

Esempi di algoritmi

1. Gli array

Un **array** o **vettore** è una struttura dati complessa usata in molti linguaggi di programmazione e chiaramente ispirata alla nozione matematica di vettore. Più precisamente, l'array è un insieme di variabili che occupano locazioni consecutive in memoria (vedi figura 1).

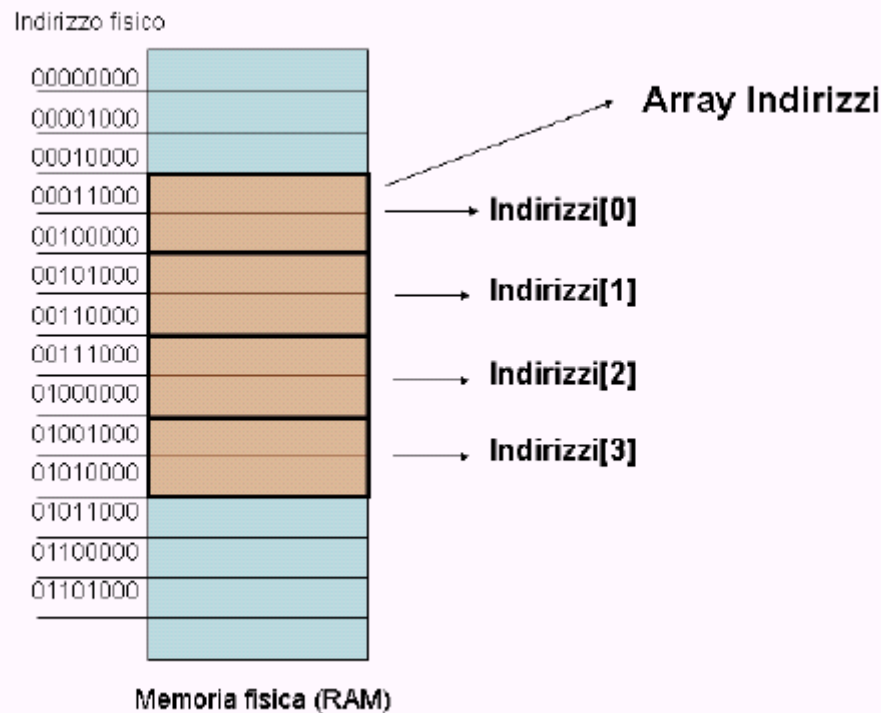


Figura 1

Molte applicazioni possono memorizzare e gestire un numero elevato di informazioni che costituiscono delle banche dati. In particolare tali strutture possono contenere indirizzi dei clienti, codici di conti correnti bancari, numeri di telefono, ...

Normalmente informazioni tipo quelle descritte sopra vengono memorizzate negli array. Un array viene memorizzato in memoria occupandone una zona contigua. La contiguità della struttura offre un duplice vantaggio: in primo luogo le operazioni di scrittura e lettura nell'array risultano più efficienti (minor tempo per accedere alle informazioni) e, in secondo luogo, l'array permette di accedere ai singoli elementi tramite indicizzazione facilitando lo scorrimento di tutta la struttura.

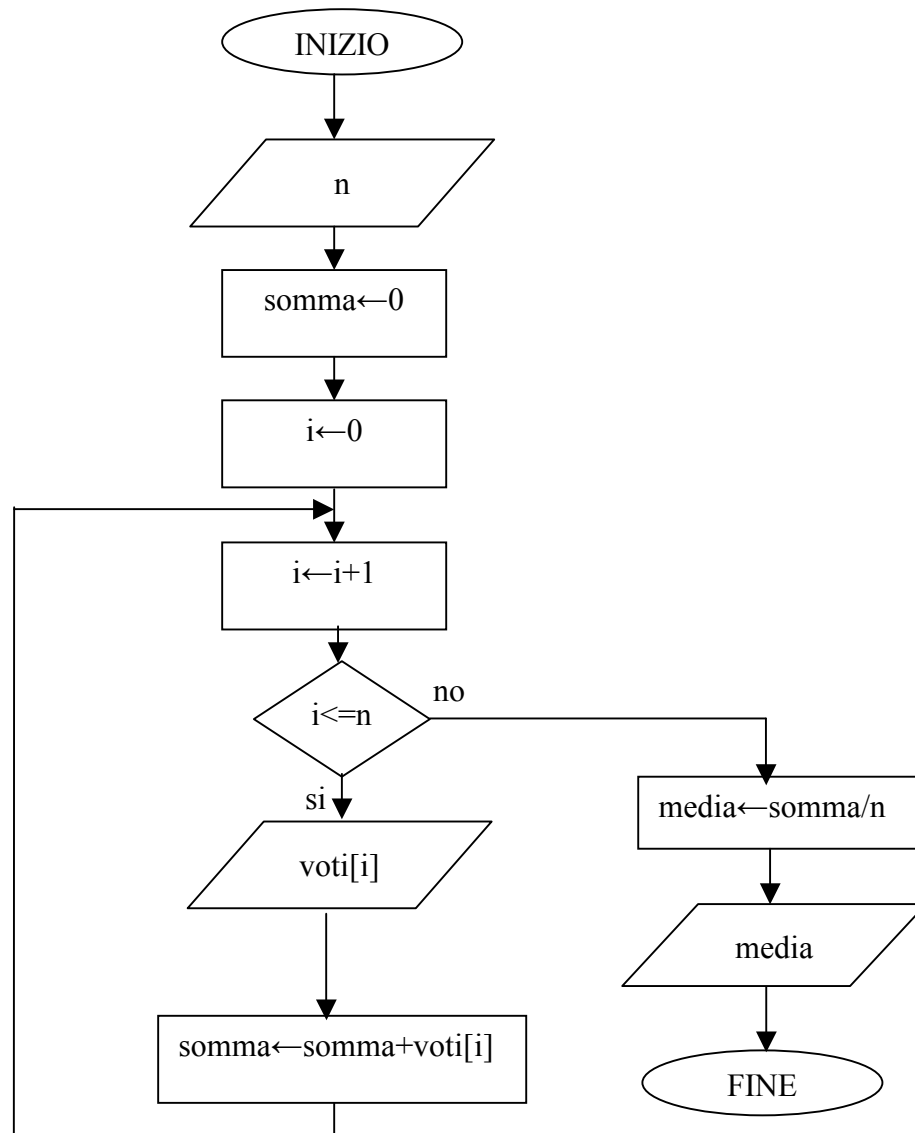
Si può immaginare un array come una sorta di casellario, le cui caselle sono dette **celle** dell'array stesso. Ciascuna delle celle si comporta come una variabile tradizionale; tutte le celle sono variabili di uno stesso tipo preesistente, detto **tipo base** dell'array. Si parlerà perciò di tipi come "array di interi", "array di stringhe", "array di caratteri" e così via.

Ciascuna delle celle dell'array è identificata da un valore di **indice**. L'indice è generalmente numerico e i valori che gli indici possono assumere sono numeri interi *contigui* che partono da 0 o da 1. Nel caso del linguaggio MATLAB, gli indici partono da 1. Si potrà quindi parlare della cella di indice 0, di indice 1, e, in generale,

di indice N, dove N è un intero compreso fra 0 (o 1) e il valore massimo per gli indici dell'array.

2. Diagramma di flusso della media aritmetica

Supponiamo di voler effettuare la media degli n esami superati da uno studente. Gli input del nostro problema sono due: il numero n di esami sostenuti ed i relativi punteggi espressi in trentesimi.



L'algoritmo inizia con la lettura da input del numero n di esami. Successivamente si ha l'inizializzazione a 0 delle due variabili somma ed i. Si entra quindi in un ciclo nel quale la variabile i viene incrementata di 1, di conseguenza alla prima iterazione i assume valore pari ad 1. Quest'ultimo viene confrontato con il valore di n: se 1 risulta essere minore o uguale ad n, viene letto il primo elemento dell'array voti[1] (è da notare che in questo caso l'array parte con l'indice 1). Quindi nella variabile somma viene memorizzato il risultato dell'espressione somma+voti[1], con il valore di somma pari a 0. Successivamente inizia una nuova iterazione con l'incremento della

variabile i . Per chiarire meglio le idee, riportiamo i valori assunti dalla variabile somma con valori di i pari a 1, 2 e 3.

$i=1$

somma=0+voti[1];

$i=2$

somma=voti[1]+voti[2];

$i=3$

somma=voti[1]+voti[2]+voti[3];

Il ciclo continua fino a quando i diventa più grande di n . Alla fine del ciclo, viene calcolata la media e quindi stampata.

3. Valore atteso di una variabile random

In teoria delle probabilità, una **variabile random** può essere pensata come il risultato numerico di un esperimento quando questo non è prevedibile con certezza (ossia non è *deterministico*). Ad esempio, il risultato del lancio di un dado a sei facce può essere matematicamente modellizzato come una variabile casuale che può assumere uno dei sei possibili valori 1, 2, 3, 4, 5 e 6.

La distribuzione di probabilità di una variabile random discreta è la lista di tutti i possibili valori della variabile e la probabilità che ogni valore occorra. La somma di queste probabilità è uno.

Per esempio, sia M il numero di volte che un computer vada in crash mentre si sta scrivendo un articolo. La distribuzione di probabilità della variabile M è la lista di probabilità di ogni possibile risultato: la probabilità che M sia uguale a zero (denotata con $\Pr(M=0)$) è la probabilità che il computer non vada in crash; $\Pr(M=1)$ è la probabilità che il computer vada in crash una volta, $\Pr(M=2)$ due volte, ... e così via.

In tabella 1 è riportato un esempio di distribuzione di probabilità per M .

	Risultati (numero di volte che il computer è andato in crash)				
	0	1	2	3	4
Distribuzione di probabilità	0.80	0.10	0.06	0.03	0.01

Tabella 1

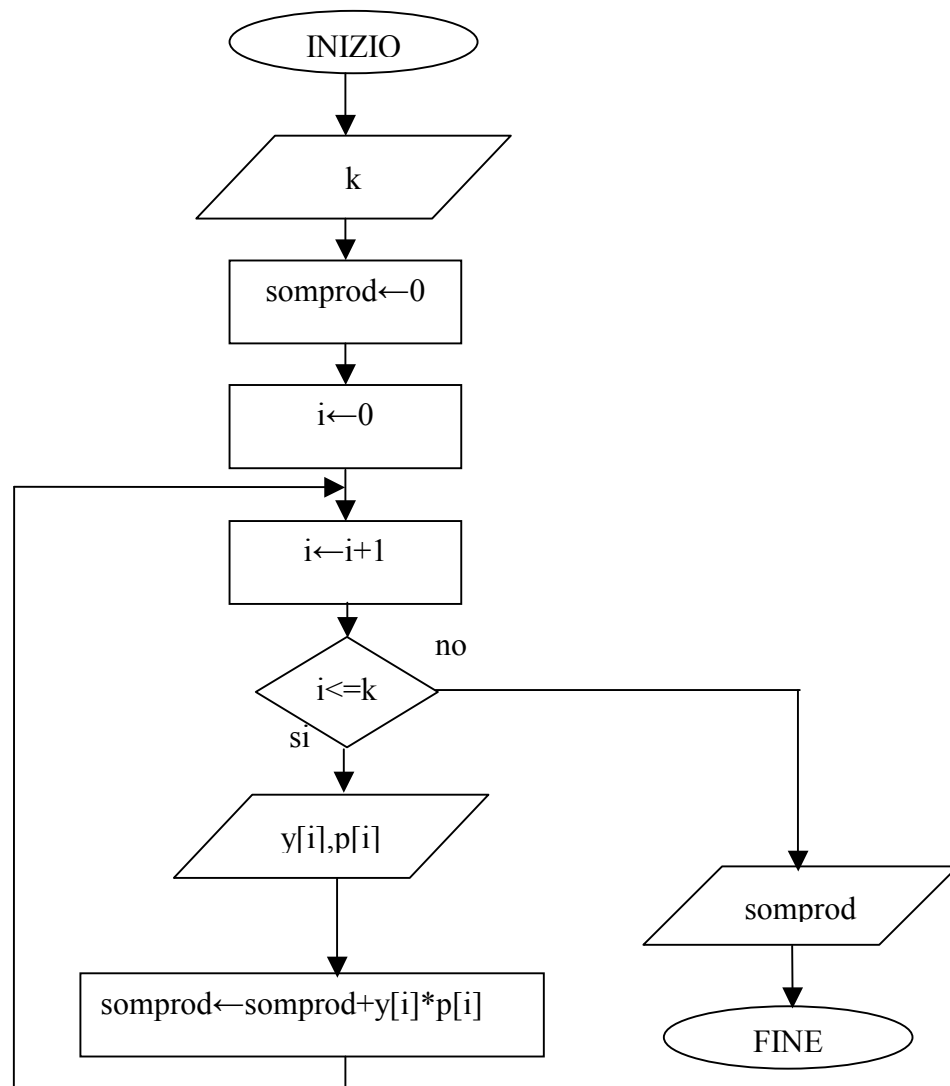
In accordo alla distribuzione di tabella 1, la probabilità di nessun crash è dell'80%, la probabilità di un crash è del 10%; le probabilità di due, tre o quattro crash sono rispettivamente: 6%, 3% e 1%. La somma di tutte le probabilità fornisce il valore 100%.

Supponiamo che la variabile random Y possa assumere k possibili valori $y_1, y_2, \dots, y_k$, dove y_1 denota il primo valore, y_2 il secondo valore, ... e che la probabilità che Y valga y_1 è p_1 , la probabilità che Y valga y_2 è p_2 e così via.

Il valore atteso di Y , denotato con $E(Y)$, è

$$E(Y) = y_1 p_1 + y_2 p_2 + \dots + y_k p_k = \sum_{i=1}^k y_i p_i .$$

Creiamo ora il diagramma di flusso che consente di calcolare e stampare il valore atteso di Y.



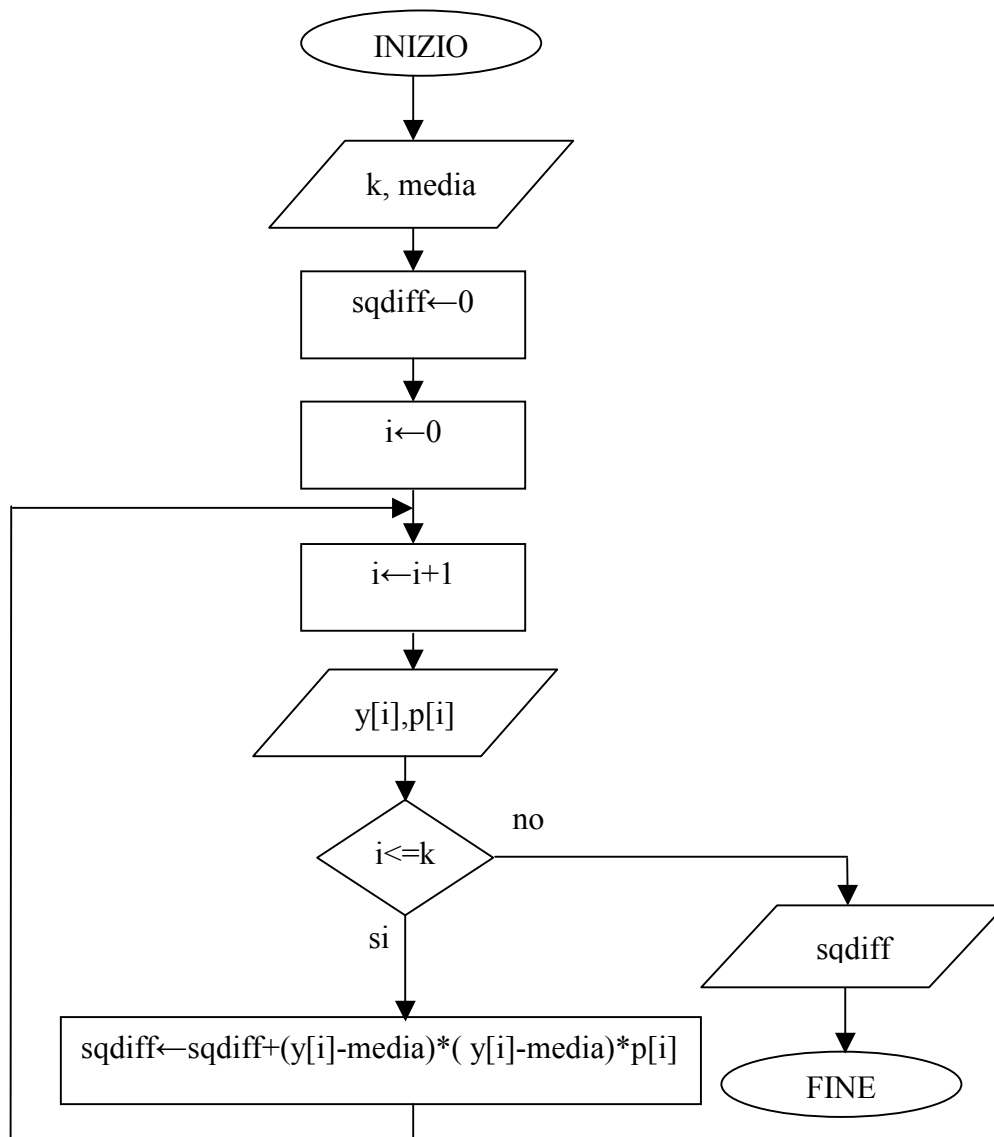
Il diagramma di flusso è analogo al precedente, con la differenza di avere due vettori in input (uno contenente i valori della variabile random Y e un altro contenente i relativi valori di probabilità) ed una somma di prodotti.

Un utile esercizio può essere quello di riprodurre i passi del ciclo per valori di i da 1 a 5.

4. Varianza

La varianza misura la dispersione di una distribuzione di probabilità. La varianza di una variabile random, denotata con $\text{var}(Y)$, è il valore atteso del quadrato della deviazione di Y dalla sua media μ_y , cioè

$$\text{var}(Y) = E[(Y - \mu_y)^2] = \sum_{i=1}^k (y_i - \mu_y)^2 p_i$$

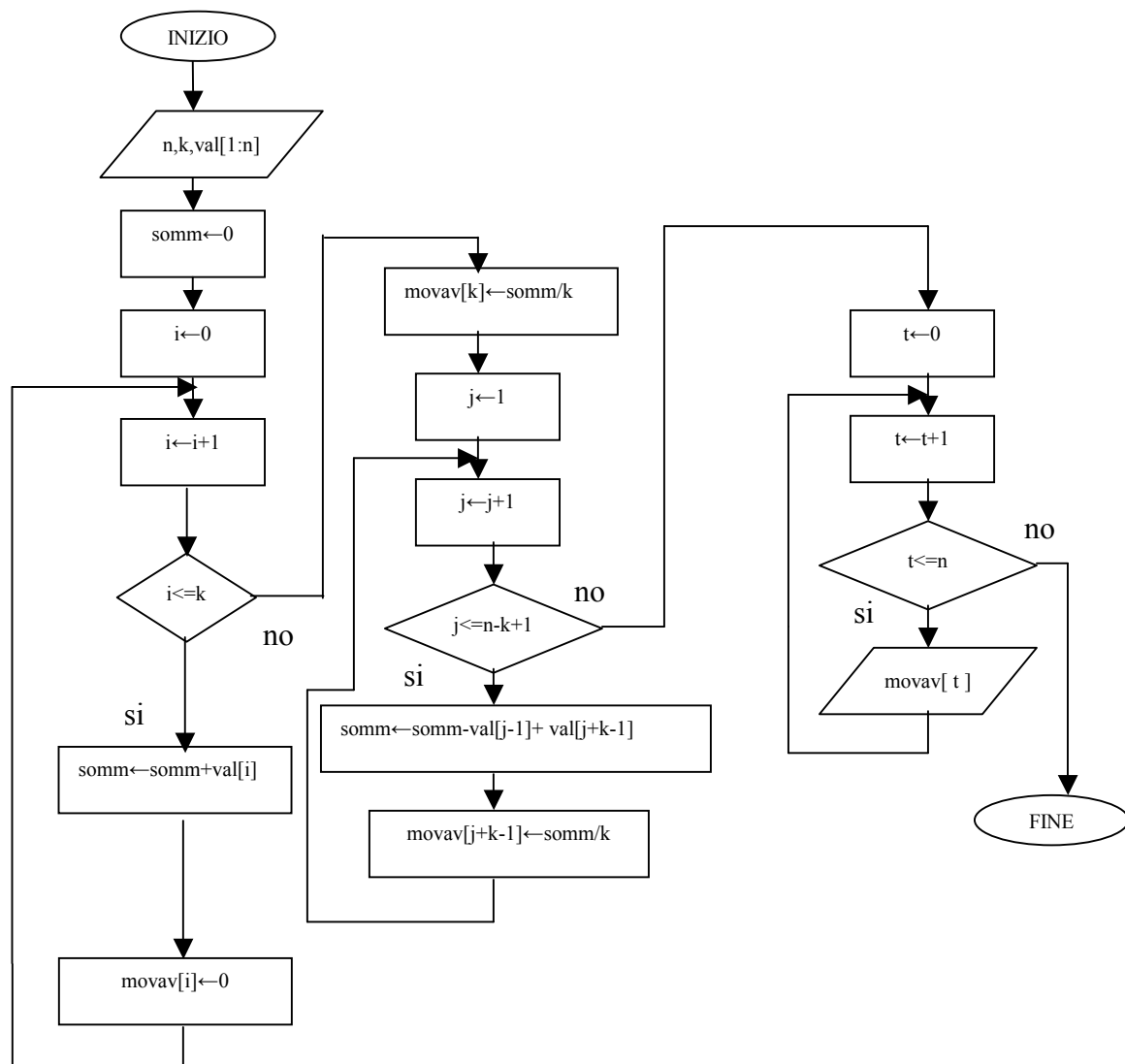


La struttura del diagramma di flusso è simile a quella degli esempi precedenti.

5. Media mobile

La *media mobile* è, in assoluto, l'indicatore tecnico più utilizzato; il suo scopo non è quello di anticipare prezzi o di prevederne le evoluzioni future, ma di seguirli passo passo nei loro movimenti.

Come suggerisce il termine *media*, si tratta di una media di una determinata quantità di dati (solitamente vengono considerati quelli di chiusura, ma ovviamente possono essere utilizzati anche quelli di massimo, minimo, piuttosto che apertura); il termine *mobile* invece si riferisce al fatto che vengono considerati soltanto gli ultimi prezzi desiderati: in una media mobile a k giorni, calcolato sui prezzi di chiusura, verranno presi in considerazione gli ultimi k close di seduta. Mediando i prezzi quindi si può ottenere una linea smussata che rende più semplice la visione del trend.



Consideriamo n valori $(x_1, x_2, \dots, x_n)$ e supponiamo di voler costruire il diagramma di flusso che calcola la media mobile degli x_i a k valori. Supponendo di voler memorizzare i risultati delle medie mobili in un vettore ad n dimensioni, si avrà:

$movav[1]=0$

$movav[2]=0$

...

```

movav[k-1]=0
movav[k]=(x1 + x2 +...+ xk)/k
movav[k+1]=(x2 + x3 +...+ xk+1)/k
movav[k+2]=(x3 + x4 +...+ xk+2)/k
...
movav[n]=(xn-k+1+xn-k+2 + ...+xn)/k

```

La costruzione del diagramma di flusso è ottimizzata: calcolo la somma per la nuova media mobile a partire dalla somma precedente, considerando che il nuovo intervallo ha un elemento in meno (a sinistra) ed uno in più (a destra) del precedente.