazzi si devono cimentare all'inizio del biennio: un linguaggio di programmazione, dove l'uso di una corretta sintassi diventa requisito indispensabile per la fattibilità di un programma anche semplicissimo.

SCHEDA: MODELLI DI RAGIONAMENTO

Strumenti: libri di matematica ed informatica; computer

Rivedere: proposizioni, condizioni e selezione binaria, implicazione.

Esercizi tipo.

1. le proposizioni da te create nella scheda precedente sono vere o false? Commenta la risposta.
2. trasforma la seconda proposizione della scheda precedente in un'altra "sicuramente" vera, dove X,Y e Z sono variabili intere. Costruisci i relativi diagramma di flusso, diagramma sintattico e programma in turbo pascal, se possibile (7).
3. verifica al computer il precedente programma, correggendone gli eventuali errori sintattici e logici.

Il passaggio alla logica degli enunciati e, quindi, alla verità o falsità di una proposizione, indica che non ci si ferma più solo alla "struttura" di una frase, ma si comincia ad analizzare il "contenuto", oltre che il "contenitore". Del resto, già nella risoluzione della scheda interdisciplinare, in particolare per le traduzioni delle frasi in latino e in greco, è emersa l'importanza del contesto in cui la frase è inserita. Tutti sappiamo come, spesso, una frase, estrapolata dal discorso, può assumere un altro significato.

E per la matematica?

Il problema:

Se tre volte quindici fa cinquantuno, allora quante volte fa sei volte sedici?

non può essere risolto se prima non si specifica la base numerica in cui si opera.

Per trovare la base, si può procedere così:

$15_x = 1x^1 + 5x^0 = x + 5$; $3_x = 3x^0 = 3$; $51_x = 5x + 1x^0 = 5x + 1$;
per cui risolvendo l'equazione $3(x + 5) = 5x + 1$, si ha $x = 7$.

13 moltiplicato 6, in base 7, dà come risultato **114₇**.

Se, però, ci si riferisce alla logica degli enunciati, si può restare in base 10 e considerare solo la struttura logica della frase, del tipo: "Se X, allora Y".

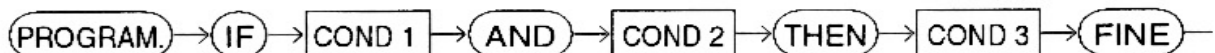
Una simile frase è vera quando X è falsa comunque sia Y (vera o falsa) (8). Per cui, la risposta esatta al quesito posto dal problema può essere **"un numero qualsiasi"**.

La ricerca della soluzione del problema precedente è un semplice esempio di come, anche in matematica, sia necessario chiarire in quale "contesto" si opera.

In questo lavoro, che rappresenta solo l'inizio di un itinerario didattico sperimentale del biennio e qui sintetizzato tramite alcune schede di verifica, la matematica non è vista come astruso esercizio mentale, ma al pari di altre discipline, concorre ad una maggiore comprensione di parole quali **grammatica**, **sintassi**, **semantica** ed acquista dignità culturale nel momento in cui assume proprie grammatiche, proprie sintassi, proprie semantiche.

NOTE

- (1) elaborata insieme alla prof.ssa Carla Mimmo, docente di lettere al ginnasio
- (2) Traduzioni:
- a1. L'uomo saggio crede che non è affatto utile la conoscenza delle cose future.
- a2. L'uomo saggio non crede affatto che è utile la conoscenza delle cose future.
- (3) Traduzioni:
- b1. Dicono che una volta in questa regione c'erano grandissime belve.
- b2. Dicono in questa regione che c'erano grandissime belve
- (4) Traduzioni:
- a1. La lingua è causa di molti mali.
- a2. La lingua di molti è causa di mali.
- (5) Traduzioni
- b1. Considera fratelli gli amici sinceri.
- b2. Considera i fratelli sinceri amici.
- (6) Nel caso che PIROTTI è un sostantivo maschile plurale, CARULIZZANO è voce del verbo CARULIZZARE, ELATICAMENTE è un avverbio, la frase è corretta sia da un punto di vista grammaticale che sintattico. Così, invece, è soprattutto sintatticamente inesistente.
- (7) Esempi: "Le volpi corrono velocemente", " $X \succ Y$ e $Y \succ Z$. Dunque $X \succ Z$ ".



- (8) "ex falso quodlibet". Dal falso può discendere qualsiasi cosa si voglia, dicevano i logici medioevali.

BIBLIOGRAFIA

- [1] W. MARASCHINI, M. PALMA, Matematica di base 1, Paravia, Torino, 1987.
- [2] G. MELZI, Perché la matematica, Ed. La Scuola, Brescia, 1978.

MODELLI GEOMETRICI PER $\mathbf{Z}$, $\mathbf{Q}$.

Raffaele Bruno
Istituto Tecnico Industriale "L. Da Vinci" - Foggia

INTRODUZIONE

Nell'insegnamento della Matematica, alla fine del Biennio delle Superiori o all'inizio del triennio, si sente la necessità di dare, per gli insiemi numerici, una presentazione non banale che non motivi la loro introduzione solo per far sì che determinate operazioni aritmetiche siano sempre possibili, ma che si connetta ad un edificio matematico più saldo e razionale e che sfrutti le moderne concezioni di prodotto cartesiano, di relazione, di isomorfismo, etc.

Una tale presentazione viene fatta su molti testi dell'ultima generazione, ma risulta spesso di difficile comprensione per gli studenti perché appesantita da troppo simbolismo o perché riportata, da testi universitari, senza particolari accorgimenti didattici.

Non voler rinunciare al rigore e nel contempo farsi capire dagli studenti, con l'ambizione ulteriore di ricreare parte della trama disciplinare in cui sono collegati molti concetti matematici, mi ha portato, dopo la lettura del "Manuale dei numeri e delle figure" (Maraschini - Palma / Editori Riuniti 1985), ad una presentazione degli insiemi numerici, scientificamente corretta, supportata da modelli geometrici, e dal gioco.

La Ricerca, a partire da $\mathbf{N}$, si è estesa a $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, ai numeri del calcolatore, ai numeri random ed è stata per il sottoscritto ricca di riflessioni ed approfondimenti.

Nelle pagine seguenti, in maniera molto sintetica (raccolgendo anche la richiesta di colleghi di poterne trarre dei lucidi), vengono introdotti $\mathbf{Z}$ e $\mathbf{Q}$ secondo la linea di ricerca effettuata.

L'insieme $\mathbf{Z}$ può essere costruito, in maniera non banale, a partire da $\mathbf{N}$ mediante la relazione d così definita:

$$\forall (a,b), (c,d) \in \mathbf{N}^2$$

$$(a, b) d (c, d) \Leftrightarrow a + d = b + c$$

d è una relazione di equivalenza.

L'insieme quoziente $\mathbf{N}^2 / d$ viene chiamato $\mathbf{Z}$.

Ogni intero relativo è una classe di equivalenza.

$$(5, 3) d (7, 5) d (23, 21) d (2, 0) \rightarrow +2$$

$$(4, 7) d (10, 13) d (100, 103) d (0, 3) \rightarrow -3$$

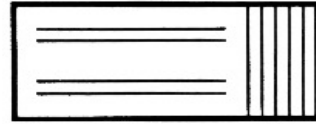
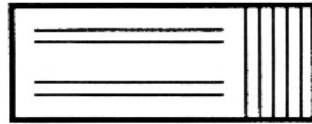
$$\mathbf{Z} = \mathbf{N}^2 / d = \{[0, 0], [1, 0], [2, 0], \dots [0, 1], [0, 2], \dots\}$$

STUDENTI ?

LA RELAZIONE $\leq$ È UNA RELAZIONE DI EQUIDIFFERENZA

IN CLASSE:

Carte di due
semi diversi+due
jolly cui è attribuito il valore zero.



REGOLE DI GIOCO:

RIPETI:

- 1) Estrai una carta da ciascun mazzo;
- 2) Dalla partenza avanza, verso destra, di tante casella di quante ne indica il valore stabilito dalla prima carta e retrocedi, verso sinistra, di quanto ne indica la seconda carta;
- 3) Annota sotto la casella in cui sei giunto la coppia di numeri che ad essa ti ha portato
- 4) Rimetti le carte nei rispettivi mazzi e mischia

FINO ALLO STOP.

UNA POSSIBILE SITUAZIONE:

						Start							
0,6	0,5	0,4	0,3	0,2	0,1	0,0	1,0	2,0	3,0	4,0	5,0	6,0	
1,7	1,6	1,5	1,4	1,3	1,2	1,1	2,1	3,1	4,1	5,1	6,1	7,1	
	2,7	2,6		2,4		2,2	3,2	4,2		6,2		8,2	
	3,8			3,5		7,7		5,3				9,3	
	4,9					9,9							

Invece di due mucchi di 14 carte, se si hanno a disposizione due sacchetti per il gioco della tombola o due urne con tante palline su cui sono disegnati "tutti" i numeri naturali ... ed un tracciato illimitato!
CHE SI PUO' DEDURRE ?

Tutte le coppie (a, b) che portano ad una stessa casella a destra della partenza sono tali che:

$$a - b = \text{cost}$$

e simmetricamente tutte le coppie (a, b) che portano ad una stessa casella a sinistra sono tali che:

$$b - a = \text{cost}$$

e tutte le coppie che portano alla casella di partenza sono tali che:

$$a - b = b - a = 0$$

Le coppie di numeri che portano ad una stessa casella sono equidifferenti.

$(1, 0), (2, 1), (3, 2), \dots (101, 100), \dots$ $\rightarrow Dx(1) \rightarrow +1$

$(2, 0), (3, 1), (4, 2), \dots (108, 106), \dots$ $\rightarrow Dx(2) \rightarrow +2$

$(n, 0), (2n, n), (3n, 2n), \dots$ $\rightarrow Dx(n) \rightarrow +n$

$(0, 1), (1, 2), (2, 3), \dots (126, 127), \dots$ $\rightarrow Sn(1) \rightarrow -1$

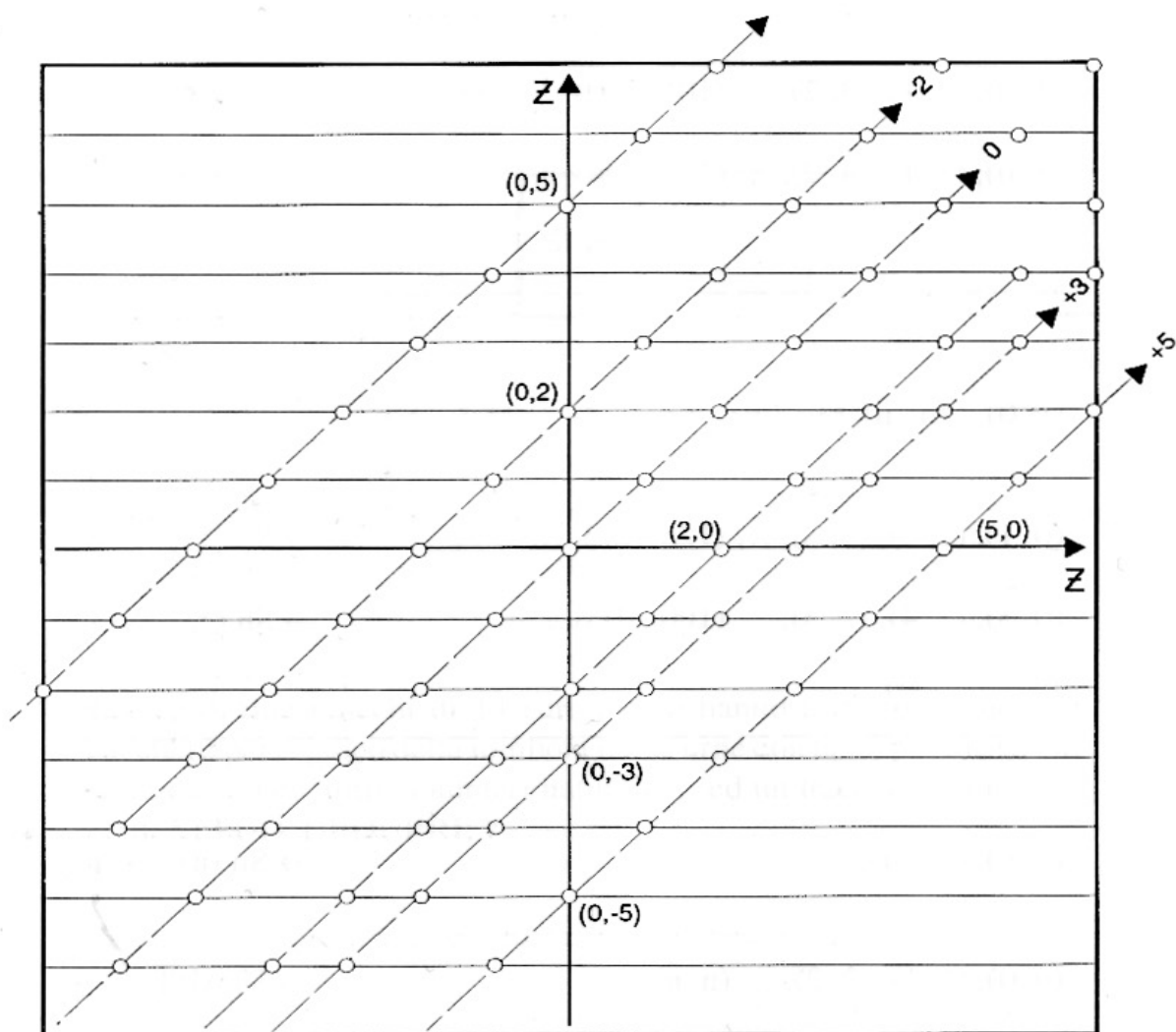
$(0, 3), (1, 4), (2, 5), \dots (100, 103), \dots$ $\rightarrow Sn(3) \rightarrow -3$

$(0, n), (n, 2n), \dots$ $\rightarrow Sn(n) \rightarrow -n$

$(0, 0), (1, 1), (2, 2), \dots (n, n), \dots$ $\rightarrow \text{START} \rightarrow 0$



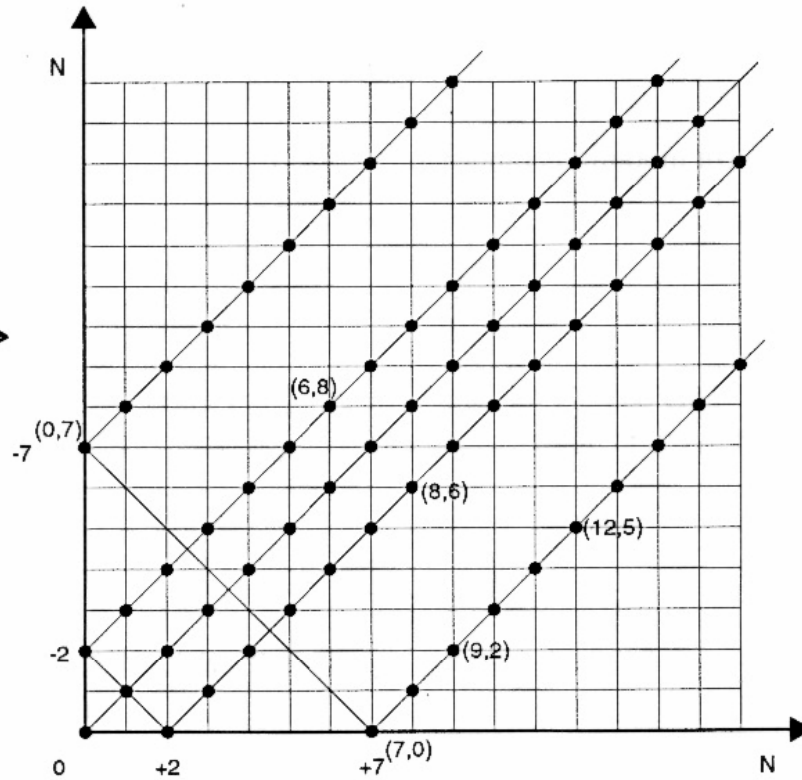
-9 -8 -7 -6 -5 -4 -3 -2 -1 0 +1+2+3+4+5+6+7+8+9+10



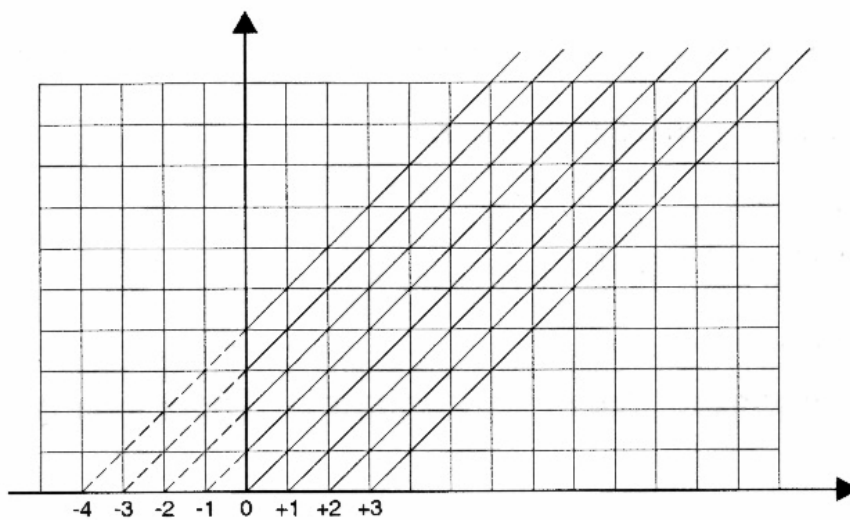
In tale modello, una classe di equivalenza $[x, y]$ risulta formata da punti allineati lungo la direzione della bisettrice I - III quadrante.

I rappresentanti delle varie classi di equivalenza sono le coordinate delle intersezioni delle rette puntiformi con i semiassi positivi.

RAPPRESENTAZIONE CARTESIANA
DELLE COPPIE EQUIDIFFERENTI



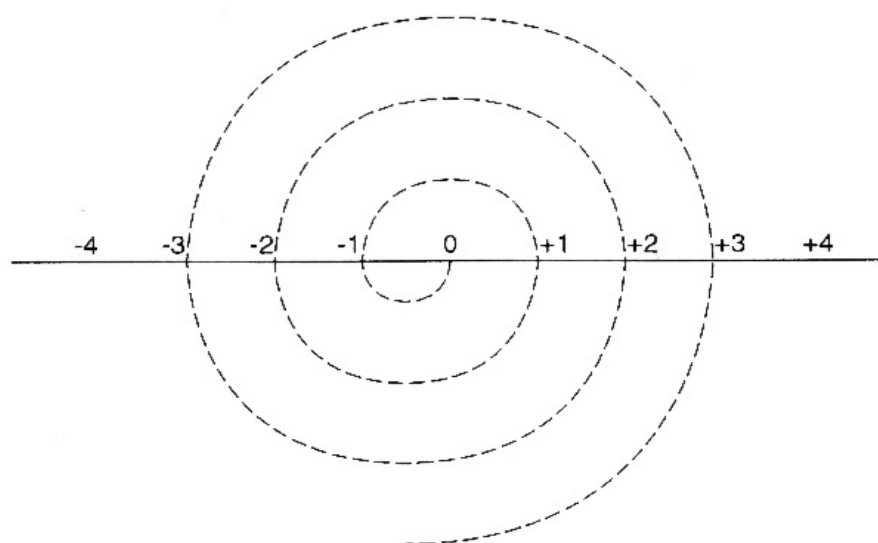
COSTRUZIONE LINEARE DI Z



$\mathbf{Z}$ è equipotente ad $\mathbf{N}$

$$f: n \in \mathbf{N} \mapsto f(n) \in \mathbf{Z} = \begin{cases} n/2 & \text{se } n \text{ pari} \\ -(n+1)/2 & \text{se } n \text{ dispari} \end{cases}$$

Un modello geometrico:



Una tale introduzione di $\mathbf{Z}$ rende meno traumatica per gli studenti anche delle definizioni, in $\mathbf{Z}$ delle operazioni di:

Somma $\forall (a, b), (c, d) \in \mathbf{Z}$
 $(a, b) + (c, d) = (a + c, b + d) \in \mathbf{Z}$

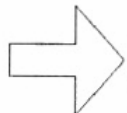
Prodotto $(a, b) * (c, d) = (ac + bd, ad + bc) \in \mathbf{Z}$

Permette la verifica delle proprietà e delle strutture algebriche indotte e di dare una giustificazione non banale della "REGOLA DEI SEGNI".

COSTRUZIONE DI $\mathbb{Q}$ a partire da $\mathbb{Z}$, introducendo la relazione ς

così definita: $\forall (a, b), (c, d) \in \mathbb{Z}^2 \quad b, c \neq 0$
 $(a, b) \varsigma (c, d) \leftrightarrow ad = bc$

ς è una relazione di equivalenza;

 possiamo considerare l'insieme quoziente $\mathbb{Z}^2 / \varsigma$ che chiamiamo $\mathbb{Q}$.

OGNI NUMERO RAZIONALE È UNA CLASSE DI EQUIVALENZA

$$(6, 8) \varsigma (12, 16) \varsigma (-3, -4) \dots \rightarrow [3, 4] \rightarrow 3/4$$

$$(-2, 5) \varsigma (2, -5) \varsigma (6, -15) \dots \rightarrow [-2, 5] \rightarrow -2/5$$

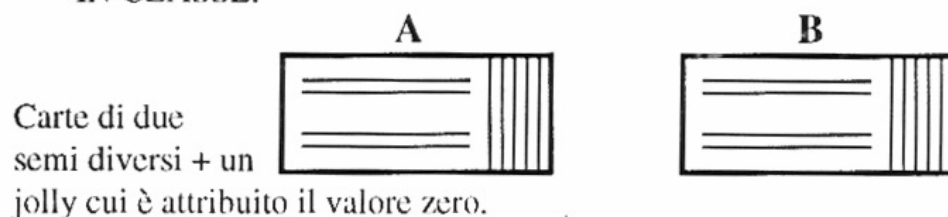
$$(2, 4) \varsigma (-6, -12) \varsigma (10, 20) \dots \rightarrow [1, 2] \rightarrow 1/2$$

$$\mathbb{Q} = \mathbb{Z}^2 / \varsigma = \{[a, b] \mid (b \neq 0) \ (a, b) \in \mathbb{Z}^2 \wedge a \text{ primo con } b\}$$

LA RELAZIONE È UNA RELAZIONE DI EQUIRAPPORTO

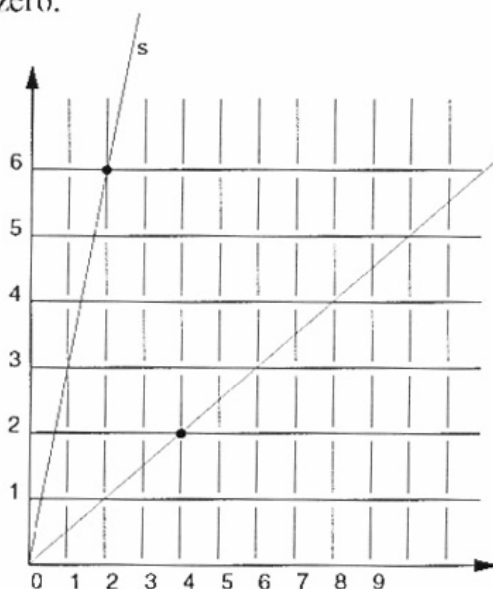
LA RELAZIONE ζ È UNA RELAZIONE DI EQUIRAPPORTO:

IN CLASSE:



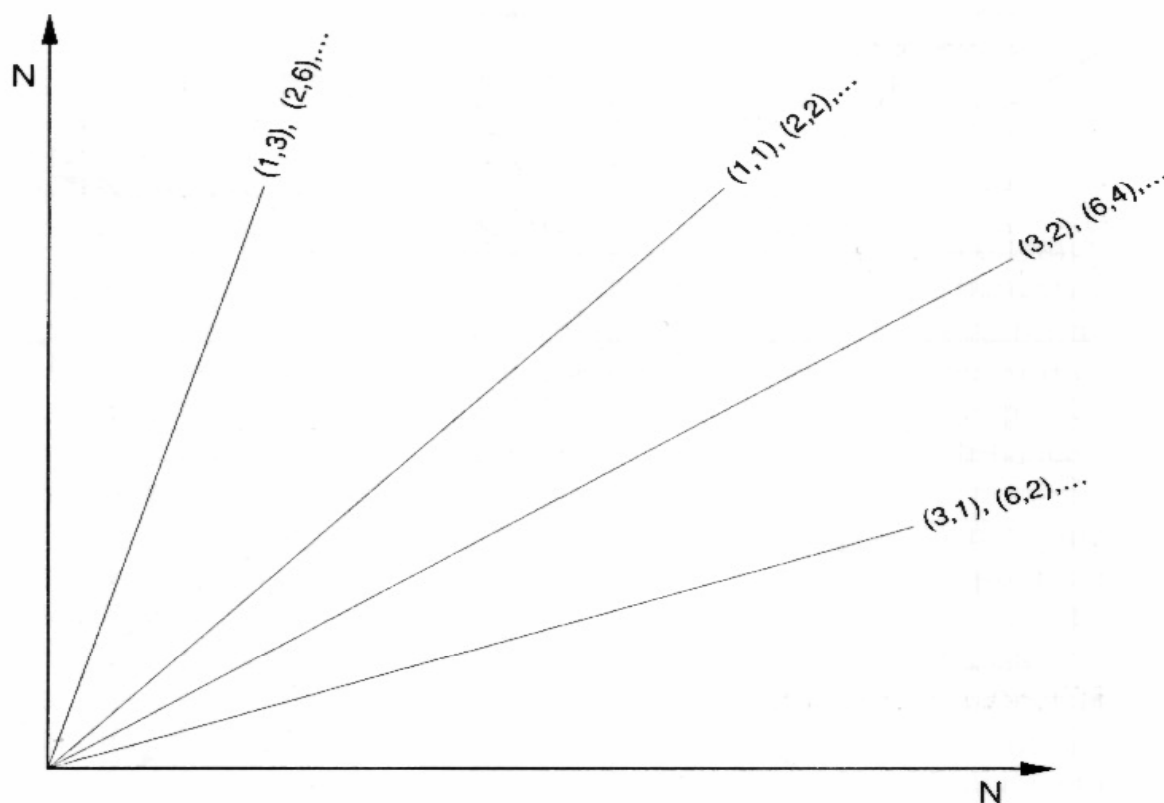
REGOLE DI GIOCO:

RIPETI:



- 1) Estrai una carta dal mucchio A e una dal mucchio B;
- 2) Dalla partenza avanza, in orizzontale, di tanti posti di quanto è il valore a della prima carta e da tale punto, avanza, in verticale di tanti posti di quanto è il valore b della seconda carta;
- 3) Segna il punto in cui sei giunto;
- 4) Disegna una semiretta dall'origine a tale punto;
- 5) Annota, dopo aver dato un nome alla semiretta, la coppia di numeri che ti ha permesso di disegnarla;
- 6) Metti le carte sui due rispettivi mucchi e mischia;
- 7) FINO ALLO STOP DEL DOCENTE.

**CI SONO PIÙ COPPIE DI NUMERI CHE TI FANNO
DISEGNARE LA STESSA SEMIRETTA ?**

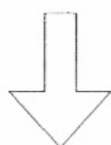


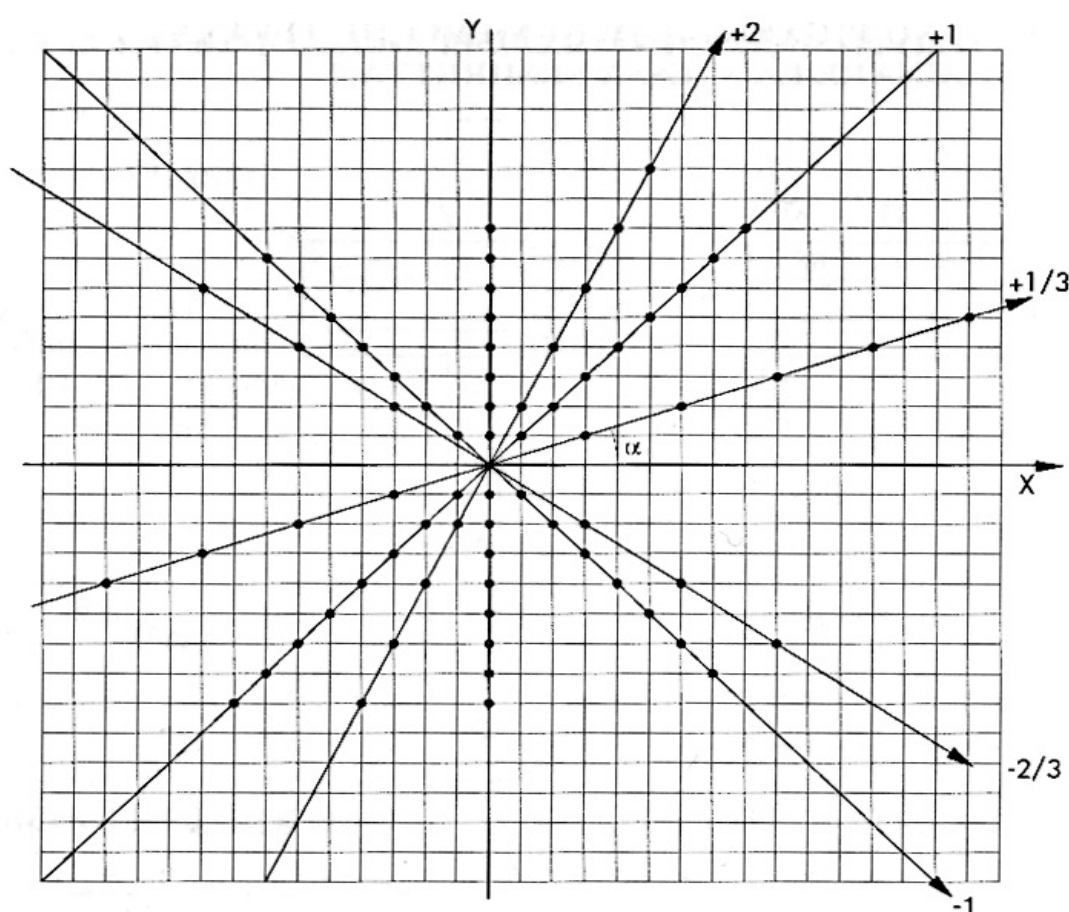
Estendendo il gioco a $\mathbb{N} \times \mathbb{N}$...

- * Ogni Coppia (a, b) porta ad una sola semiretta;
- * due coppie $(a, b), (c, d)$ individuano una stessa semiretta se:

$$a : b = c : d \rightarrow ac = bd$$

Estendendo il gioco (con operazioni di simmetria) a $\mathbb{Z} \times \mathbb{Z}$





COPPIE DELLA STESSA CLASSE HANNO PER IMMAGINE PUNTI ALLINEATI LUNGO UNA DIREZIONE PASSANTE PER L'ORIGINE DEGLI ASSI

* I punti dell'asse delle ascisse non sono immagine di alcuna coppia (a, b) dovendo essere $b = 0$.

Un numero $Q + \leftrightarrow$	la retta puntiforme forma un angolo $\alpha < 90^\circ$
Un numero $Q - \leftrightarrow$	" " " " $90^\circ < \alpha < 180^\circ$
Un numero $= \leftrightarrow$	" " " " $\alpha = 90^\circ$

STUDIO DI UNA POLIZZA ALLO SPORTELLLO

*Filippo Di Leonardo * Giuseppina Pagliuca ***

** Cassa Rurale ed Artigiana di Castignano e Rotella (AP)*

*** Dipartimento di Metodi Quantitativi e Teoria Economica
Università "G. D'Annunzio" viale Pindaro - Pescara*

1. INTRODUZIONE

A seguito della grande diffusione, riscontrata negli ultimi tempi, di prodotti assicurativi distribuiti attraverso gli Sportelli Bancari e con lo scopo di studiarne la convenienza rispetto ad altre forme assicurative "similari" e più tradizionali già in circolazione, si tratterà in questo lavoro di una polizza assicurativa, tra le tante, proposta dalla Assimoco Vita S.p.A. attraverso tutti gli sportelli bancari delle Casse Rurali ed Artigiane Italiane: la PREVI.C.R.A.

Nel mare magnum delle polizze vita offerte dalle Compagnie Italiane e straniere operanti nel nostro Paese, la polizza PREVI.C.R.A. si colloca su buoni livelli di economicità e rendimento, e presenta un ampio ventaglio di possibilità per l'assicurato, connesse tra l'altro, al riscatto anticipato ed alla opzione per la rendita finale. Ma la vera novità di questo "Servizio Assicurativo", che sta già rivelando la sua forza commerciale nel Nord Italia, sono altre: la semplicità e la duttilità. Semplicità e duttilità che tra l'altro investono, trovandoli complici, anche gli addetti alla distribuzione, che nella fattispecie sono bancari e non assicuratori e pertanto posseggono meno padronanza, dal punto di vista tecnico, di tali prodotti finanziari. L'idea di presentare una polizza che fosse di facile comprensione e che quindi si potesse proporre e vendere meglio, sembra una scelta vincente, anche se non sono disponibili dati statistici a questo riguardo, studiata da esperti di marketing, prima ancora che dai tecnici finanziari che hanno poi creato questa nuova "polizza allo sportello": la PREVI.C.R.A..

Tecnicamente si tratta di una assicurazione di capitale differito a premio annuo medio costante, con controassicurazione a prestazione rivalutabile. I premi sono a taglio fisso (mensile o annuale), per una durata contrattuale di 10 anni e validi per ogni età d'ingresso

compresa tra i 18 e 55 anni: la particolarità sta nel fatto che il premio da corrispondere è indipendente dall'età dell'assicurato, essendo comunque calcolato per una testa di 40 anni di età.

È importante sottolineare che le prestazioni del contraente proposte dalla Compagnia in combinazioni fisse, sono determinate, tenendo conto, oltre che dei criteri demografico-finanziario, anche di aspetti psicologici e sociali. È alquanto probabile che l'Istituto si sia preoccupato anche di creare un prodotto che, dal punto di vista dell'entità del costo, fosse ben collocabile sul mercato e contemporaneamente fosse facilmente percepito dal pubblico. Ciò facendo sono stati certamente privilegiati gli aspetti prettamente finanziari nonostante l'esigenza di proporre un prodotto assicurativo.

Per snellezza di trattazione, si omettono disquisizioni su alcune condizioni generali e speciali di polizza, questioni legali e fiscali e quant'altro non sia di precipuo interesse di questo lavoro: per queste ed altre considerazioni, si rimanda alla bibliografia.

La presente analisi rivolgerà invece l'attenzione alla particolare tipologia di questo nuovo prodotto, anche confrontandolo con uno tradizionale, per trovare i segreti della sua peculiarità, facendone trasparire le caratteristiche marcatamente finanziarie, che forse mortificano le radici assicurative.

2. - PREVI.C.R.A.: descrizione del contratto e modello.

Nel presente paragrafo ci occuperemo della presentazione delle clausole del prodotto assicurativo in esame e della costruzione del modello matematico che, molto verosimilmente, ha condotto l'Ente alla formazione delle basi contrattuali con lo scopo precipuo di effettuare un confronto successivo con il modello "classico" di una assicurazione di capitale differito con controassicurazione.

Il prodotto PREVI.C.R.A. prevede un contratto che si articola come segue:

- il premio annuo costante, pagato anticipatamente, è determinato in modo indipendente dal sesso e dall'età del contraente. Tale premio è comunque quello riferito ad una testa di 40 anni di età. Le età di ingresso variano da 18 a 55 anni;

- la durata è fissata in 10 anni. Alla scadenza di tale termine il contraente può scegliere tra il pagamento di un capitale, in unica soluzione, o di una rendita annua rivalutabile;
- in caso di morte dell'assicurato prima della scadenza del contratto, i beneficiari designati dal contraente, potranno recuperare i premi versati rivalutati.

Inoltre il contraente, nonostante la durata fissa del contratto, dispone di altre possibilità:

- riscattare la polizza già al termine del primo anno, pagando una penalizzazione dello 0,25% per ogni anno mancante al decimo;
- interrompere il piano dal secondo al decimo anno ed attendere la scadenza del contratto scegliendo il capitale o la rendita rivalutati e commisurati ai premi versati.

I premi pagati dall'assicurato, dopo aver dedotto le imposte e le commissioni previste, vengono investiti in un apposito fondo "PREVI.C.R.A." costituito per la maggior parte da obbligazioni e titoli di Stato. PREVI.C.R.A. riconosce ogni anno almeno l'85% del rendimento conseguito dalla gestione del fondo, garantendo un rendimento minimo del 4%.

Analizzate le clausole essenziali del contratto, studiamo ora gli aspetti tecnico-finanziari di tale prodotto assicurativo.

Per brevità di esposizione, approfondiremo solo il caso in cui il contraente opti per il pagamento di un capitale finale.

2.1. IL CONTRATTO PRINCIPALE

Trattandosi di una assicurazione di capitale differito con controassicurazione a premio annuo costante, la Compagnia si impegna a corrispondere ad un individuo, con età compresa tra 18 e 55 anni, un determinato capitale se alla scadenza del contratto, dopo n anni ($n = 10$), l'assicurato sarà in vita (assicurazione principale), altrimenti, in caso di decesso durante il differimento, verrà corrisposto ai suoi eredi un capitale commisurato al cumulo dei premi versati rivalutati al netto delle tasse (controassicurazione).

Il capitale assicurato a scadenza si forma attraverso la rivalutazione annuale delle somme versate.

Il capitale rivalutato anno per anno e relativo al t -esimo periodo, fermo restando l'ammontare annuo del premio, viene determinato sommando al capitale in vigore nel periodo annuale precedente, (il $(t - 1)$ -esimo):

- i) un importo ottenuto rivalutando il capitale C_0 inizialmente assicurato e riducendolo nella proporzione in cui il numero degli anni trascorsi sta al numero degli anni di durata del differimento;
- ii) un ulteriore importo ottenuto rivalutando la differenza tra il capitale in vigore nel periodo annuale precedente e quello inizialmente assicurato.

Pertanto, secondo le clausole di PREVI.C.R.A., la relazione ricorrente per il capitale rivalutato anno per anno è:

$$C_t = C_{t-1} + C_0 \omega r \frac{1}{n} + \omega r (C_{t-1} - C_0) \quad (t + 1, 2, \dots, n) \quad (1)$$

da cui

$$C_t = C_{t-1} (1 + \omega r) - C_0 \omega r (1 - \frac{1}{n})$$

dove

C_t è il capitale rivalutato all'epoca t ;

C_0 è il capitale assicurato inizialmente;

ω_r è la misura della rivalutazione, essendo w ed r rispettivamente il tasso di retrocessione ed il tasso di rendimento del fondo;

n è la durata del differimento.

Il capitale C_0 inizialmente assicurato e che costituisce il punto di partenza per la rivalutazione annuale, è stato ottenuto moltiplicando il premio netto di tariffa, ovvero il premio effettivamente versato dal contraente al netto delle tasse e al lordo del caricamento (indicato con $P_N^{(T)}$ per un fattore h costante per tutte le età e per tutti gli importi. Tale fattore è ravvisabile nel valore attuale medio

di una rendita vitalizia anticipata, immediata, temporanea per 10 anni su una testa $x = 40$.

Si ha pertanto

$$C_0 = P_N^{(T)}$$

La (1), dopo semplici passaggi, può essere scritta come segue:

$$C_t = C_0(1 + \omega r)^t \left[1 - \omega r \sum_{k=1}^t 1 + \omega r^{-k} \left(1 - \frac{k}{n} \right) \right]$$

o meglio

$$C_t = C_0 \left[1 + \frac{\omega r}{n} (I_s)_{t|\omega r}^- \right] \quad (2)$$

dove $(I_s)_{t|\omega r}^-$ rappresenta il montante di una rendita a rate variabili in progressione aritmetica in cui è unitario il primo termine e unitaria anche la ragione.

Data la possibilità dell'assicurato di riscattare la polizza prima della scadenza del contratto è interessante vedere come si determina il capitale riscattabile ad ogni epoca intermedia.

L'importo del riscatto si ottiene pagando una penalizzazione pari a $i = 0,25\%$ per ogni anno mancante al termine del differimento e valutata sul capitale ridotto relativo alla stessa epoca:

$$C_{RISC_t} = C_{RID_t}(1 + i)^{t-n}$$

Il capitale ridotto viene calcolato sommando i seguenti due importi:

- il prodotto del capitale inizialmente assicurato per il rapporto tra il numero dei premi pagati, comprese le eventuali frazioni di anno, ed il numero dei premi pattuiti;
- la differenza tra il capitale rivalutato all'ultima ricorrenza di polizza ed il capitale inizialmente assicurato

da cui

$$C_{RISC_t} = C_t - C_0 \left(1 - \frac{t}{n} \right).$$

È ovvio che, nel caso in cui la polizza sia riscattata al termine del contratto, l'importo del riscatto coincide sia con il capitale ridotto, sia con il capitale rivalutato alla medesima epoca come peraltro si evidenzia nella tabella che segue. Viene utilizzata una misura di rivalutazione pari al 10,74% quale quella ipotizzata dalla Compagnia sulla base del rendimento medio dei titoli di Stato e delle obbligazioni nel periodo ventennale dal 1971 al 1990, secondo le prescrizioni e la misura di riduzione fissate dall'ISVAP e un premio fisso pari a £. 2.000.000.

Tab. 1.

EPOCHE	CAP. RIVALUTATO	CAP. RIDOTTO	CAP. DI RISCATTO
0	17.970.009		
1	18.163.006	1.989.999	1.945.778
2	18.569.730	4.193.723	4.110.785
3	19.213.134	6.634.128	6.519.183
4	20.118.637	9.336.632	9.197.800
5	21.314.389	12.329.385	12.176.417
6	22.831.563	15.643.560	15.488.097
7	24.704.680	19.313.677	19.169.546
8	26.971.966	23.377.965	23.261.512
9	29.675.758	27.878.757	27.809.234
10	32.862.934	32.862.934	32.862.934

Dalla (2), per $t = n$, si ha:

$$C_{RISC_n} = C_n = C_0 \left[1 + \frac{\omega r}{n} (I_s)_n | \omega r \right]$$

che rappresenta l'importo del riscatto a scadenza ovvero il capitale rivalutato al termine del differimento.

2.2. LA CONTROASSICURAZIONE

In caso di premorienza, la controassicurazione consiste, come già accennato, nel corrispondere agli eredi un importo uguale al premio netto pattuito, moltiplicato per il numero dei premi annui pagati e rivalutati nella proporzione in cui il capitale, quale risulta rivalutato all'epoca che precede la data del decesso, sta a quello inizialmente assicurato.

Pertanto il capitale pagabile alla fine dell'anno in cui avviene la morte sarà pari a $P_N^{(T)}$, oppure $2 P_N^{(T)} \frac{C_1}{C_0}$, oppure $2 P_N^{(T)} \frac{C_2}{C_0}$,

a seconda che l'individuo di età x muoia tra le età x e $x + 1$, oppure $x + 1$ e $x + 2$, oppure $x + 2$ e $x + 3$ e così via.

Stabiliti tutti i termini del contratto, calcoliamo il premio puro dell'assicurazione in esame:

$$P_{/n} a_x = C_{n/n} E_x + \sum_{k=1}^n k P_N^{(T)} \frac{C_{k-1}}{C_0} {}_{k-1/1} A_x \quad (3)$$

dove

k è il numero dei premi pagati

$P_N^{(T)}$ è il premio di tariffa al netto delle tasse ma al lordo del caricamento

C_{k-1} è il capitale rivalutato all'epoca precedente la data del decesso

${}_n E_x$ è, come noto, il fattore di sconto demografico-finanziario

${}_{k-1/1} A_x$ è il valore attuale medio di un capitale unitario pagato alla fine dell'anno in cui avviene il decesso di una testa x , se esso avviene tra le età $x + k - 1$, $x + k$.

Conglobando, per semplicità, il caricamento nell'unica aliquota β del premio netto di tariffa, si ha:

$$P_N^{(T)} = P + \beta P_N^{(T)} \quad \text{da cui} \quad P_N^{(T)} = \frac{P}{1 - \beta} \quad (4)$$

dalla (3), tenendo presente la (4), si ricava l'espressione del premio puro

$$P = \frac{C_n \cdot {}_nE_x}{{}_n a_x - \frac{1}{1-\beta} \sum_{k=1}^n k \frac{C_{k-1}}{C_0} {}_{k-1/1} A_x}. \quad (5)$$

Il secondo membro della (5), calcolato per un individuo di età $x = 40$ anni, fornisce esattamente il taglio fisso proposto dalla Compagnia.

È inoltre necessario precisare che il capitale a scadenza C_n , è determinato secondo una misura di rivalutazione minima, pari al 4%.

Sulla base di semplici calcoli abbiamo provato che, se la Compagnia adottasse la (5) per calcolare i premi effettivamente pagati dall'assicurato per ciascuna età, lo scarto tra il premio versato da un diciottenne e quello versato da un cinquantacinquenne, sarebbe irrilevante tanto da rendere accettabile il premio fisso da parte di tutti gli individui con età compresa nella fascia 18-55 anni. Tale lieve differenza, viene compensata dalla maggiore facilità di presentazione del prodotto che risulta più comprensibile anche per il cliente.

Il premio lordo effettivamente corrisposto dal contraente è:

$$P_L^{(T)} + P_N^{(T)} + \alpha P_N^{(T)}$$

ed inoltre, essendo $P_N^{(T)} = \frac{P}{1-\beta}$, risulterà:

$$P_L^{(T)} = P \frac{1+\alpha}{1-\beta} \quad (6)$$

Riportiamo nella Tab. 2 di pagina seguente gli ammontari dei premi lordi calcolati secondo la (6) e tenendo presente la (5), avendo assunto per α e β rispettivamente i valori 2,5% e 7,792%.

Tab. 2.

ETÀ	PREMIO	ETÀ	PREMIO
18	1.998.122	37	1.999.397
19	1.998.126	38	1.999.580
20	1.998.130	39	1.999.781
21	1.998.136	40	<u>2.000.000</u>
22	1.998.147	41	2.000.242
23	1.998.164	42	2.000.513
24	1.998.186	43	2.000.813
25	1.998.215	44	2.001.140
26	1.998.252	45	2.001.499
27	1.998.297	46	2.001.884
28	1.998.353	47	2.002.303
29	1.998.419	48	2.002.765
30	1.998.494	49	2.003.273
31	1.998.581	50	2.003.845
32	1.998.678	51	2.004.483
33	1.998.791	52	2.005.201
34	1.998.921	53	2.006.005
35	1.999.065	54	2.006.892
36	1.999.225	55	2.007.871

3. CONFRONTO CON UNA POLIZZA TRADIZIONALE DELLO STESSO TIPO.

Consideriamo ora un contratto classico di assicurazione di capitale differito con controassicurazione. Tale contratto prevede il pagamento, da parte dell'Istituto assicuratore, di un capitale ad un individuo di età x , se è in vita alla scadenza cioè all'età $x + n$, o il rimborso agli eredi, in caso di premorienza dei premi pagati. Pertanto, l'assicurazione principale è quella di capitale differito, mentre la controassicurazione è un'assicurazione temporanea caso morte.

Nel caso in cui il contraente versi dei premi annui, la controassicurazione comporta il rimborso dei premi di tariffa $\bar{P}^{(T)}$ pagati. Precisamente, la Compagnia rimborsa $\bar{P}^{(T)}$, se la morte accade

accade nel primo anno, $2\bar{P}^{(T)}$, se la morte accade nel secondo, $3\bar{P}^{(T)}$ se accade nel terzo e così via.

Si tratta quindi di una assicurazione di morte a capitale variabile in progressione aritmetica di primo termine e ragione $\bar{P}^{(T)}$.

Se il premio è pagabile per tutta la durata del contratto si ha l'equazione:

$$\bar{P}_{/n} a_x = C_n E_x + \bar{P}^{(T)}_{/n} (IA)_x \quad (7)$$

avendo indicato con $\bar{P}$ il premio puro di detta assicurazione. La (6) può essere scritta come segue:

$$\bar{P}_{/n} a_x = C_n E_x + \bar{P}^{(T)} \sum_{k=1}^n k_{k-1/1} A_x \quad (8)$$

Per analogia al precedente contratto, si ha: $\bar{P}^{(T)} = \bar{P} + \beta \bar{P}^{(T)}$

da cui $\bar{P}^{(T)} = \frac{\bar{P}}{1 - \beta}$, dalla (8) ricaviamo il premio puro

$$\bar{P} = \frac{C_n E_x}{_{/n} a_x - \frac{1}{1 - \beta} \sum_{k=1}^n k_{k-1/1} A_x} \quad (9)$$

Per effettuare il confronto con il contratto di assicurazione PREVI.C.R.A., ipotizziamo, per esigenze di omogeneità, un capitale a scadenza pari a quello determinato dalla Compagnia ASSI-MOCO e confrontiamo i premi puri relativi ai due contratti dati rispettivamente dalla (5) e dalla (9).

Dal confronto, si evince innanzi tutto la diversità riguardo alla controassicurazione che nella polizza PREVI.C.R.A. si configura in modo atipico in quanto prevede pur sempre il rimborso dei premi di tariffa pagati dal contraente ma rivalutati.

Inoltre, poichè ovviamente si ha $C_{t-1} \geq C_0 \quad \forall t = 1, 2, \dots, n$, risulta $P > \bar{P}$ per ogni età come si evidenzia nella tabella di pagina seguente.

ETÀ	POLIZZA CLASSICA	POLIZZA ASSIMOCO
18	1.996.871	1.998.122
19	1.996.863	1.998.126
20	1.996.854	1.998.130
21	1.996.839	1.998.136
22	1.996.816	1.998.147
23	1.996.785	1.998.164
24	1.996.742	1.998.186
25	1.996.690	1.998.215
26	1.996.624	1.998.252
27	1.996.543	1.998.297
28	1.996.444	1.998.353
29	1.996.328	1.998.419
30	1.996.198	1.998.494
31	1.996.052	1.998.581
32	1.995.890	1.998.678
33	1.995.704	1.998.791
34	1.995.491	1.998.921
35	1.995.257	1.999.065
36	1.994.999	1.999.225
37	1.994.723	1.999.397
38	1.994.433	1.999.580
39	1.994.119	1.999.781
40	1.993.781	<u>2.000.000</u>
41	1.993.410	2.000.242
42	1.992.993	2.000.513
43	1.992.528	2.000.813
44	1.992.018	2.001.140
45	1.991.457	2.001.499
46	1.990.859	2.001.884
47	1.990.213	2.002.303
48	1.989.508	2.002.765
49	1.988.736	2.003.273
50	1.987.869	2.003.845
51	1.986.898	2.004.483
52	1.985.803	2.005.201
53	1.984.572	2.006.005
54	1.983.218	2.006.892
55	1.981.734	2.007.871

4. CONSIDERAZIONI FINALI

Dallo studio effettuato nei paragrafi precedenti, la Previcra è risultato un prodotto singolare di indubbio interesse in campo assicurativo-finanziario.

Dal confronto con una polizza tradizionale dello stesso tipo, abbiamo riscontrato il maggiore onere della Previcra, per qualunque età dell'assicurato. Ciò tuttavia non significa una minore competitività economica, perchè si tratta di prodotti non perfettamente collimanti, in quanto, in caso di premorienza, la controprestazione assicurativa della Previcra è più gratificante. Pertanto la minore economicità del prodotto analizzato, è compensato da altri fattori.

La caratteristica più importante di questa polizza allo sportello, che prescinde dagli aspetti economici e legali, molto rilevanti per il contraente all'atto della stipula di una polizza tradizionale, è costituita dalla veste finanziaria.

A partire dagli importi predeterminati dei premi mensili o annuali, che rappresentano chiaramente delle soglie psicologiche di "risparmio" e con la rivalutazione dei premi recuperati in caso di premorienza, tutti gli aspetti di questo prodotto fanno pensare ad un investimento finanziario: non a caso, chi sottoscrive la Previcra in uno degli sportelli bancari delle Casse Rurali ed Artigiane Italiane, discute sul "quanto renda" questa polizza.

La forza di questo prodotto, che porta a trascurare il suo limite principale, ovvero la sua minore competitività economica, è quella di garantire un certo rendimento legato agli andamenti della Borsa ed in generale del mercato finanziario: andamenti che determinano il rendimento del Fondo Assicurativo e quindi il rendimento retrocesso al cliente, oltre al recupero fiscale legato alle normative vigenti (attualmente il 27%). Il tutto esattamente come una qualunque forma di investimento proposta in Banca, ad esempio certificati di deposito, pronti contro termine, titoli di Stato.

In effetti è scomparso quasi completamente il quadro psicologico legato alla assicurazione: con la Previcra il verificarsi di un evento aleatorio, quale la premorienza dell'assicurato, è, rispetto ad una assicurazione tradizionale dello stesso tipo, praticamente irrilevante.

Alla luce pertanto dei fatti suesposti, la Previcra appare innanzi tutto come un prodotto finanziario prima che assicurativo; pertanto dobbiamo ragionevolmente ritenere che la buona diffusione avutasi di questa polizza allo sportello fino ad oggi, si sia avuta, almeno in parte, a scapito di prodotti bancari tradizionali e Titoli del debito Pubblico, piuttosto che di polizze assicurative tradizionali, che ad una prima analisi erano sembrate i prodotti succedanei concorrenti.

BIBLIOGRAFIA

- [1] B.P.A.M., 1992 - Circolari interne.
- [2] CERQUETTI - DI LEONARDO, 1992 - " *Convenienza di Polizze Assicurativo-Bancarie* " - Pubblicazione del Dipartimento di Metodi Quantitativi e Teoria Economica.
- [3] E. LEVI, 1964 - " *Corso di matematica finanziaria e attuariale* " - Ed. Giuffrè.
- [4] G. OTTAVIANI, 1984 - " *Lezioni di Matematica Finanziaria* ".
Veschi - Roma.

I PROCESSI COGNITIVI DELL'ARITMETICA IN RUSSEL E PIAGET

Silvana D'ANDREA
USSL Giulianova (TE)

1. INTRODUZIONE

Beltrand Russel (1872-1970) filosofo inglese dai vastissimi interessi culturali, attento e profondo critico dei fondamenti della Matematica e Jean Piaget (1896-) psicologo svizzero e profondo elaboratore dei meccanismi preposti alla base del pensiero e del linguaggio del fanciullo: due mostri sacri del pensiero contemporaneo. Entrambi si sono occupati di Matematica e precisamente del come nasce il concetto di numero. Il primo se ne è occupato guardando la genesi del numero nella mente dell'uomo in una *critica a posteriori e dall'interno* della disciplina matematica. In altre parole Russel si pone dal punto di vista di una persona di vasta cultura matematica e il suo obiettivo è quello di ricostruire per intero l'edificio matematico su basi puramente logico deduttive e sostanzialmente prescindendo dall'individuo. Tuttavia egli si allontana dalle idee della scuola di Bourbaki, in quanto desidera che i suoi metodi abbiano una qualche attinenza con il reale, al contrario dei Bourbakisti che non si preoccupano e volutamente ignorano il reale.

L'approccio del Piaget è totalmente diverso essendo la sua una *critica dall'esterno* della matematica e dall'interno dell'uomo anzi del fanciullo. Questo punto di osservazione, non ci sembra contrapposto al Russel ma differente. Piaget parte dunque dall'osservazione del reale e dell'individuo nella sua tenera età e va a studiarne i meccanismi mentali e le relative conquiste, tra cui il concetto di numero.

La nostra tesi è semplice. Confrontando le strade seguite si ottiene un risultato interessante. I termini dell'uno sono diversi da quelli dell'altro ma i meccanismi con i quali si crea il concetto sono i medesimi. La via di Russel a posteriori per riforgiare l'edificio coincide con quella via che il Piaget segnala come il percorso pedagogico che naturalmente si crea nel passare dall'osservato naturale al momento più rigido del conseguito momento di apprendimento

realizzato. Così dalla corrispondenza termine a termine si passa al riconoscere le classi di equivalenza o se si preferisce alle seriazioni Piagetiane e così via fino alla cardinalità e all'operatività.

2. RUSSEL E PIAGET, IL NUMERO E IL BAMBINO

Il lavoro fondamentale del Piaget riguardo l'apprendimento della Matematica da parte del bambino è l'opera "La genesi del numero nel bambino" del 1947, riproposto da svariati punti di vista negli anni successivi, precisamente nel '53, nel '55, nel '64. Dice il Piaget (Szemiska, 1941):

"Il problema di cui ci siamo occupati è quello della costruzione del numero in rapporto alle operazioni logiche. L'ipotesi dalla quale siamo partiti è naturalmente, che questa costruzione è correlativa allo sviluppo della logica stessa e che al livello prelogico corrisponde un periodo prenumerico. "

"Le operazioni logiche e aritmetiche ci sono dunque apparse come un solo sistema totale e psicologicamente naturale, risultando le seconde dalla generalizzazione e fusione delle prime."

Questa è di fatto la tesi fondamentale del Piaget, in sintesi dunque il bambino inizierebbe una attività di pre-logica, non essendo le strutture logiche strutture innate, e gradualmente ci sarebbe una conquista sempre maggiore di atteggiamenti logici. Inoltre nessuna attività di tipo razionale può prescindere da costruzioni di tipo aritmetico. Per andare nel dettaglio dell'età il Piaget è convinto che un bambino di cinque anni non è, nella norma, capace di comprendere le relazioni logiche più elementari. Ad esempio non dovrebbe essere in grado di sapere che:

$$2 + 3 = 3 + 2, 4 + 5 = 5 + 4$$

e passare alla astratta proprietà commutativa asserente che

"cambiando l'ordine degli addendi la somma non cambia".

Altresì può essere veramente difficoltosa la comprensione delle onnipresenti relazioni di equivalenza e della proprietà transitiva in particolare.

Secondo Beltrand Russel l'idea cardinale del numero si conquista secondo una serie di fasi successive. In primo luogo il bambino impara ad osservare gli oggetti e i gruppi (o insiemi concreti "per ora") di oggetti. Un primo riconoscimento è nel "di più" e "di meno". Così capisce quando un gruppo è di più di un altro. In una fase successiva comprende l'agognato "tanto quanto". Dopo questa fase dice il Russel vi è un intervento costante esterno degli adulti e l'operazione nuova che si compie è la seguente:

il bambino osserva "tre caramelle, tre dita, tre pupazzi" ma il "tre" è suggerito dall'esterno come simbolo nominale comune. Quindi il numero è conquistato prima nominalmente e poi concettualmente. Probabilmente, e Piaget è di questo avviso, è solo vicino ai sette-otto anni che il bambino sviluppa strutture operative che gli consentiranno sia pure con grande gradualità di passare da un gioco-intuitivo ad una comprensione delle relazioni di inclusione prima e di equivalenza poi.

Le relazioni di equivalenza sono relazioni su un insieme che godono delle tre proprietà: riflessiva, simmetrica e transitiva.

La più importante relazione da conquistare è la *relazione di equicardinalità* tra insiemi concreti.

Gli insiemi concreti sono quelli costruiti con gli oggetti del mondo dell'infanzia, palline più o meno colorate, caramelle, piccoli pupazzi ed animali e tante altre cose che si presentano al fanciullo.

Capire la *equicardinalità* significa capire il "tanto quanto", significa capire che :

- se il gruppo delle caramelle è "tanto quanto" quello delle dita,
- se il gruppo delle dita è tanto quanto quello del gruppo delle palline,

ALLORA

- il gruppo delle caramelle è "tanto quanto" quello delle palline.

Infine si capisce, si intuisce in maniera più o meno cosciente che vi è un “QUID” di astratto in comune a gruppi equicardinali e questo quid di astratto conquistato è il NUMERO.

Anche secondo Piaget il processo ultimo di formazione del concetto di numero sarebbe una sorta di *principio di conservazione della quantità* (o equicardinalità che dir si voglia) . Ogni attività razionale non può che partire da questo stadio e l'ipotesi dell'invarianza della quantità, la vera conquista dell'equicardinale, del numero come misura della quantità o numero naturale , pitagorico, intero etc. etc. si compone passando attraverso tre stadi logici:

- *Un primo stadio nel quale il bambino si rende conto più delle variazioni o differenze di quantità. Questo stadio non è molto dissimile dal primo stadio Russelliano.*
- *Un secondo stadio nel quale ci si rende conto o si incomincia a rendersi conto al più solo delle piccolissime equicardinalità (piccolissime nel senso del numero delle dita di una mano).*
- *Un terzo stadio nel quale l'invarianza della quantità è accettata e conquistata come cosa naturale, non più limitata a piccole cose osservabili con le mani direttamente, ma come una conquista del pensiero logico. Dice il Piaget (Szemiska,1941) a riguardo: “ Il bambino non ha più da riflettere per assicurarsi della conservazione delle quantità totali : ne è certo a priori” .*

Così i processi indicati dal Russell come individuazione delle relazioni di equivalenza e di passaggio all'astratto per la conquista del numero in Piaget sono individuati come atti di *classificazione* e *seriazione*. Nello specifico per Piaget seriare significa distinguere ogni classe-elemento-gruppo in quanto non equivalente ad altri mentre classificare significa la fase preliminare del riunire gli oggetti in classi di equivalenza.

Questi processi si ripetono in altri ambienti, ad esempio quando si classificano gli animali. Mettendo insieme animali di dato tipo si ottengono delle classi, si sono classificati. Il concetto astratto ad esempio di “felini” o “equini” o più in generale di “mammiferi” è una seriazione magari con inclusione che naturalmente ne deriva.

Il punto di vista di Piaget ha subito profonde critiche negli ultimi anni. L'insoddisfazione nasce da considerazioni tipicamente pedagogiche: la suddivisione in periodi di apprendimento **non** sarebbe di alcuna utilità per comprendere i processi di apprendimento, anzi andrebbero maggiormente e più tecnicamente differenziati. In altre parole la teoria di Piaget si presenterebbe didatticamente irrilevante (cfr. Engelmann 1983 e Hughes 1986) in quanto essa non fornirebbe ad un docente alcuna indicazione su strategie da seguire. Inoltre per parecchi studiosi sperimentali l'idea del numero sarebbe ben presente in alcuni in età precedente ai 5 anni del Piaget ed assente in età più elevate per altri, così che l'assolutismo del Piaget ne uscirebbe contestato.

Ancora un aspetto sulle difficoltà della comprensione della Matematica è dato dal fatto che, oltre all'impadronirsi dei concetti, il bambino deve anche impadronirsi del simbolismo scritto e parlato dell'aritmetica e successivamente deve imparare a capire che questa disciplina ha un grosso riscontro con molti fatti concreti del mondo.

3. CONCLUSIONI

Entrambe le strade sembrano dare una medesima risposta riguardo al grande problema dell'insegnamento dell'aritmetica e delle grandi difficoltà di questo compito. E' paradossale ma l'aritmetica, pur sembrando così naturale e semplice, costituisce di per se con il suo formalismo e le sue regole un codice complesso. Il bambino che in età pre-scolare ha imparato qualche piccola nozione non le ritrova nel linguaggio scolastico, spesso lontano dal mondo reale.

Questo è spesso a ragione per il fatto che la realizzazione di salti di qualità può avvenire ma solo a prezzi di sforzi immani, che a volte devono compiersi anche contro quelli che dovrebbero essere i naturali alleati e cioè i docenti e i genitori, allora che manifestino indisponibilità e intolleranze, sia pure comprensibili. È vero che occorre portare il bambino non solo ad una abilità computazionale ma addirittura ad una abilità applicativa per la soluzione sia pure di semplici problemi (problem solving). Le difficoltà vengono anche da fattori di ordine linguistico e l'astrattezza della Matematica

richiede a volte una vera e propria opera di traduzione da parte dei bambini stessi che da veri autodidatti devono cercare delle forme a loro più comprensibili, magari scambiando l'esempio con la regola e il particolare con il generale. Una ulteriore foresta di difficoltà si presenta allora che si pensi alla Matematica come disciplina in cui non sono ammessi salti logici. Tutto è strettamente concatenato e coerente. L'incoerenza è errore. Inoltre per arrivare al momento magico del Problem solving occorre imparare una serie di procedure precise, l'uso di algoritmi efficienti e la capacità di prendere le decisioni più plausibili nel caso di alternative.

Così non siamo autorizzati, nel caso di disabilità matematica, ad una politica di non intervento per le età pre-anni sette, alla luce del "tanto mancano le capacità logiche" di piagetiana memoria. Oggi sappiamo che così non è ed occorre dunque una politica di intervento e le strategie per operare in questa direzione.

BIBLIOGRAFIA

- [1] E.AMBRISI-B.RIZZI, *Interdisciplinarietà e nuovo fusionismo*, Archimede 1-2 (1978), 49-51.
- [2] S.ENGELMANN, *Piaget e la didattica: promessa o delusione?*, Psicologia e scuola 16 (1983), 6-12.
- [3] F.EUGENI, *Le due rivoluzioni del secolo: da Bourbaki alla Matematica del discreto*, Per. di Mat. 1 (1992), 3-21.
- [4] M.HUGHES, *What is difficult about learning arithmetic?*, in Donaldson Editor: "Early Childhood Development and Education, Basil Blackwell, Oxford (1983).
- [5] J.PIAGET-A.SZEMISKA, *La genesi del numero del bambino*, La Nuova Italia (1968). (I Ed. 1941).
- [6] G.PETTER, *Conversazioni psicologiche con gli insegnanti*, Ediz. Bemporad Marzocco, Firenze, 1966.

SISTEMI DINAMICI CAOTICI: UN ESEMPIO IN CAMPO NEUROBIOLOGICO

Serena Doria

*Facoltà di Scienze Matematiche, Fisiche e Naturali
Università "G. D'Annunzio" - Chieti*

Un sistema dinamico è costituito da più variabili $h_1(t), h_2(t), \dots, h_N(t)$, funzioni del tempo, che interagiscono tra loro.

Esempi di sistemi dinamici possono essere trovati in Economia, Ecologia, Biologia.

In particolare il sistema nervoso è un esempio di sistema dinamico.

Le variabili di tale sistema sono rappresentate dai neuroni. Per capire come essi interagiscono è necessario schematizzare brevemente la struttura ed il funzionamento delle cellule nervose.

L'attività di un neurone è determinata dalla membrana neuronale, che per la sua diversa permeabilità ad alcuni ioni, determina una diversa concentrazione di cariche elettriche tra l'interno e l'esterno, producendo una differenza di potenziale elettrico nel neurone.

Quando la differenza di potenziale supera un valore soglia, il neurone emette un impulso elettrico, cioè risulta attivo. Altrimenti rimane quiescente.

L'impulso elettrico, emesso dal neurone viene trasmesso alle altre cellule nervose attraverso un prolungamento che prende il nome di assone; inoltre ogni neurone ha altri prolungamenti, i dendriti, che hanno la funzione di ricevere gli eventuali impulsi provenienti dalle altre cellule nervose.

Assoni e dendriti di differenti neuroni sono congiunti tra loro dalle sinapsi, degli ispessimenti che permettono il passaggio degli impulsi elettrici tra i neuroni.

Le sinapsi possono avere un effetto inibitorio o eccitatorio sull'impulso elettrico trasmesso. Inoltre ogni neurone non è collegato con tutte le altre cellule del sistema nervoso.

L'interazione tra le cellule neuronali è determinata da altre numerose variabili, molte delle quali ancora sconosciute. Ai fini dello studio del sistema nervoso e di una sua schematizzazione in

un modello matematico alcune variabili vengono considerate *essenziali*, cioè determinanti rispetto ad altre e prendono il nome di *componenti* del sistema.

Il sistema nervoso viene rappresentato come un sistema dinamico ad N componenti $h_1(t)$, $h_2(t)$, ... $h_N(t)$, ciascuna delle quali rappresenta il potenziale elettrico di un neurone ad un certo istante ed N è il numero totale di neuroni.

Ad ogni istante di tempo t il sistema è caratterizzato da una configurazione determinata dai valori dei potenziali elettrici assunti dai neuroni.

Tale configurazione prende il nome di *stato del sistema* ed è rappresentata da un vettore ad N componenti

$$H(t) = (h_1(t), h_2(t), \dots, h_N(t)).$$

Le configurazioni assunte dal sistema nervoso dipendono dagli impulsi esterni ricevuti.

Ad esempio un odore o la visione di una immagine producono configurazioni differenti perché diversi sono i neuroni stimolati dalle due percezioni.

Al variare del tempo variano le configurazioni assunte dal sistema, poiché i neuroni interagendo tra loro mediante le sinapsi modificano i loro potenziali.

Tutte le successive configurazioni assunte dal sistema possono essere rappresentate mediante punti di uno spazio ad N dimensioni, che prende il nome di *spazio delle fasi* o *spazio delle configurazioni*.

L'insieme delle successive configurazioni assunte dal sistema al variare del tempo forma una curva nello spazio delle fasi, che prende il nome di *traiettoria* del sistema.

Si definisce condizione iniziale del sistema la configurazione che esso assume all'istante in cui inizia l'osservazione

$$H(t_0) = (h_1(t_0), h_2(t_0), \dots, h_N(t_0)).$$

La regola che descrive l'evoluzione del sistema nel tempo e che rappresenta le modalità di interazione tra i neuroni, prende il nome di *dinamica* del sistema.

no fino ad allontanarsi esponenzialmente. Tale fenomeno prende il nome di divergenza esponenziale delle traiettorie.

Nello spazio delle fasi il comportamento asintotico di un sistema caotico è rappresentato da un ente geometrico che prende il nome di *attrattore strano o caotico*.

Generalmente la presenza di caos è interpretata negativamente perché fonte di imprevedibilità, questo accade spesso nei sistemi dinamici applicati all'economia. Tale giudizio non è in genere estendibile ai sistemi dinamici neuronal, dove la presenza di caos può essere vista come fonte di intuizione. In particolare nel modello studiato in [2] si è verificato, mediante simulazioni al calcolatore, come il fenomeno della divergenza esponenziale delle traiettorie non alteri la capacità di riconoscimento della rete ed inoltre l'aumento del parametro di non linearità del sistema sia determinante ai fini del riconoscimento dell'informazione. L'attività caotica che caratterizza il modello teorico è stata riscontrata anche su cortecce cerebrali animali da esperimenti effettuati dal Prof. Freeman, docente di neurobiologia all'Università della California di Berkeley.

La presenza di caos potrebbe essere una delle proprietà che distingue il cervello umano dall'intelligenza artificiale.

Un notevole vantaggio che il caos può conferire al cervello è di renderlo capace di produrre continuamente nuovi tipi di attività, che potrebbero essere decisive per lo sviluppo di raggruppamenti di neuroni diversi da quelli già stabiliti.

BIBLIOGRAFIA

- [1] D. J. AMIT *Modelling brain function: the world of attractor neural network*, N. Y, Cambridge University Press ads., 1989
- [2] TIROZZI-DORIA *Caos in rete neuronale altamente diluita* Atti III Conferenza di Locarno "Stochastic Process - Geometry and Physics"

IL CODICE DI LEON BATTISTA ALBERTI

Franco Eugeni* e Diana Eugeni**

*Facoltà di Ingegneria della Terza Università di Roma

**Facoltà di Ingegneria dell'Università dell'Aquila

INTRODUZIONE

La necessità di trasmettere messaggi segreti è antichissima. Svetonio (140-69 a.C.) nella sua "Vita dei Cesari", presenta un codice usato da Giulio Cesare (100-44 a.C.): ogni lettera è sostituita con quella che la segue di un certo numero di posti

a b c d e f g h i l m n o p q r s t u v z - chiaro
U V Z A B C D E F G H I L M N O P Q R S T - cifrato

Ad esempio:

testo chiaro: romaeunagrandemetropoliurbana.....

testocifrato: OLHUBRI..... ROVUIU...etc.

Sul finire del Medio Evo, l'inizio di relazioni diplomatiche tra stati e staterelli, conduce all'uso dei codici segreti e con esso la ricerca della rottura dei codici in uso. Nasce quindi "la necessità di costruire sempre nuovi codici in grado di rimpiazzare quelli ormai noti agli avversari".

Inizia anche l'idea graduale di sostituire le lettere degli alfabeti latini con quelle di altre culture (greche, ebraiche o inventati) o con parole (Nomenclatori). Si era infatti intuita la debolezza dei codici *monoalfabetici*, cioè degli alfabeti in biezione con un alfabeto classico, e si voleva di tanto in tanto sostituire a singole lettere intere parole, forse con un incoscio desiderio di impedire intuitive statistiche: il tallone di Achille dei codici monoalfabetici. Nacque l'uso dei Nomenclatori. La difficoltà principale era il trasporto del Nomenclatore. I tentativi di superamento di questi sistemi condussero all'idea di trovare metodi facilmente memorizzabili e trasportabili, cioè algoritmi e chiavi.

IL CONTRIBUTO DELL'ALBERTI

I vecchi codici sono più o meno legati al fatto che ad un alfabeto, diciamo in chiaro, se ne sostituisce uno nuovo, il cifrato, in una corrispondenza di simboli "uno ad uno". Il sistema usato si dice monoalfabetico.

Leon Battista Alberti - architetto, urbanista, ed anche matematico e crittografo si accorge per primo di due fatti fondamentali: (cfr. "La cifra" in "Opuscoli morali").

- a) Considerata una qualsiasi lingua, e come esempio la lingua italiana, possiamo esaminare la percentuale con cui una data lettera si presenta in un discorso "abbastanza" lungo. Si scopre che ogni lettera si presenta con una frequenza che è propria della lingua. Tale frequenza delle singole lettere nella lingua, possedendo un testo di messaggio cifrato abbastanza lungo, si ripresenta nel messaggio stesso così come era per la lingua. I simboli più frequenti sono le vocali a, e, i, o ed è altamente probabile che la frequenza propria si conservi. Di fatto è possibile "indovinare" il messaggio codificato con un codice monoalfabetico.
- b) Il segreto del codice deve risiedere solo in un meccanismo facilmente mutabile e questo segreto deve essere facilmente trasportabile.
- c) Occorre trovare una metodologia che permetta di mutare l'alfabeto per codificare, in maniera da rompere le statistiche.

Nel Rinascimento così troviamo un cambiamento radicale, nascono i primi codici segreti che non fanno uso di un solo alfabeto cifrante, ma di molti alfabeti cifranti. Tali codici si chiamano codici polialfabetici.

Nuovi sistemi di cifratura furono ideati oltre che da Leon Battista Alberti, sul cui codice torneremo più avanti, da Giovan Battista Della Porta, celebre multiingegno napoletano (1540-1615), inventore anche della camera oscura, da Gerolamo Cardano (1501-1576), proprio lui: il medico-mago, quello dell'equazione di terzo grado, che, durante la sua sfortunata esistenza, trattò anche di argomenti crittografici di interesse notevole. Fuori d'Italia troviamo il

tedesco Tritemio (Johannes da Trittenheim, 1462-1516), ed infine il francese Blaise de Vigenere (1522-1596) che “fece il grande passo” e raccogliendo le eredità di tutti coloro che l’avevano preceduto realizzò un codice che doveva resistere ad attacchi dei più grandi geni del tempo per più di 300 anni.

L’esame dell’opera di questi Matematici, ma anche personaggi in altri campi, costituisce la migliore analisi dello sviluppo della Scienza dell’Informazione del XV e XVI Secolo.

Leon Battista Alberti nasce a Genova in una data non sicura tra il 1404 e il 1407. Proveniente da famiglia di abili commercianti a livello internazionale si dedica profondamente allo studio. Si addottora a Bologna nel 1428 ed entra al servizio del Cardinal Albergati ai cui servizi era pure Enea Silvio Piccolomini futuro Papa Pio II, lo segue a Roma ove nel 1432 viene assunto come abbreviatore Apostolico da Papa Eugenio IV (Papa dal 1431 al 1447). La situazione dei Papi ed antipapi del tempo è complessa come appare dalla tabella che segue:

PAPI

URBANO VI
1378 - 1389

BONIFACIO IX
1389 - 1404

INNOCENZO VII
(Cosimo Migliorati
da Sulmona)
1404 - 1406

GREGORIO XII
1406 - 1417
(abdica)

ANTIPAPI

CLEMENTE VII
1378 - 1394

BENEDETTO XIII
(Pietro De Luna)
1394 - 1414 deposto

ALTRI ANTIPAPI

ALESSANDRO V
1409 - 1410
(muore avvelenato)

GIOVANNI XXIII
(Cardinale Cassa)
1410 - 1414 deposto

Il Concilio di Costanza, promosso da Sigismondo Malatesta da Rimini inizia nel 1414 e finisce nel 1418, mette fine allo Scisma d'Occidente. Gregorio XII abdica e i rimanenti due Antipapi Benedetto XIII e Giovanni XXIII sono deposti dal Concilio. Il Concilio elegge il Cardinale Colonna che prende il nome di Martino V, nel 1417.

MARTINO V	(1417-1431)	
EUGENIO IV	(1431-1447)	
NICCOLO V	(1447-1455)	
CALLISTO III	(1455-1458)	(Alfonso Borgia)
PIO II	(1458-1464)	(Enea Silvio Piccolomini)
PAOLO II	(1464-1471)	

In questa ripresa dopo lo Scisma Leon Battista Alberti entra alla Corte Papale e rimane al servizio di vari Papi quali Nicolo V Di Sarzana, Callisto III (Alfonso Borgia), Pio II (E. S. Piccolomini) ed infine Paolo II che lo licenzia dal lavoro presso il Vaticano. Il codice (crittografico) di Alberti fu elaborato, sotto Papa Piccolomini, su commissione del segretario pontificio Leonardo Dato, intorno al 1466.

Dopo il licenziamento come abbreviatore l'Alberti si dedica ai suoi lavori di Architettura. Muore a Roma nel 1472. Nel "Trattato di pittura" formula i concetti fondamentali della geometria descrittiva e può essere considerato il padre della Prospettiva. Tra le sue opere architettoniche ricordiamo il Tempio Malatestiano a Rimini, Palazzo Rucellai e la facciata di S.Maria Novella a Firenze e infine S.Andrea a Mantova.

ANALISI E CRITTOANALISI DELLA MACCHINA DI ALBERTI

Il codice crittografico di Alberti fu elaborato, sotto Papa Piccolomini e su richiesta di Leonardo Dato, segretario pontificio, intorno al 1466.

Esso era una macchina di legno mediante costituita da due cerchi concentrici. Nel disco più esterno appare l'alfabeto in chiaro formato da 24 caselle 20 delle quali contenenti lettere (mancano

per ragioni di sicurezza crittografica, le lettere che si presentano con minore frequenza, cioè h, j, k, q, w, x, y) e le rimanenti quattro contenenti i numeri 1, 2, 3, 4. Il disco interno contiene una permutazione di altre 23 lettere (mancano le lettere J e W ed U=V) formanti l'alfabeto cifrante. Esso può ruotare rispetto al primo disco.

Possiamo simularlo utilizzando, per ragioni tipografiche, un modello equivalente costituito da due strisce delle quali la prima fissa e la seconda mobile:

a b c d e f g i l m n o p r s t u v z 1 2 3 4 - chiaro
A B C D E F G H I K L M N O P Q R S T U X Y Z - cifrato

Il segreto del codice consiste nel fissare una permutazione del secondo alfabeto concordata tra trasmettitore e ricevitore e quindi una coppia di lettere che "azzerano" la macchina.

B A C H I K L D E F G M N O T U P Q R S X Z Y - permutazione
(a, F) coppia di azzeramento

La macchina simulata dalla doppia striscia è pronta a lavorare secondo lo schema che segue:

a b c d e f g i l m n o p r s t u v z 1 2 3 4 - chiaro
F G M N O T U P Q R S X Z Y B A C H I K L D E - cifrato

Si prenda il messaggio : "romaeunagrandemetropoliurbana....."
e lo si copi inserendo di tanto in tanto dei numeri di cambio di alfabeto nel modo casuale che segue:

rom1ae3unagr2andeme1tropoliur4bana.....

Seguendo il rapporto chiaro - cifrato iniziamo a codificare:

a b c d e f g i l m n o p r s t u v z 1 2 3 4 - chiaro
F G M N O T U P Q R S X Z Y B A C H I K L D E - cifrato

rom1aeunagr2andeme1tropoliur4bana..... testo chiaro
YXRK testo cifrato

Arrivati ad 1 si codifica 1 come se fosse una lettera però si sposta la lettera F di azzeramento sotto il numero 1 in modo da cambiare alfabeto:

a b c d e f g i l m n o p r s t u v z 1 2 3 4 - chiaro
O T U P Q R S X Z Y B A C H I K L D E E G M N - cifrato

e si continua a crittografare con il nuovo alfabeto:

rom1aeunagr2andeme1tropoliur4bana.....	testo chiaro
YXRKOQLBOSH <u>G</u>	testo cifrato

Al nuovo cambio di alfabeto corrisponde un secondo spostamento che consiste nel portare o (che corrisponde all'azzeramento a) verso destra fino al numero 2.

a b c d e f g i l m n o p r s t u v z 1 2 3 4	- chiaro
P Q R S X Z Y B A C H I K L D E F G M N <u>Q</u> T U	- cifrato

e si continua a crittografare con il nuovo alfabeto:

rom1aeunagr2andeme1tropoliur4bana.....	- chiaro
YXRKOQLBOSHGHPSX <u>CXN</u>	- testo cifrato

Si deve ancora cambiare alfabeto portando P sotto 1 verso destra e continuando a codificare, ma il procedimento è chiaro. La decodifica avviene a rovescio disponendo della macchina azzerata.

a b c d e f g i l m n o p r s t u v z 1 2 3 4	- chiaro
F G M N O T U P Q R S X Z Y B A C H I K L D E	- cifrato

e del messaggio da decodificare

YXRKOQLBOSHGHPSXCN	testo cifrato
roml	decodifica fino al primo numero!

Si porta allora F sotto 1 e si continua:

a b c d e f g i l m n o p r s t u v z 1 2 3 4	- chiaro
O T U P Q R S X Z Y B A C H I K L D E <u>E</u> G M N	- cifrato

YXRKOQLBOSHGHPSXCN	testo cifrato
romlaeunagr2	seguito della decodifica!

È ormai ben chiaro il meccanismo della macchina di Alberti e il nostro discorso volge al termine.

Quando è robusto il codice di Alberti è facile da dire! Se ci mettiamo nell'ipotesi che il "nemico" ovvero l'utente non autorizzato conosca la permutazione dell'alfabeto cifrante il codice appare di una debolezza estrema perché occorrono soltanto i 23 tentativi con le coppie azzeranti:

(a, A), (a, B),, (a, F),, (a, Z)

per scoprire che la coppia (a,F) ha un significato!

Se è segreta anche la permutazione iniziale i tentativi da fare sono:

$$24! \cdot 24$$

che è un grande numero. Tuttavia nasce la difficoltà del comunicarsi la permutazione, con accordi di vario tipo. Per una sovracifatura il codice di Alberti è praticamente un ottimo rinforzo.

Eravamo nella metà del 1400 ed un genio ha prodotto ed ideato tre grandi idee nuove:

la statistica

(scoprendo le frequenze di lettere in una lingua)

i codici polialfabetici

(che erano l'idea vincente ancora oggi valida);

è il caso di concludere:

il GENIO ha prodotto!

BIBLIOGRAFIA

- [1] L.B.ALBERTI, *Opuscoli morali* (VI - La cifra), pubblicato postumo da Cosimo Bartoli, Venezia, 1568. (Archivi Vaticani).
- [2] E.AMBRISI-F.EUGENI, *Il problema della protezione della informazione, I: cenni storici e metodi statistici per la decrittazione*, Ratio Math. 1 (1990), 15-37.
- [3] L.BERARDI, *Algebra e Teoria dei Codici correttori*, Edz.F.Angeli, 1994.
- [4] A.BEUTELSPACHER-F.EUGENI, *Geometrie Finite e crittosistemi: stato dell'arte e problematiche*, Atti del II Simposio Nazionale su : "Stato e prospettive della ricerca crittografica in Italia" a cura della Fondazione Bordini, Roma, 1989.
- [5] C.B.BOYER, *Storia della Matematica*, Mondadori Editore, 1987.
- [6] G.CARDANO, *De subtilitate*, Basilea, 1547.(Vaticano-Apostolica)
- [7] B.DE VIGENERE, *Traicte des chiffres ou secretes manieres d'escrire*, Parigi, 1586, (Roma - Biblioteca Nazionale).

- [8] F.EUGENI-B.K.DASS, *How to share secrets: the idea of threshold games*, Journal of Information & Optimization Sciences, 12 (1991),451-458.
- [9] R.L.MEYER, *Teoria della Comunicazione e struttura urbana*, Il Saggiatore, Edz. Mondadori,Milano, 1969.
- [10] G.B. PORTA, *De Furtivis Literarum notis*, Napoli, 1563, (Roma - Biblioteca Nazionale)
- [11] G.TRITEMIO, *Steganografia*, (traduzione dall'originale latino di F.Benedetti e A.Dupré), Nardini Editore, Firenze, 1982