

Qualche esempio di applicazioni

Domizia Orestano

orestano@fis.uniroma3.it

webusers.fis.uniroma3.it/~orestano/applicazioni.pdf

Molti esempi nel testo

Barone, Marinari, Organtini, Ricci-Tersenghi - Programmazione Scientifica - Pearson Education

- Ricerca degli 0 di una funzione mediante il metodo di bisezione e mediante il metodo di Newton (4.3.2)
- Ricerca dei numeri primi (4.3.3)
- Problemi di arrotondamento (4.5)
- Riordinamento degli elementi di un vettore: Bubblesort (5.2.1)
- Ricerca binaria (5.2.2)
- Soluzione di sistemi di equazioni lineari (5.4, 7.3.2)
- Generazione di numeri casuali (5.5)
- Manipolazione di testi (5.6.3, 6.5.1)
- Istogrammi (7.4.1) – Visto
- Calcolo del χ^2 di una distribuzione (7.4.2)
- Calcolo di una derivata (7.6)
- Interpolazione e integrazione numerica (cap. 8)

soluzione di equazioni $f(x)=0$

- Uno dei problemi matematici di base è quello di trovare le soluzioni (o radici) di un'equazione $f(x)=0$, o che è lo stesso gli zeri della funzione $y = f(x)$;
- esistono per questo formule risolutive "esatte" e metodi iterativi "approssimati" o numerici.
- solo alcuni tipi molto particolari di equazioni ammettono formule risolutive (p.es. equazioni algebriche fino al quarto grado) e anche quando esistono queste formule sono spesso molto pesanti dal punto di vista computazionale.
- vediamo qui due metodi numerici: il metodo di bisezione e il metodo di Newton

metodo di bisezione

- e' un metodo **iterativo** ed e' il metodo piu' semplice
- ricerca del cambiamento di segno:
 - si calcola la funzione in un dato intervallo $[a, b]$ cominciando da a e con un passo h si percorre tutto l'intervallo calcolando per ogni x il corrispondente $y = f(x)$;
 - quando si verifica un cambiamento di segno della y , si può concludere che c'è almeno uno zero in $[x - h, x]$, a condizione che la funzione sia continua nell'intervallo
- **iterazioni**: da ripetere fin tanto che la lunghezza dell'intervallo $[a_i, b_i]$ e' inferiore a una precisione ϵ data
 - partendo dall'intervallo $[a_0, b_0]$ contenente uno zero
 - si determina $c_0 = (a_0 + b_0)/2$
 - si valuta il prodotto $p_0 = f(a_0) * f(c_0)$
 - se $p_0 < 0$ la soluzione e' compresa in $[a_0, c_0]$, si pone $a_1 = a_0$ $b_1 = c_0$
 - se $p_0 > 0$ la soluzione e' compresa in $[c_0, b_0]$, si pone $a_1 = c_0$ $b_1 = b_0$
 - si passa all'iterazione successiva partendo dall'intervallo $[a_1, b_1]$

metodo di Newton

- converge piu' velocemente del metodo di bisezione ma richiede una prima stima abbastanza buona della radice
- anche questo e' un metodo **iterativo**
 - data x_n stima ottenuta all'iterazione n
 - la stima successiva si costruisce ponendo uguale a zero $f(x_{n+1})$, ottenuta espandendo in serie la $f(x)$ intorno a x_n e fermandosi al primo ordine:
 - $f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0$
 - da cui $x_{n+1} = x_n - f(x_n)/f'(x_n)$
 - se la derivata f' non e' nota **la si stima** come $f'(x_n) = (f(x_n) - f(x_{n-1})) / (x_n - x_{n-1})$
 - ci si ferma quando la distanza tra due stime successive e' inferiore alla precisione ε prescelta

Interpolazione ed estrapolazione

- Il problema consiste nel calcolare il valore assunto da una funzione $f(x)$ in un punto x arbitrario, noti i valori della funzione in un insieme finito di punti x_i con $i=1, \dots, N$.
- Se x è compresa tra due punti noti si parla di **interpolazione**, se è all'esterno dell'intervallo coperto dai punti noti si parla di **estrapolazione**.
- Illustriamo qui uno dei metodi più semplici.

Interpolazione mediante polinomi

Dato un qualsiasi insieme di N punti esiste un unico polinomio di ordine $N-1$ passante per questi punti. È definito dalla formula di Lagrange:

$$P_{N-1}(x) = \sum_i \left(\prod_j (x-x_j)/(x_i-x_j) \right) f(x_i)$$

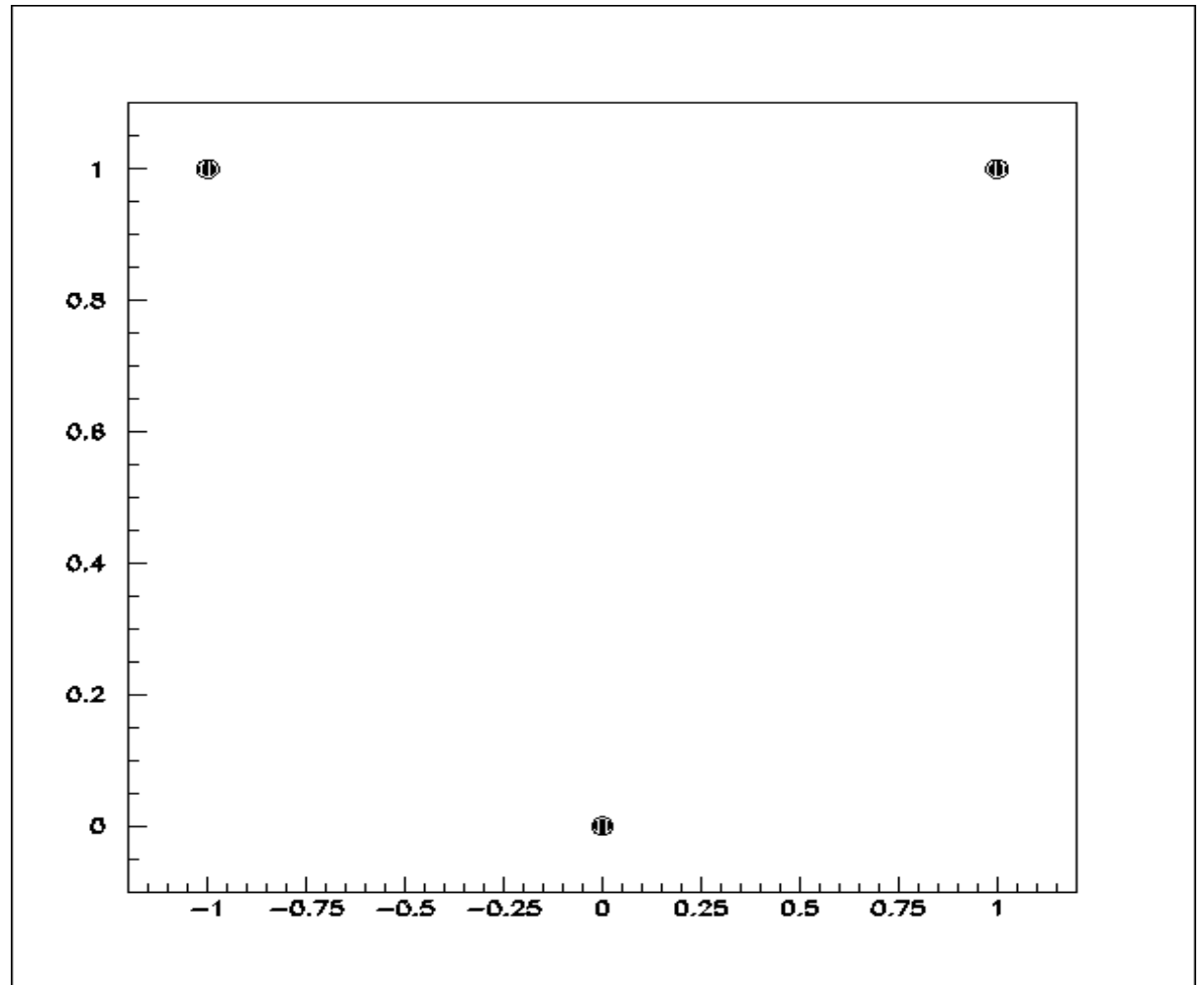
dove la sommatoria è sull'indice i e va da 1 a N , la produttoria è sull'indice j e va da 1 a N con j diverso da i

Errore

- L'errore commesso sull'interpolazione della funzione f nel punto x è
 $r(x) = f(x) - P_n(x)$ con n grado del polinomio interpolatore
- Se la funzione f è derivabile almeno n volte abbiamo un teorema che afferma che esiste almeno un punto ξ appartenente all'intervallo $x_1 \dots x_N$ nel quale
 $r(x) = \frac{f^{(n)}(\xi)}{(n!) \prod_{i=1}^n (x - x_i)}$, dove la produttoria è su i compreso tra 1 e N , e g è la derivata n -ma ($n=N-1$) della funzione f

Esempio

Supponiamo di cercare una funzione passante per i punti $(0,0)$, $(-1,1)$ e $(1,1)$.

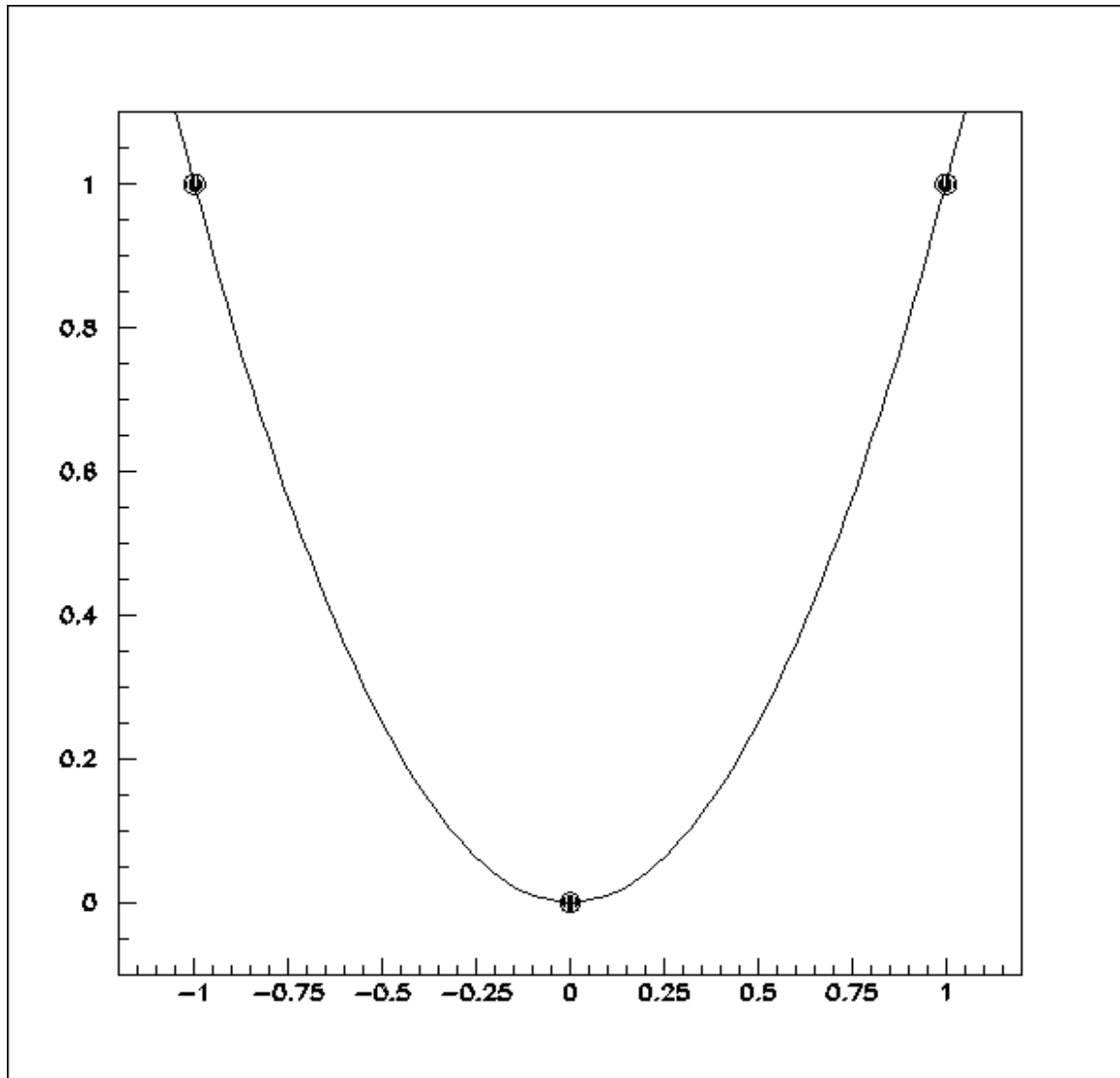


Esempio

Il polinomio di Lagrange avrà grado 2 e sarà dato da

$$\begin{aligned} P_2(x) &= (x+1)/(0+1)*(x-1)/(0-1)*0 \quad // \quad i=1 \\ &\quad + (x-0)/(-1-0)*(x-1)/(-1-1)*1 \quad // \quad i=2 \\ &\quad + (x-0)/(1-0)*(x+1)/(1+1)*1 \quad // \quad i=3 \\ &= x*(x-1)/2 + x*(x+1)/2 = x^2 \end{aligned}$$

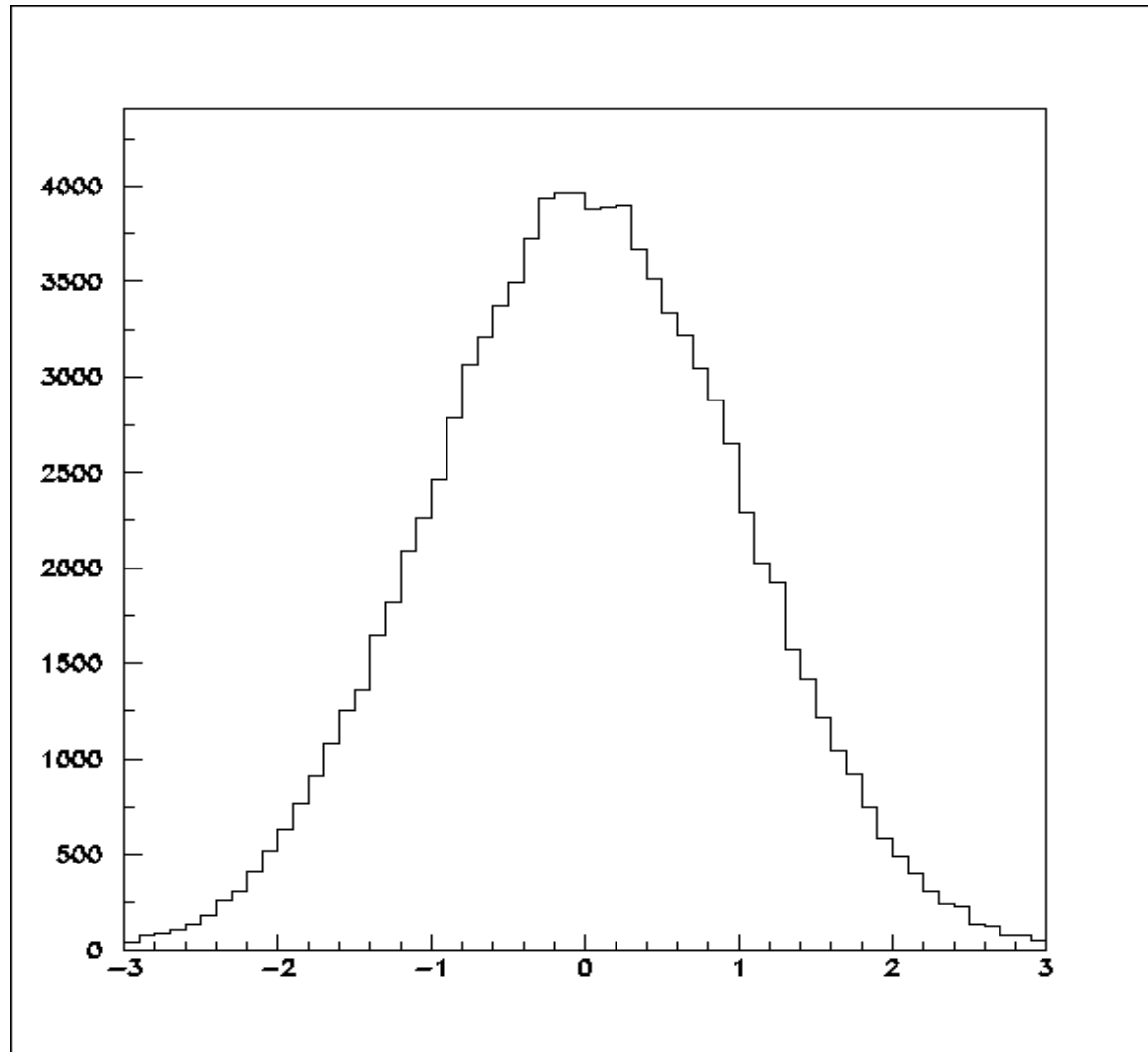
Esempio

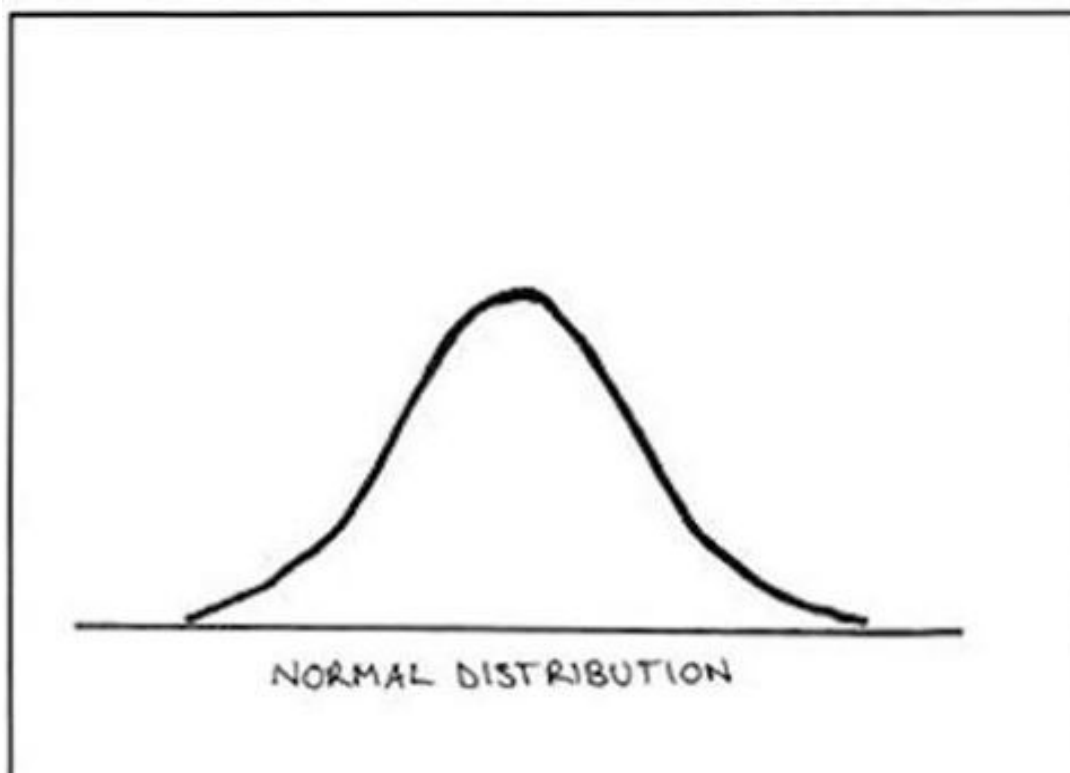


Regressione (fit)

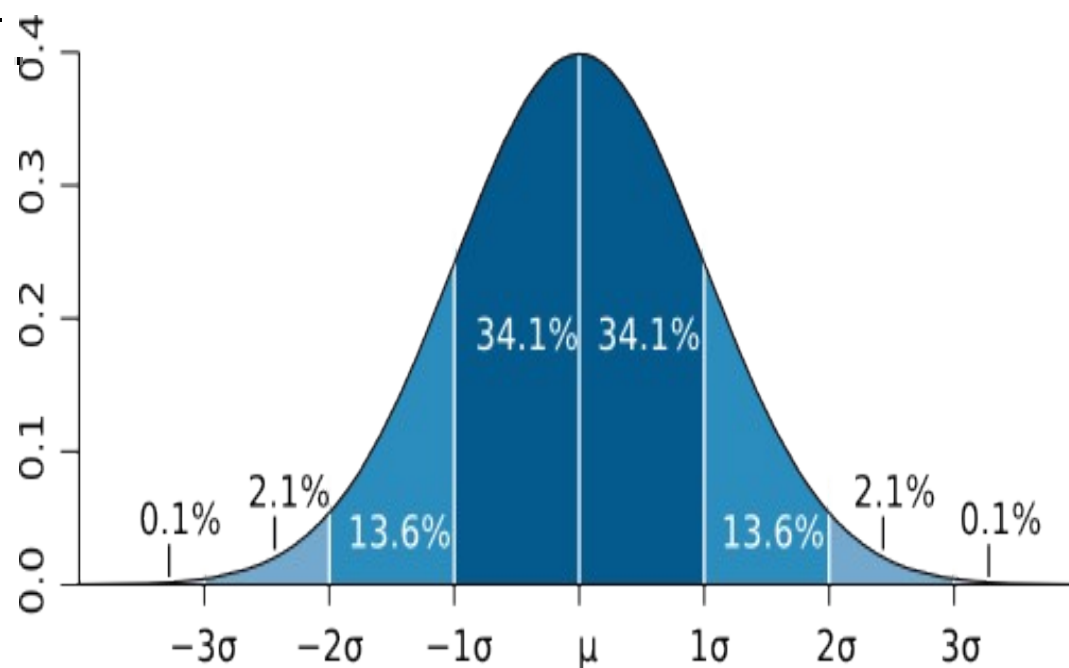
- se invece abbiamo un insieme di dati sperimentali che seguono un andamento funzionale previsto da un modello teorico dipendente da un insieme di parametri possiamo usarli per
 - estrarre i parametri del modello
 - stimare l'errore su di essi
 - valutare la bonta' della descrizione fornita dal modello

dati prodotti da un nostro generatore
di numeri pseudo-normali
(100000 numeri)





$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$





NORMAL DISTRIBUTION

da non confondere con:



PARANORMAL DISTRIBUTION

Freeman.

Fit gaussiano

ipotesi: dati gaussiani

parametri: termine
costante, valor medio e
sigma

risultati:

media=0.0001

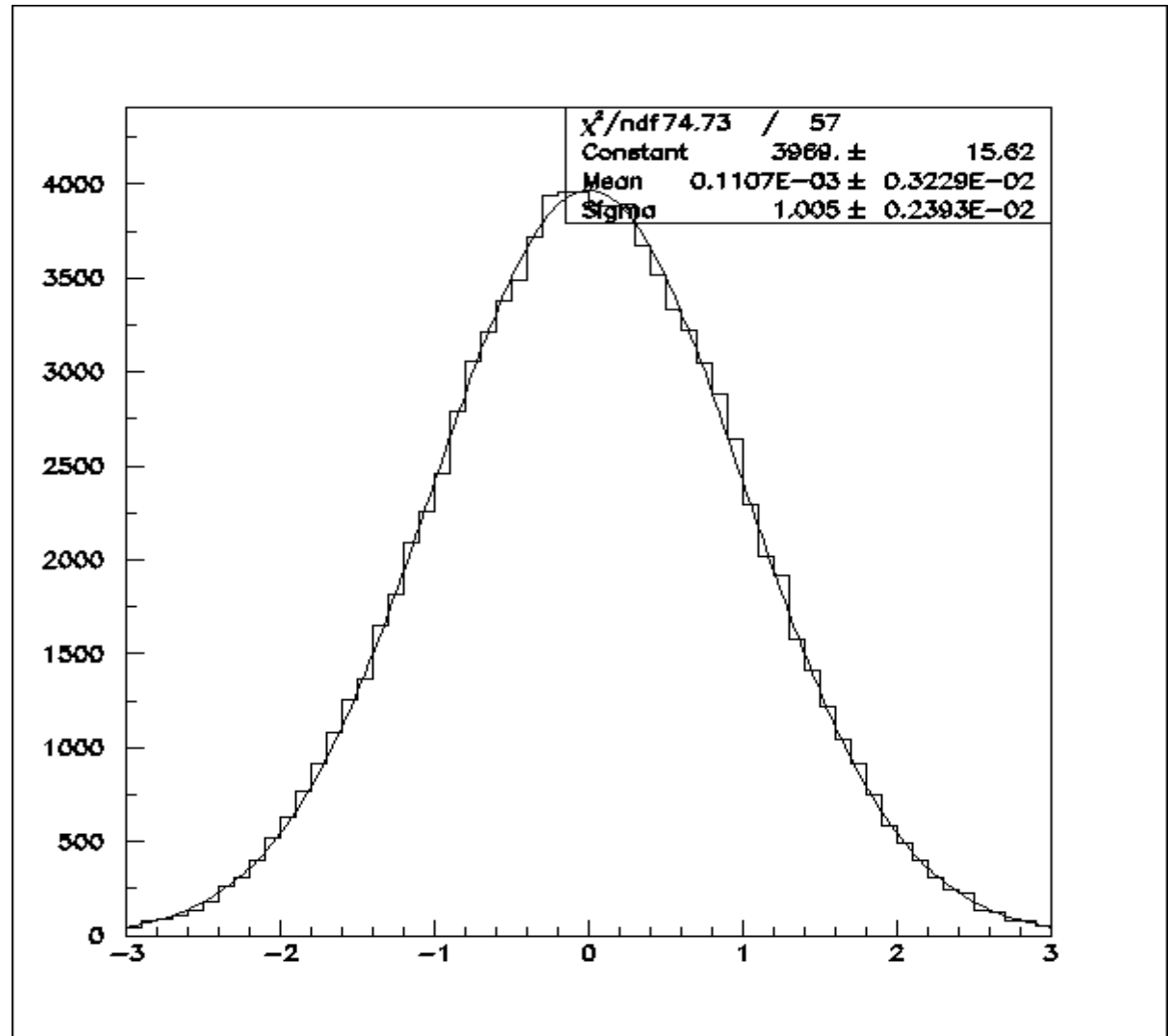
errore sulla media=0.003

sigma=1.005

errore su sigma=0.002

χ^2 per grado di liberta' =
74.7/57

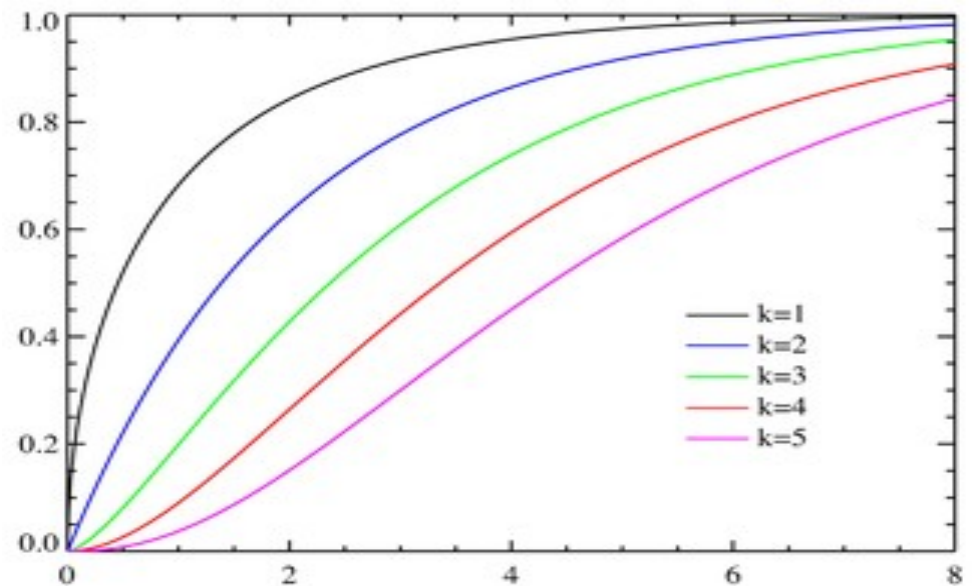
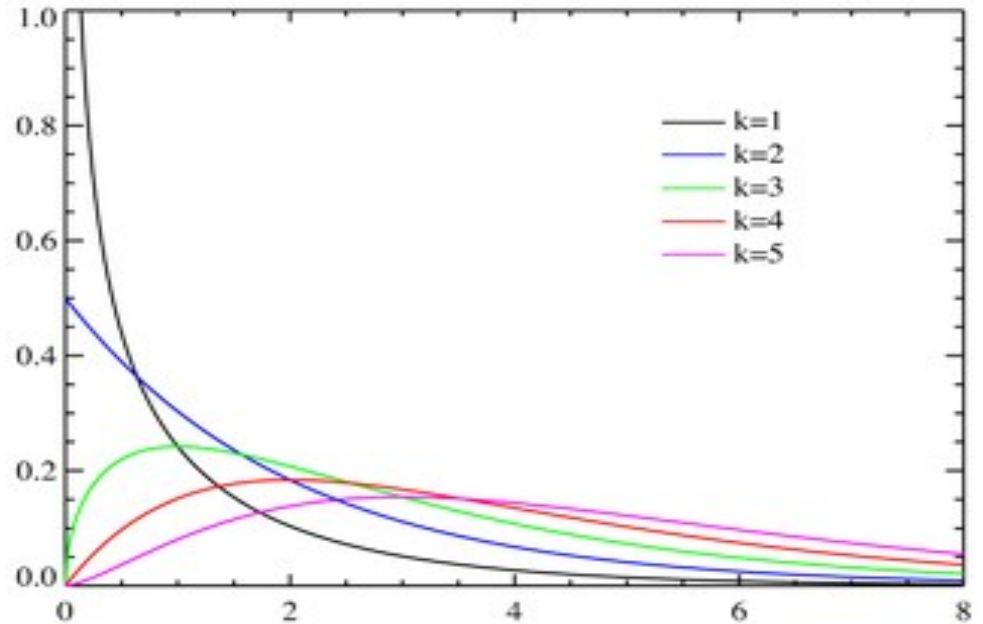
misura la bonta' del fit



χ^2

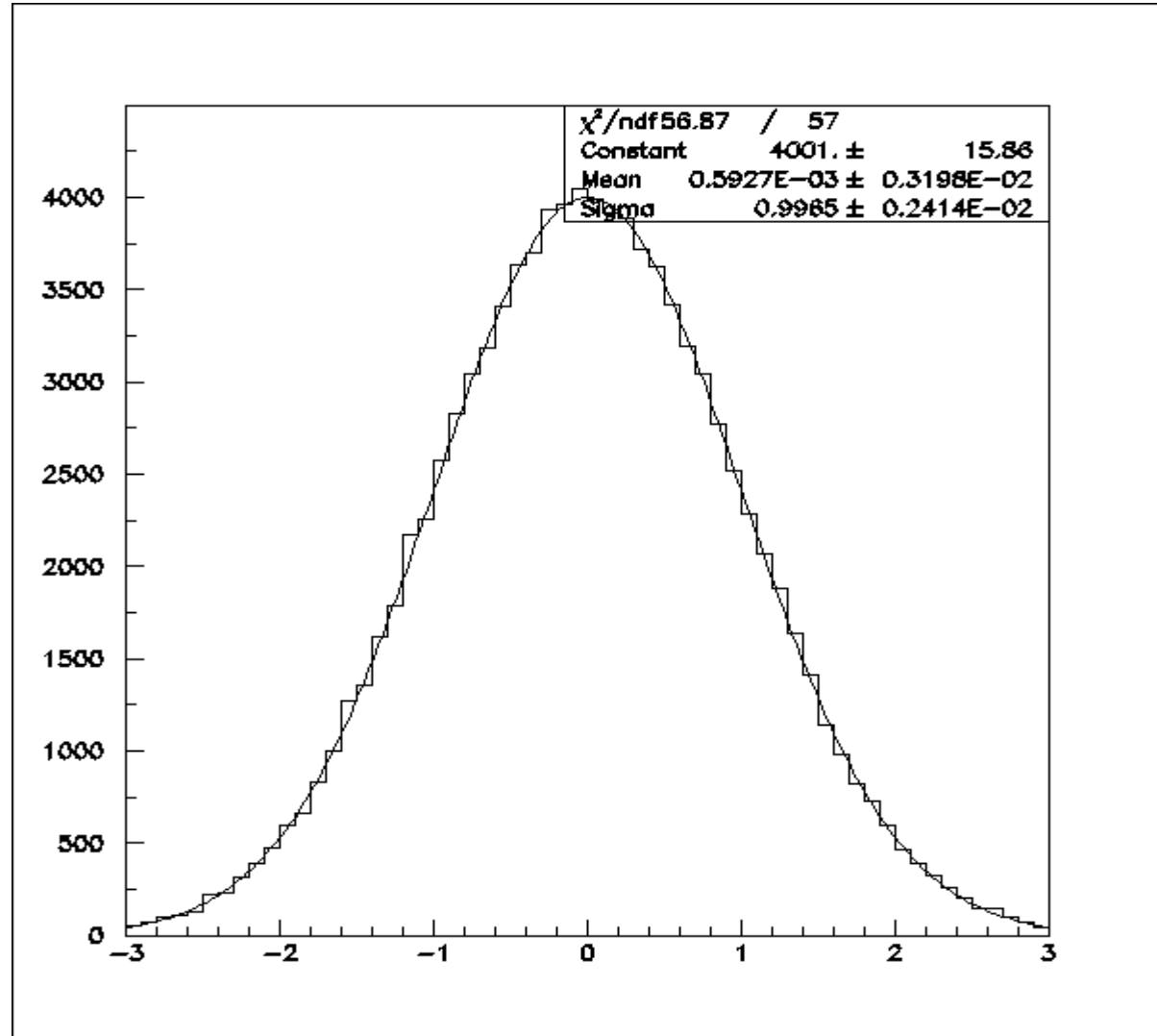
$$x^2 = \sum_{i=1}^k x_i^2 = x_1^2 + \dots + x_k^2$$

- Somma di variabili normali con valore medio 0 e sigma 1
- Segue una distribuzione ben nota e quindi ad un valore di χ^2 si associa un contenuto probabilistico



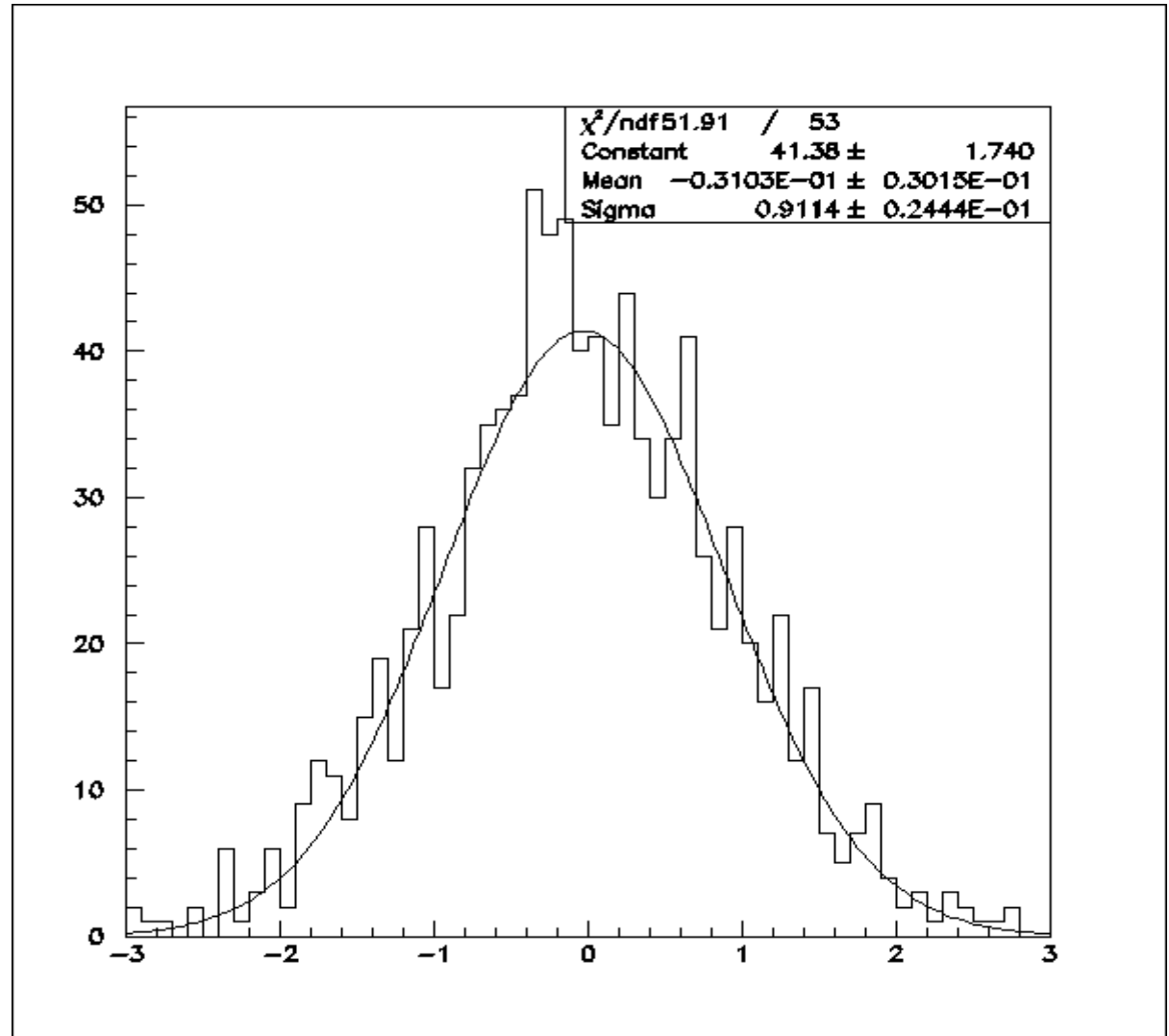
Fit gaussiano (2)

modificando il nostro generatore (sommando 48 numeri pseudo-random invece di 12 e dividendo per 2 il risultato) si ottiene una distribuzione gaussiana migliore: χ^2 per grado di liberta' = 56.7/57



Fit gaussiano (3)

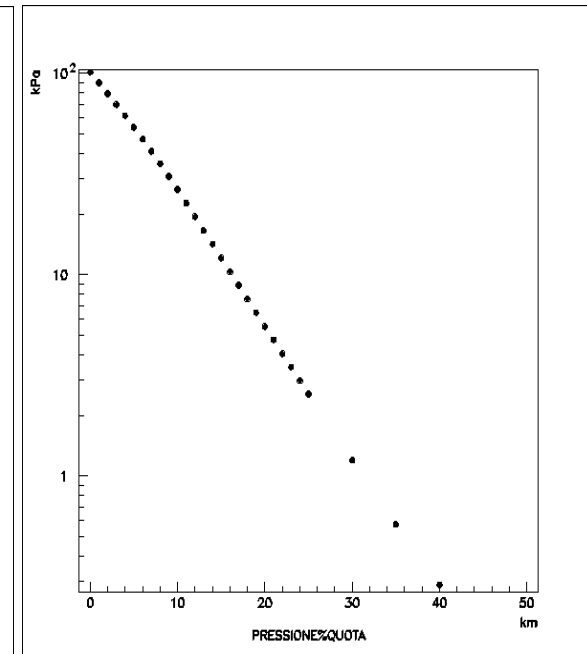
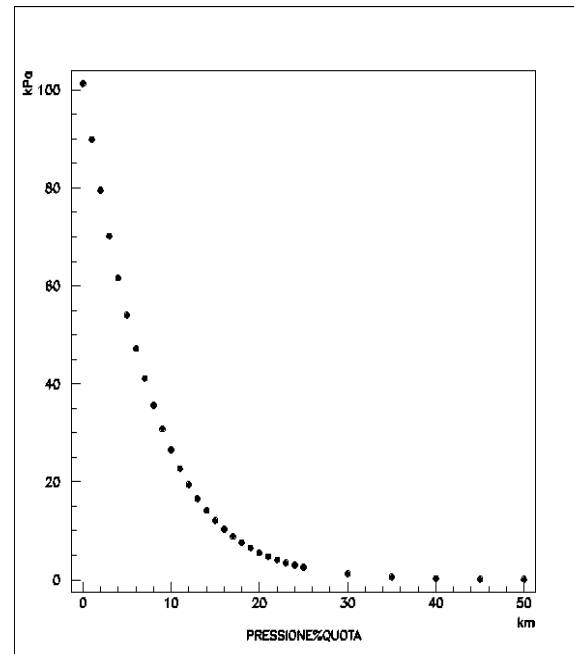
Se con la stessa versione del generatore estraiamo solo 1000 valori invece di 100000 otteniamo praticamente gli stessi risultati ma gli errori sono piu' grandi di un fattore $\sqrt{100}=10$. Cio' e' dovuto ai maggiori errori statistici sul contenuto N di ogni bin (i quali dipendono da \sqrt{N})



Esempio: pressione atmosferica

in molti casi e' possibile ricondursi mediante semplici trasformazioni a dipendenze lineari

Quota (km)	Pressione (kPa)	Quota	Pressione
0	101.325	15	12.111
1	89.876	16	10.350
2	79.501	17	8.850
3	70.121	18	7.565
4	61.660	19	6.467
5	54.048	20	5.529
6	47.217	21	4.729
7	41.105	22	4.048
8	35.651	23	3.467
9	30.800	24	2.972
10	26.499	25	2.549
11	22.699	30	1.197
12	19.399	35	0.575
13	16.579	40	0.287
14	14.170	45	0.149
		50	0.080



fit lineare

- il caso piu' semplice e piu' diffuso di fit e' quello lineare. La funzione che si suppone che possa descrivere i dati sperimentali e' una retta $y=Ax+B$.
- I dati sperimentali sono n coppie di punti (x_i, y_i) e le incertezze (errori) su queste grandezze. Per semplicita' assumiamo che le incertezze sulle x_i siano trascurabili rispetto a quelle sulle y_i che chiamiamo σ_i .
- I parametri A e B si determinano minimizzando rispetto a A e B la quantita' $\chi^2 = \sum_{i=1, n} ((y_i - (Ax_i + B)) / \sigma_i)^2$
- le formule risultanti nella versione piu' idonea per il calcolo numerico si trovano sul testo al paragrafo 8.1.1

Integrazione di funzioni

- Concentriamoci su integrali in una dimensione del tipo

$$I = \int_D f(x) dx$$

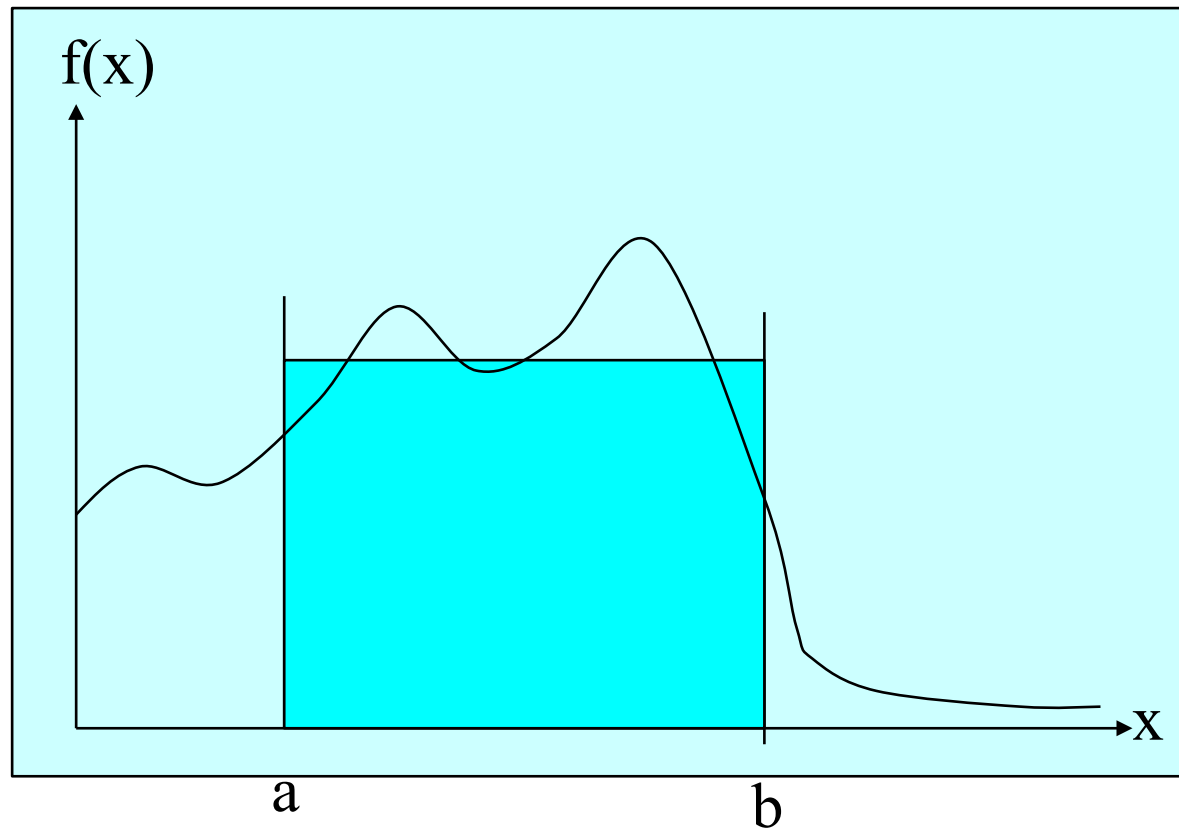
dove D è un dominio di integrazione definito (ad esempio l'intervallo $[a,b]$ dell'asse x)

- Esistono diversi metodi, in questa lezione mostreremo i più semplici (per una trattazione più completa si rinvia al cap.8 del testo).

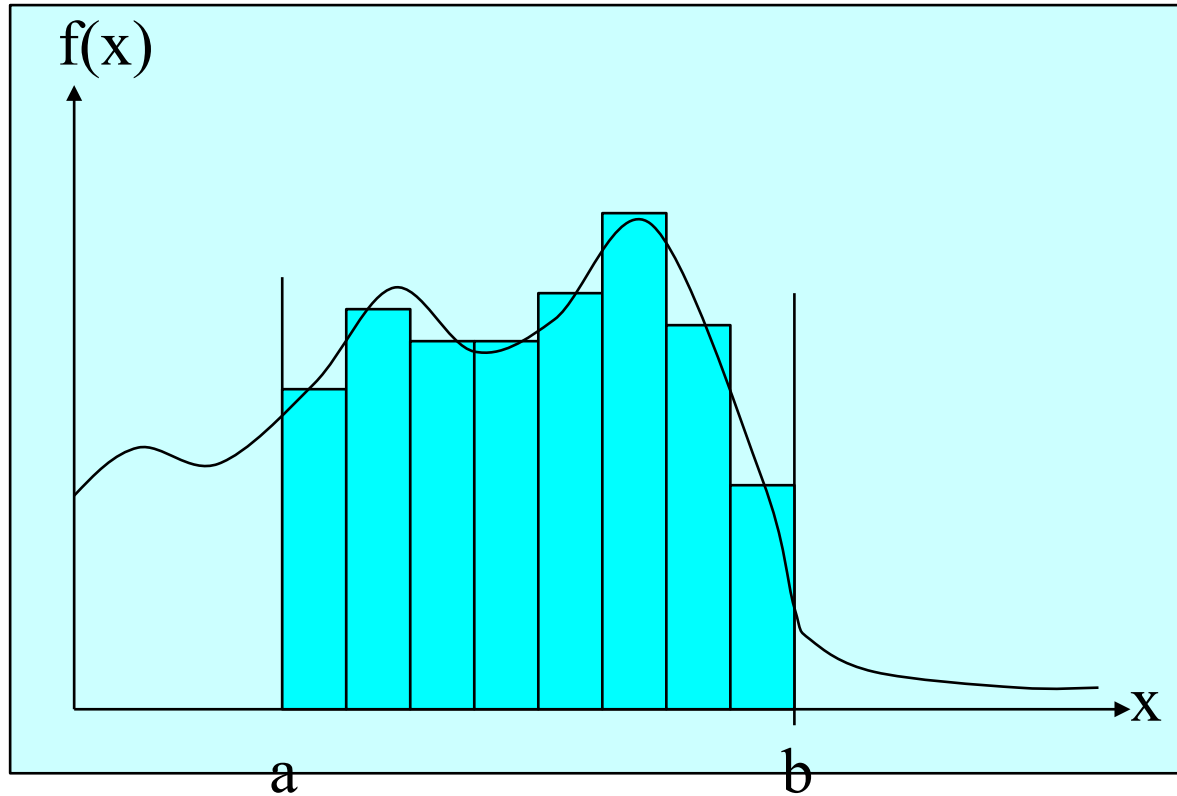
Metodo dei rettangoli

- Si approssima la funzione con un valore costante (polinomio di Lagrange di grado 0) in un intervallo $[a,b]$:
 $f(x) \approx f((a+b)/2)$ (caso particolare del metodo del punto di mezzo)
L'integrale esteso a tale intervallo vale
 $f((a+b)/2) (b-a)$
- L'errore commesso è pari all'integrale della $r(x)=f(x)- f((a+b)/2)$ ed è tanto più grande quanto più esteso è l'intervallo di integrazione (va col quadrato della dimensione dell'intervallo)
- Per ridurre l'errore si suddivide l'intervallo di integrazione in tanti intervallini.

Metodo dei rettangoli



Metodo dei rettangoli



Metodo dei trapezi

- Si approssima la funzione con un polinomio di Lagrange di grado 1 (retta) in un intervallo $[a,b]$:

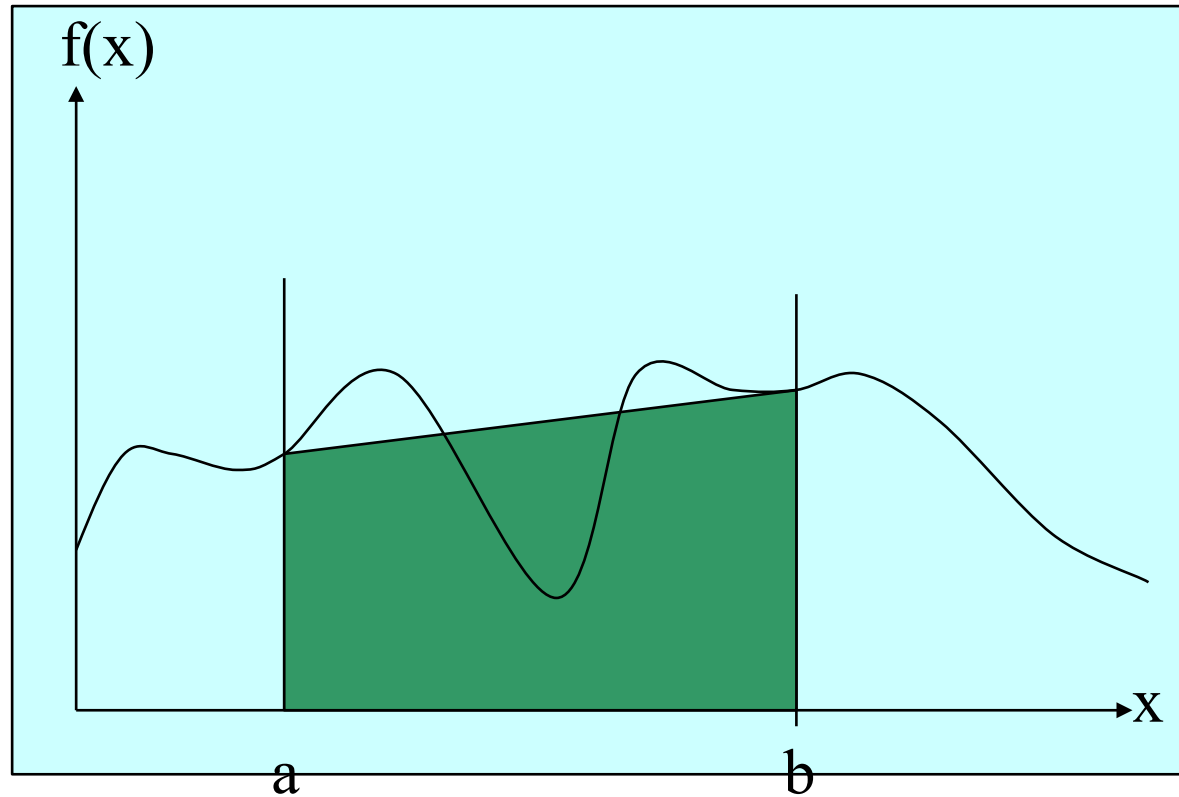
$$f(x) \approx \left((f(b)-f(a))*x + (b*f(a)-a*f(b)) \right) / (b-a)$$

L'integrale esteso a tale intervallo vale

$$(f(a)+f(b))*(b-a)/2$$

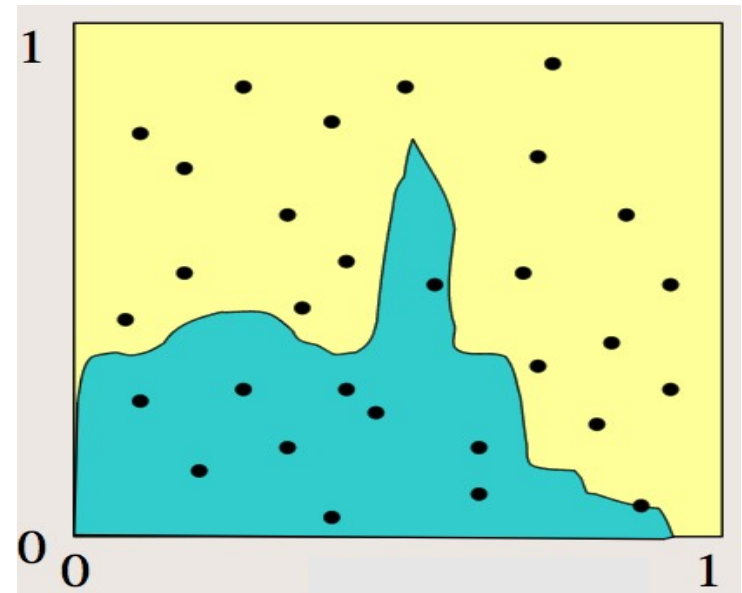
- Per l'errore valgono le considerazioni precedenti: l'errore è dell'ordine di $(a-b)$ al cubo per la derivata seconda della f calcolata in un punto interno all'intervallo.
- Per ridurre l'errore si suddivide ancora l'intervallo di integrazione in tanti intervallini.

Metodo dei trapezi



Metodo Monte Carlo: cos'e'?

- Metodo per la risoluzione numerica di problemi matematici mediante estrazione di numeri casuali
- Si applica sia a processi di natura statistica che deterministica
- Esempio: misura di un'area
- Sequenze di numeri casuali
 - Da processi caotici
 - Da processi stocastici
 - Da algoritmi (pseudo-random !)



Numeri pseudo-random: come si usano?

- Nelle librerie del C (usate anche dal C++) troviamo le funzioni `rand()` e `random()`
- Doc:
 - `man rand`
 - `man random`
- Uso: includere `<stdlib.h>`
- Attenzione: ritornano interi distribuiti uniformemente tra
 0 e $RAND_MAX=2^{31}-1=2147483647$
- Inizializzazione: `srand(int seed)`

Integrali con metodo Monte Carlo

- Come nella misura dell'area... ma servono 2 estrazioni per punto e ci sono inefficienze!
- Oppure se si estraggono casualmente N valori x_i della variabile x nell'intervallo $[a,b]$, per N sufficientemente grande e distribuzione uniforme (ma il metodo vale anche per diverse distribuzioni di probabilita') e' come se
 - si fosse suddiviso l'intervallo di integrazione in N intervallini di dimensioni simili, ed in media pari a $(b-a)/N$
 - si fosse scelto a caso un punto all'interno di ogni intervallo nel quale valutare la funzione $f(x)$
 - l'integrale della funzione f nell'intervallo $[a,b]$ puo' quindi essere approssimato con la sommatoria

$$\sum_{i=1,n} f(x_i) * (b-a)/N$$