993SM - Laboratory of Computational Physics lecture 6 April 6, 2022

Maria Peressi
Università degli Studi di Trieste - Dipartimento di Fisica
Sede di Miramare (Strada Costiera 11, Trieste)
e-mail: peressi@units.it
tel.: +39 040 2240242

Numerical integration - II

- other deterministic methods
- more on errors in deterministic methods

- comparison of efficiency between deterministic and Monte Carlo methods
- Central Limit Theorem and related algorithms

M. Peressi - UniTS - Laurea Magistrale in Physics
Laboratory of Computational Physics - Unit VI

Other deterministic methods

Numerical integration; other deterministic methods:

- in the simplest equally-spaced-point methods, we choose weights to calculate the average of the function:

$$\int_a^b f(x)dx \approx F_N = \sum_{i=1}^{N} v_i f(x_i)$$

rectangular rule: $x_i = a + \frac{b-a}{N} i, \qquad v_i = \frac{b-a}{N} \quad \forall i = 1, \ldots N-1$

trapezoidal rule: $x_i = a + \frac{b-a}{N} i, \qquad v_i = \frac{b-a}{N} \quad \forall i \neq 1, N; \quad v_1 = v_N = \frac{b-a}{2N}$

...

(at variance with these methods, in MC methods such as the 'importance sampling', we choose only points, not weights)

Numerical integration; other deterministic methods:

- in the simplest equally-spaced-point methods, we choose weights to calculate the average of the function:

$$\int_a^b f(x)dx \approx F_N = \sum_{i=1}^{N} v_i f(x_i)$$

idea: choose not only weights but also points: more degrees of freedom!

$$x_i = ?, \quad v_i = ?$$

Another deterministic method: Gaussian quadrature - I

Consider $\int_a^b f(x)dx$ and a function $W(x)$ defined on $[a,b]$

We can always formally write:

$$\int f(x)dx = \int W(x)F(x)dx \approx \sum_{j=1}^{N} w_j F(x_j) = \sum_{j=1}^{N} w_j \frac{f(x_j)}{W(x_j)} = \sum_{j=1}^{N} v_j f(x_j)$$

with $v_j = w_j / W(x_j)$

to be determined, depending on $W(x)$ (*)

(This will be convenient in particular if the resulting $F(x)$ is smooth, but not necessarily)

(*) in general: $w_j \neq W(x_j)$

Another deterministic method: Gaussian quadrature - II

Consider $\int_a^b f(x)dx$ and a function $W(x)$ defined on $[a,b]$

We can always formally write:

$$\int f(x)dx = \int W(x)F(x)dx \overset{(*)}{\approx} \sum_{j=1}^{N} w_j F(x_j) = \sum_{j=1}^{N} w_j \frac{f(x_j)}{W(x_j)} = \sum_{j=1}^{N} v_j f(x_j)$$

For a given $W(x)$, the N points and weights $\{x_j\}, \{w_j\}$ can be chosen to make the approximate relationship (*) an exact equality if $F(x)$ is a 2N-1 degree polynomial.

Another deterministic method: Gaussian quadrature - III

Consider $$\int_{x_1}^{x_2} W(x)F(x)dx = \sum_{j=1}^{N} w_j F(x_j)$$

$F(x)$ a 2N - 1 degree polynomial. Which are the N $\{x_j\}, \{w_j\}$?

If there is a set of polynomials $\{p_N(x)\}$ which are orthogonal in the same interval and for the same weight function $W(x)$:

$$\langle p_N | p_{N'} \rangle_W = \delta_{N,N'} \ , \ i.e., \ \int_{x_1}^{x_2} W(x) p_N(x) p_{N'}(x) dx = \delta_{N,N'}$$

the points $\{x_i\}$ are exactly the roots of the $p_N(x)$ polynomials. The weights $\{w_j\}$ are related to them, but in general $w_j \neq W(x_j)$.

Gauss-Legendre quadrature

Consider $\int_{x_1}^{x_2} W(x)F(x)dx = \sum_{j=1}^{N} w_j F(x_j)$

with $F(x)$ a 2N - 1 degree polynomial.

If: $W(x) = 1$ and $x_1 = -1, \quad x_2 = 1$,

the Legendre polynomials $\{P_N(x)\}$ defined by:

$$(j+1)P_{j+1} = (2j+1)xP_j - jP_{j-1}$$

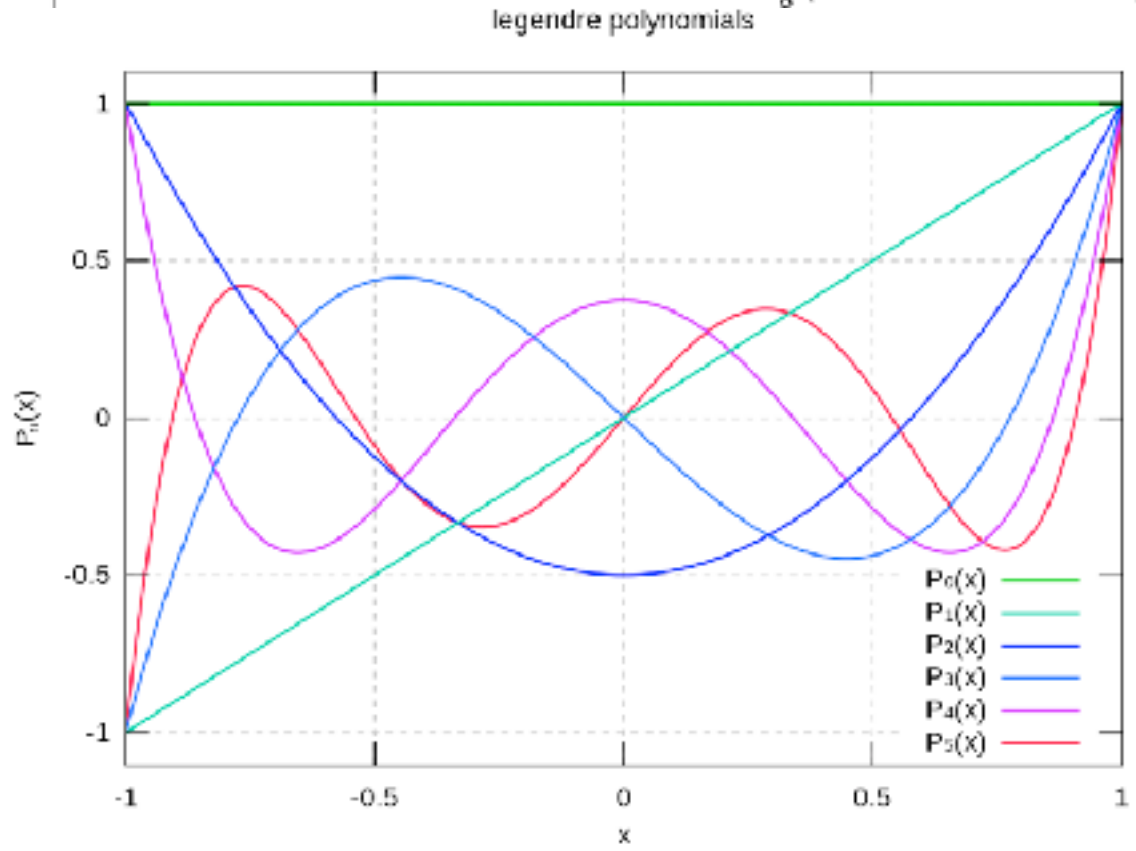
are orthogonal in [-1,1] with $W(x)$=1;

$\{x_j\}, \{w_j\}$ are such that $P_N(x_i) = 0$ and

$$w_i = \frac{2}{(1-x_i^2)[P'_N(x_i)]^2}$$

The first few Legendre polynomials are:

n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}(3x^2-1)$
3	$\frac{1}{2}(5x^3-3x)$
4	$\frac{1}{8}(35x^4-30x^2+3)$
5	$\frac{1}{8}(63x^5-70x^3+15x)$



polynomials are odd or even in x ⇨ roots are even

Legendre polynomials in Physics: examples of applications

1) For a polynomial expansion of a gravitational or coulombic potential:

$$\frac{1}{|\mathbf{x}-\mathbf{x}'|} = \frac{1}{\sqrt{r^2+r'^2-2rr'\cos\gamma}} = \sum_{\ell=0}^{\infty} \frac{r'^{\ell}}{r^{\ell+1}} P_\ell(\cos\gamma)$$

where r and r' are the lengths of the vectors $\mathbf{x}$ and $\mathbf{x}'$ respectively and γ is the angle between those two vectors.

2) solution of Laplace's equation of the static potential, $\nabla^2\Phi(\mathbf{x}) = 0$, in a charge-free region of space, if the boundary conditions have axial symmetry :

θ is the angle between the position of the observer and the $\hat{\mathbf{z}}$ axis (the zenith angle) ;

the solution for the potential will be

$$\Phi(r,\theta) = \sum_{\ell=0}^{\infty} \left[A_\ell r^\ell + B_\ell r^{-(\ell+1)}\right] P_\ell(\cos\theta)$$

3) solving Schrödinger equation in three dimensions for a central force :

the **associated Legendre polynomials** are derivatives of ordinary Legendre polynomials ($m \geq 0$)

$$P_\ell^m(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} (P_\ell(x))$$

The Legendre polynomials are closely related to the spherical harmonics

$$Y_{\ell,m}(\theta,\phi) = \sqrt{\frac{(2\ell+1)(\ell-m)!}{4\pi(\ell+m)!}} P_\ell^m(\cos\theta)\, e^{im\phi} \qquad -\ell \leq m \leq \ell.$$

Gaussian quadrature

In practice, we choose W(x) and N and use the set of N points and weights $\{x_j\}, \{w_j\}$ for the approximate integration:

$$\int f(x)dx = \int W(x)F(x)dx \approx \sum_{j=1}^{N} w_j F(x_j) = \sum_{j=1}^{N} w_j \frac{f(x_j)}{W(x_j)} = \sum_{j=1}^{N} v_j f(x_j)$$

Gauss-Legendre quadrature

For: $x_1 = -1, \quad x_2 = 1$

(see slide 10 for the list of the first polynomials)

N	i	x_i	w_i	degree
1	1	0	2	1
2	1	-0.577350269189626	1	3
	2	0.577350269189626	1	
3	1	-0.774596669241483	0.555555555555556	5
	2	0	0.888888888888889	
	3	0.774596669241483	0.555555555555556	
4	1	-0.861136311594053	0.347854845137454	7
	2	-0.339981043584856	0.652145154862546	
	3	0.339981043584856	0.652145154862546	
	4	0.861136311594053	0.347854845137454	

degree of the polynomial exactly integrable

The integration in an interval [a,b] different from [-1,1] ("old") can be easily done performing the scaling:

$$x_{new} = \frac{b-a}{2} x_{old} + \frac{b+a}{2} \quad \text{and} \quad w_{new} = \frac{b-a}{2} w_{old}$$

Gauss-Legendre quadrature

In case of classical, well known, orthogonal polynomials, ready-to-use subroutines exist for the computation of Abscissas and Weights $\{x_j\}, \{w_j\}$

e.g. GAULEG(x_1, x_2, x, w, n) of *Numerical Recipes* which, given x_1, x_2, n, provides as output the arrays $x(n), w(n)$

Some programs:

on https://moodle2.units.it/

gauleg-IIorder.f90

gauleg-others.f90 (generation of points up to 15 points in [-1,1] using GAULEG adapted from "Numerical Recipes" (self-contained) and some tests for easy-to-integrate functions)

In the subdirectory: **gauss-nr90/**

find the original routine from "Numerical Recipes" and related external routines/modules/interfaces and a main program for test (see following slide)

Gauss-Legendre from Numerical Recipes

Use of GAULEG:

In order to use the routines of Numerical Recipes, you have to compile and link the main program with:

- the subroutine **gauleg.f90** which gives points and abscissas
- **nrtype.f90** containing type declarations;
- **nrutil.f90** containing **moduli** and utilities;
- **nr.f90** containing (through the interfaces) the conventions

to call the subroutines with the main program

i) You must first compile these files with the option "-c":
this produces .mod and .o (the objects).

ii) In a second step compile the main program.

iii) Finally you link all the files *.o and produce the executable:

gfortran -c nrtype.f90 nrutil.f90 nr.f90 gauleg.f90
gfortran -c gauleg_nr_test.f90
gfortran -o a.out gauleg_nr_test.o nrtype.o nrutil.o nr.o gauleg.o

gauleg.f90 from Numerical Recipes

```
SUBROUTINE gauleg(x1,x2,x,w)
  USE nrtype; USE nrutil, ONLY : arth,assert_eq,nrerror
  IMPLICIT NONE
  REAL(SP), INTENT(IN) :: x1,x2
  REAL(SP), DIMENSION(:), INTENT(OUT) :: x,w
  REAL(DP), PARAMETER :: EPS=3.0e-14_dp
  INTEGER(I4B) :: its,j,m,n
  INTEGER(I4B), PARAMETER :: MAXIT=10
  REAL(DP) :: xl,xm
  REAL(DP), DIMENSION((size(x)+1)/2) :: p1,p2,p3,pp,z,z1
  LOGICAL(LGT), DIMENSION((size(x)+1)/2) :: unfinished
  n=assert_eq(size(x),size(w),'gauleg')
  m=(n+1)/2
  xm=0.5_dp*(x2+x1)
  xl=0.5_dp*(x2-x1)
  z=cos(PI_D*(arth(1,1,m)-0.25_dp)/(n+0.5_dp))
     ...
     ...
     ...
  x(1:m)=xm-xl*z
  x(n:n-m+1:-1)=xm+xl*z
  w(1:m)=2.0_dp*xl/((1.0_dp-z**2)*pp**2)
  w(n:n-m+1:-1)=w(1:m)
END SUBROUTINE gauleg
```

`nrtype.f90` from Numerical recipes

```
MODULE nrtype
  INTEGER, PARAMETER :: I4B = SELECTED_INT_KIND(9)
  INTEGER, PARAMETER :: I2B = SELECTED_INT_KIND(4)
  INTEGER, PARAMETER :: I1B = SELECTED_INT_KIND(2)
  INTEGER, PARAMETER :: SP = KIND(1.0)
  INTEGER, PARAMETER :: DP = KIND(1.0D0)
  INTEGER, PARAMETER :: SPC = KIND((1.0,1.0))
  INTEGER, PARAMETER :: DPC = KIND((1.0D0,1.0D0))
  INTEGER, PARAMETER :: LGT = KIND(.true.)
  REAL(SP), PARAMETER :: PI=3.141592653589793238462643383279502884197_sp
  REAL(SP), PARAMETER :: PIO2=1.570796326794896619231321691639751442098584_sp
  REAL(SP), PARAMETER :: TWOPI=6.283185307179586476925286766559005768394_sp
  REAL(SP), PARAMETER :: SQRT2=1.414213562373095048801688724209698078569671_sp
  REAL(SP), PARAMETER :: EULER=0.577215664901532860606512090082402431042159_sp
  REAL(DP), PARAMETER :: PI_D=3.141592653589793238462643383279502884197_dp
  REAL(DP), PARAMETER :: PIO2_D=1.570796326794896619231321691639751442098584_dp
  REAL(DP), PARAMETER :: TWOPI_D=6.283185307179586476925286766559005768394_dp
  TYPE sprs2_sp
     INTEGER(I4B) :: n,len
     REAL(SP), DIMENSION(:), POINTER :: val
     INTEGER(I4B), DIMENSION(:), POINTER :: irow
     INTEGER(I4B), DIMENSION(:), POINTER :: jcol
  END TYPE sprs2_sp
  TYPE sprs2_dp
     INTEGER(I4B) :: n,len
     REAL(DP), DIMENSION(:), POINTER :: val
     INTEGER(I4B), DIMENSION(:), POINTER :: irow
     INTEGER(I4B), DIMENSION(:), POINTER :: jcol
  END TYPE sprs2_dp
END MODULE nrtype
```

nr.f90 from Numerical Recipes

```
MODULE nr
  INTERFACE
     SUBROUTINE gauleg(x1,x2,x,w)
       USE nrtype
       REAL(SP), INTENT(IN) :: x1,x2
       REAL(SP), DIMENSION(:), INTENT(OUT) :: x,w
     END SUBROUTINE gauleg
  END INTERFACE
  ! ... the original file contains several other  INTERFACES ...
END MODULE nr
```

nrutil.f90 (Here only for: array_copy, arth, assert_eq, nrerror)

```
MODULE nrutil
  USE nrtype
  IMPLICIT NONE
  INTEGER(I4B), PARAMETER :: NPAR_ARTH=16,NPAR2_ARTH=8

    ...
    ...
    ...

  INTERFACE array_copy
     MODULE PROCEDURE array_copy_r, array_copy_d, array_copy_i
  END INTERFACE

    ...
  ! ... l'originale contiene ancora molte altre INTERFACEs....
CONTAINS

  SUBROUTINE array_copy_r(src,dest,n_copied,n_not_copied)
    REAL(SP), DIMENSION(:), INTENT(IN) :: src

    ...


   !  .... and many other FUNCTIONs and SUBROUTINEs ....
 END MODULE nrutil
```

Summary of numerical integration (MC and deterministic) methods

MC sample mean

$$\int_a^b f(x)dx = (b-a) < f > \approx (b-a)\frac{1}{N}\sum_{i=1}^{N} f(x_i) \quad with\ \{x_i\}\ randomly\ uniformly\ distributed\ in\ [a,b]$$

$$\left(it\ can\ be\ considered\ as\ Importance\ sampling\ with\ p(x) = \frac{1}{b-a}\ in\ [a,b]\right)$$

MC importance sampling

$$\int_a^b f(x)dx = \int_a^b \frac{f(x)}{p(x)}p(x)dx = < \frac{f(x)}{p(x)} > \int_a^b p(x)dx \approx \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)}\int_a^b p(x)dx$$

$$with\ \{x_i\}\ \ randomly\ \ distributed\ \ according\ \ p(x)$$

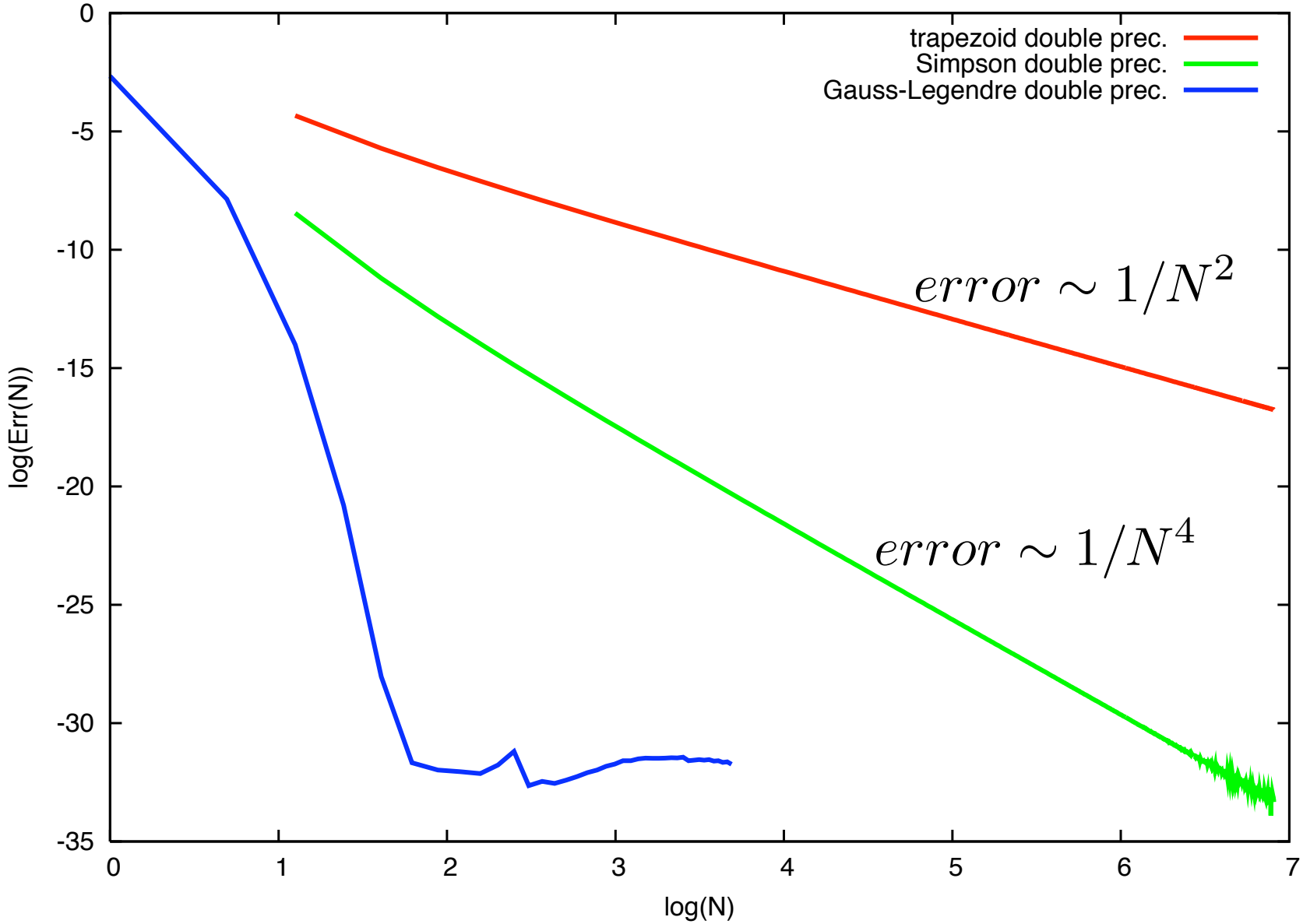
Deterministic, equispaced points

$$\int_a^b f(x)dx \approx \sum_{i=1}^{N} v_i f(x_i) \quad with\ x_i = a + \frac{b-a}{N}i\ ,\ v_i\ to\ be\ determined$$

Deterministic, non equispaced points

$$\int_a^b f(x)dx \approx \sum_{i=1}^{N} v_i f(x_i) \quad with\ \{x_i\}\ ,\ \{v_i\}\ to\ be\ determined$$

Numerical integration, deterministic methods: comparison of errors in 1D



(double precision needed to appreciate the convergence of Gauss-Legendre numerical estimate)

Error estimate: comparison between deterministic and MC methods in d-dimension

Error estimate for numerical integration with deterministic methods

(Reminder from previous Lecture)

$$\int f(x)dx = F_n + error$$

How to evaluate the error? Consider the Taylor expansion of the integrand function and then integrate:

$$f(x) = f(x_i) + f'(x_i)(x - x_i) + \frac{1}{2}f''(x_i)(x - x_i)^2 + \dots ,$$

$$\int_{x_i}^{x_{i+1}} f(x)\,dx = f(x_i)\Delta x + \frac{1}{2}f'(x_i)(\Delta x)^2 + \frac{1}{6}f''(x_i)(\Delta x)^3 + \dots (*)$$

$$\Delta x \equiv x_{i+1} - x_i$$

Error estimate for numerical integration: Rectangular approximation

(Reminder from previous Lecture)

$$\int_{x_i}^{x_{i+1}} f(x)dx \approx \boxed{f(x_i)\Delta x}$$

Compare ▭ with (*):

$$\int_{x_i}^{x_{i+1}} f(x)\,dx = f(x_i)\Delta x + \frac{1}{2}f'(x_i)(\Delta x)^2 + \frac{1}{6}f''(x_i)(\Delta x)^3 + \dots$$

error

(leading order in Δx)

For n intervals $(\Delta x = (b-a)/n)$: error is $n(\Delta x)^2 \sim 1/n$

(...and similarly for higher-order approximations)

Numerical integration:
multidimensional integrals

$$F = \int_R f(x,y)\, dxdy$$

The rectangular approximation gives $\Delta x \Delta y \sim (\Delta x)^2 \sim 1/n$, being n the number of parts (or pairs of points) of the integration domain:

$$\int_{x_i}^{x_{i+1}} \int_{y_i}^{y_{i+1}} f(x,y) dxdy \approx f(x_i, y_i)\Delta x \Delta y \qquad (*)$$

The Taylor expansion of the integrand function gives:

$$f(x,y) = f(x_i,y_i) + f'_x(x_i,y_i)(x-x_i) + f'_y(x_i,y_i)(y-y_i) + \dots$$

$$\int_{x_i}^{x_{i+1}} \int_{y_i}^{y_{i+1}} f(x,y) dxdy = f(x_i,y_i)\Delta x\Delta y + f'_x(x_i,y_i)\frac{(\Delta x)^2}{2}\Delta y + f'_y(x_i,y_i)\Delta x\frac{(\Delta y)^2}{2} + \dots \quad (**)$$

(*) against (**) => error
(leading order in Δx)

For n intervals: error is $n(\Delta x)^3 \sim 1/n^{1/2}$

Numerical integration: multidimensional integrals

Therefore for rectangular approx.:

d=1: $error \sim 1/n$ d=2: $error \sim 1/n^{1/2}$

In general:

if the error decreases as n^{-a} for $d = 1$, then the error decreases as $n^{-a/d}$ in d dimensions.

Classical formulas with equispaced points: slowly decreasing error for multidimensional integration !

Numerical integration:
error in MC methods

$$\sigma_n/\sqrt{n} \approx \sigma_m \approx \sigma_s/\sqrt{s}$$

(σ_n is roughly constant with n ; for uncorrelated points, the variance of the averages goes like $\sim 1/n^{1/2}$)

- The average function value

$$\langle f \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

- The average squared function value

$$\langle f^2 \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} f^2(x_i)$$

- Estimate of the integrand (+/- standard error)

$$\int f \, dV \approx V \langle f \rangle \pm V \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N}}$$

Numerical integration:
errors in multidimensional integrals

d	Rect.	Trap.	Simps.	MC
1	$1/n$	$1/n^2$	$1/n^4$	$1/n^{1/2}$
2	$1/n^{1/2}$	$1/n$	$1/n^2$	$1/n^{1/2}$
4	$1/n^{1/4}$	$1/n^{1/2}$	$1/n$	$1/n^{1/2}$
...	...	...	...	

if the error decreases as n^{-a} for $d = 1$, then the error decreases as $n^{-a/d}$ in d dimensions.

the error for all Monte Carlo integration methods decreases as $n^{-1/2}$ independently of the integral.

Monte Carlo convenient for multidimensional integration !

Summary: advantages of MC integration methods

- convergence as ~$N^{1/2}$ in any dimension regardless of the smoothness of the integrand
- simplicity: only two simple steps required (namely, producing a set of sampling points and evaluating the integrand function over such points)
- generality: sampling can be used even on domains that do not have a natural correspondence with the 'standard' domain $[0,1]^d$ and thus are not well-suited to numerical quadrature
- better suited than quadrature for integrands with singularities (importance sampling can handle this problem)
- flexibility: easy to add more points as needed (in the Gaussian quadrature, increasing the accuracy implies doing calculations from scratch)

The central limit theorem

The central limit theorem

- the variance of the averages
- the Gaussian distribution

Consider a continuous random variable x with probability density $f(x)$.

characterized by $\langle x^m \rangle = \int x^m f(x)\, dx$ and ${\sigma_x}^2 = \langle x^2 \rangle - \langle x \rangle^2$.

Consider y s.t. y_n corresponding to the average of n values of x:

$$y = y_n = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

Suppose that we make many measurements of y. The variable y is distributed according to a probability density $P(y) \neq f(x)$

quantities of interest are the mean $\langle y \rangle$, the variance ${\sigma_y}^2 = \langle y^2 \rangle - \langle y \rangle^2$, and $P(y)$ itself.

The central limit theorem

The random variable:

$$y = \quad y_n = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

is distributed according to:

$$P(y): \text{ gaussian distribution}$$

with: $< y > = < x >$ $\qquad \sigma_y \approx \sigma_x / \sqrt{n}$

(Therefore, the sample mean of a random sample is better than a single observation)

$$provided\ \langle x \rangle\ and\ \langle x^2 \rangle\ exist\ (finite)\ and\ n\ is\ large!$$

The central limit theorem

Analogously, is instead of considering the new random variable as the **average** we consider just the **sum:**

$$y = x_1 + x_2 + \ldots + x_n$$

it also has a gaussian distribution but with:

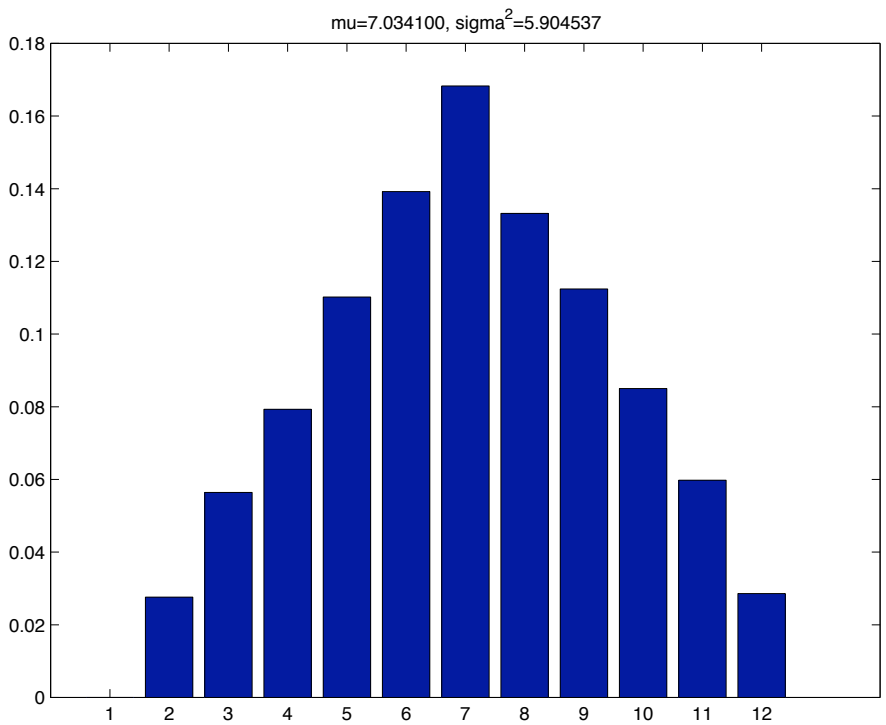
$$< y > = n < x > \quad \text{and} \quad \sigma_y \approx \sqrt{n}\ \sigma_x$$

$provided\ \langle x \rangle\ and\ \langle x^2 \rangle\ exist\ (finite)\ and\ n\ is\ large!$

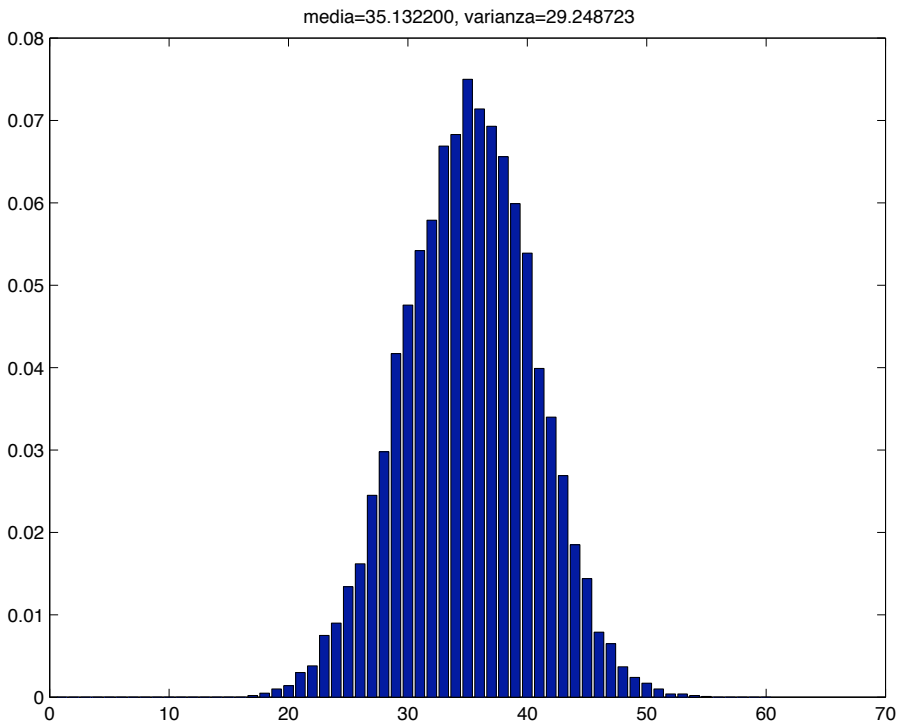
The central limit theorem

Note: large enough n needed to obtain the gaussian distribution.
Suppose that f(x) is uniform: e.g., playing dice:

n=2 not enough



n=100 OK



The central limit theorem

The previous example was for UNIFORM distribution (dice)
but the central limit theorem work also with random deviates x
with NON UNIFORM distribution; e.g. with exponential distribution:

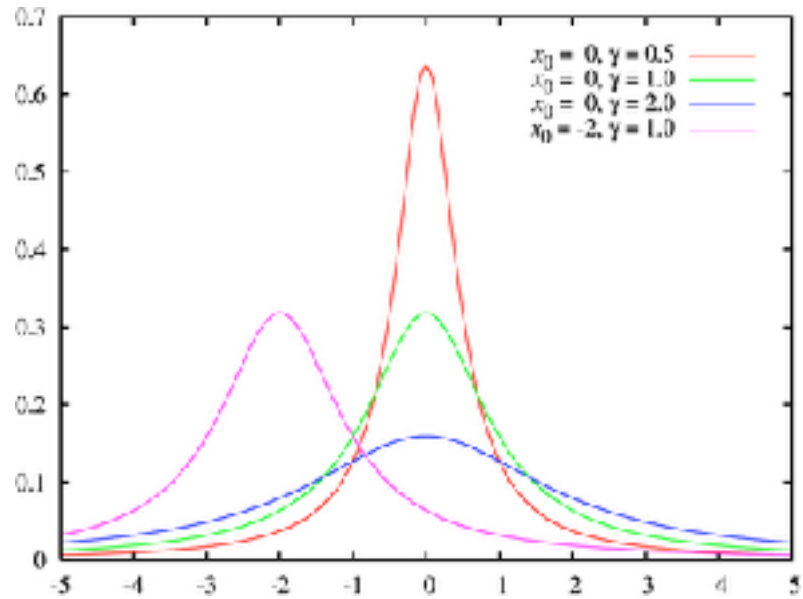
$$f(x) = \begin{cases} e^{-x}, & \text{if } x \geq 0 \\ 0, if & x < 0 \ . \end{cases}$$

(NOTE: the central limit theorem gives therefore another
operative method to generate random numbers with a gaussian distribution)

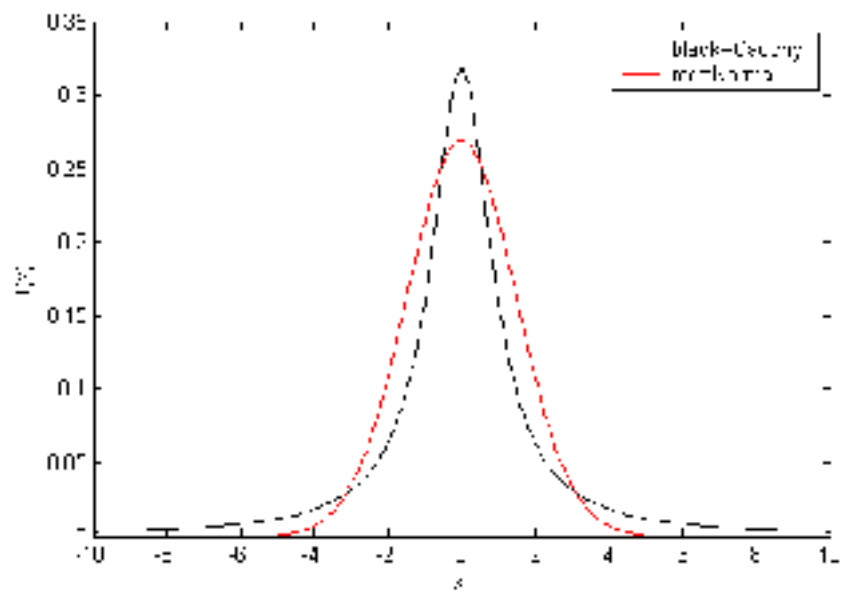
...but sometimes it doesn't work:

Cauchy-Lorentz
probability density function

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]}$$
$$= \frac{1}{\pi}\left[\frac{\gamma}{(x - x_0)^2 + \gamma^2}\right]$$



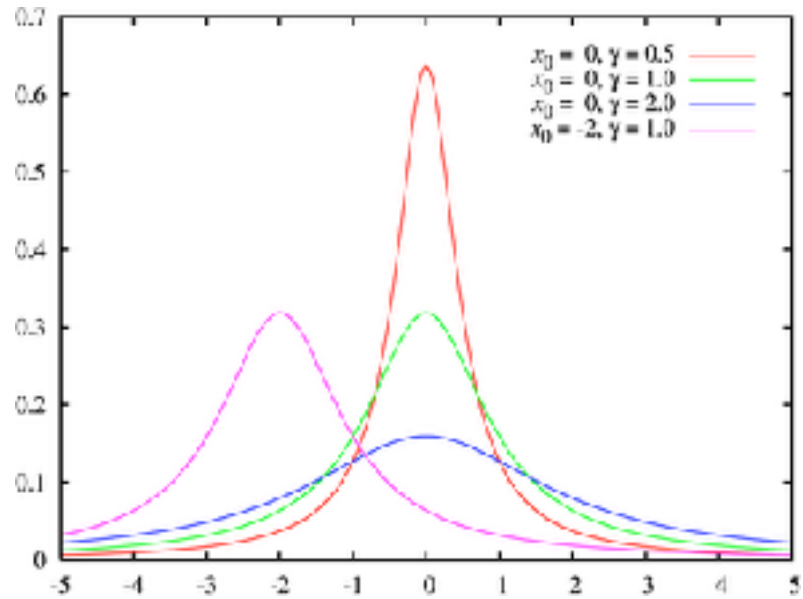
The Cauchy-Lorentz distribution is an example of "fat-tailed" distribution.
Fat-tailed distributions decay to infinity slower than exponentially.
For instance, they can decay with a power law: $f(x) \sim x^{-(1+\alpha)}$ as $x \longrightarrow +\infty$
In some cases the expression "fat-tailed" indicates distributions where $0 < \alpha < 2$.



...but sometimes it doesn't work:

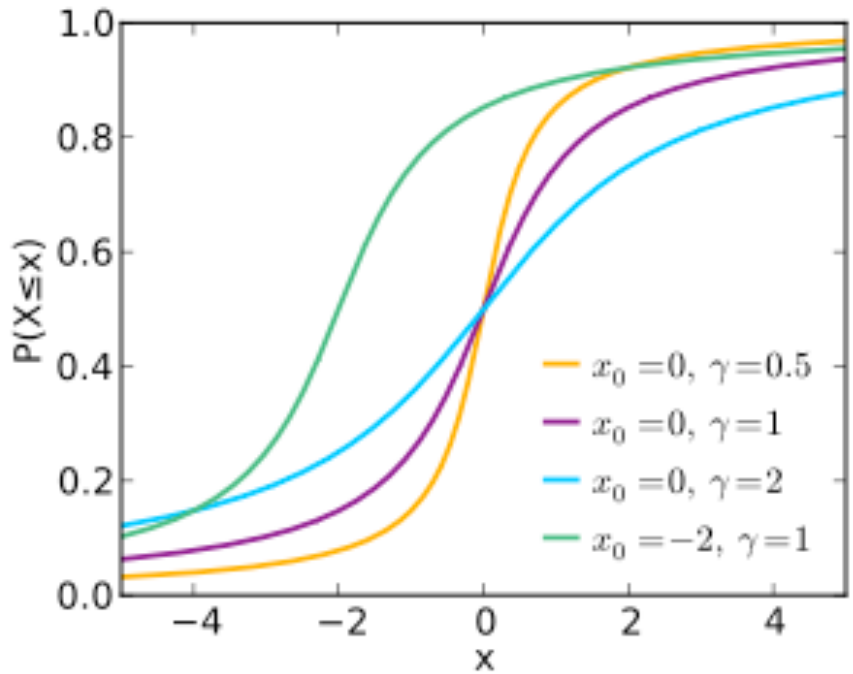
Cauchy-Lorentz
probability density function

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]}$$
$$= \frac{1}{\pi}\left[\frac{\gamma}{(x - x_0)^2 + \gamma^2}\right]$$



Cumulative distribution :

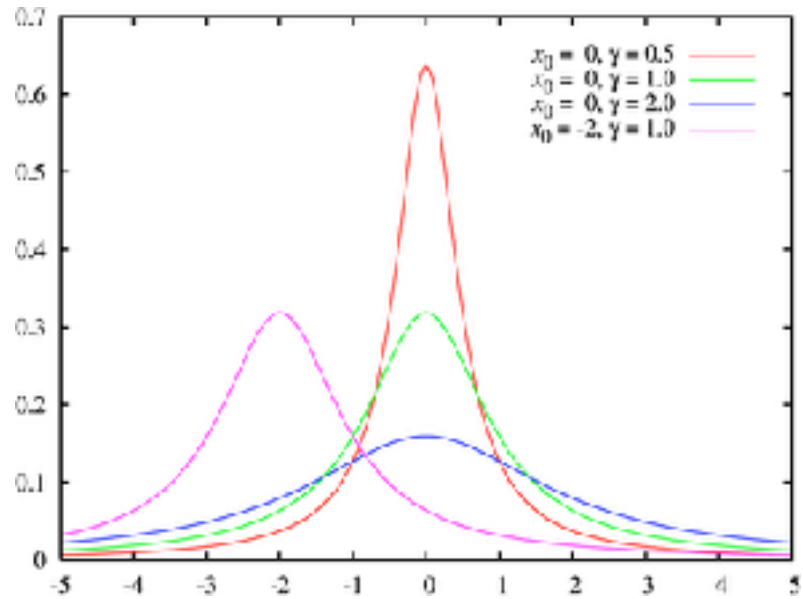
$$\frac{1}{\pi}\arctan\left(\frac{x - x_0}{\gamma}\right) + \frac{1}{2}$$



...but sometimes it doesn't work:

Cauchy-Lorentz
probability density function

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]}$$
$$= \frac{1}{\pi}\left[\frac{\gamma}{(x - x_0)^2 + \gamma^2}\right]$$



Mean and variance are **not** defined

The mean: $\int_{-\infty}^{\infty} x f(x)\, dx$ which can be rewritten as: $\int_0^{\infty} x f(x)\, dx - \int_{-\infty}^{0} |x| f(x)\, dx$
$(here\ for\ x_0 = 0)$

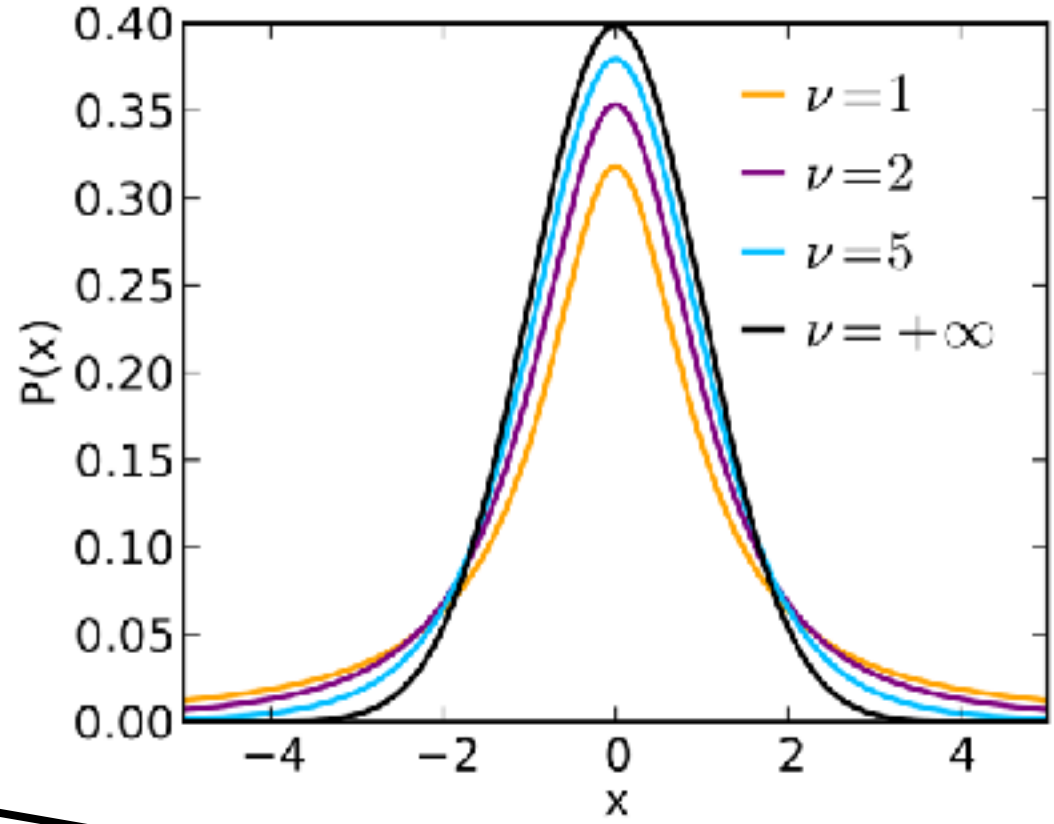
is not defined since both terms are infinite; only the Cauchy principal value is defined:

$$\lim_{a \to \infty} \int_{-a}^{a} x f(x)\, dx$$

Without a defined mean, it is impossible to define the variance (but the second moment is defined and it is infinite). Some results in probability theory about expected values, such as the law of large numbers, do not work in such cases.
Also, the mean of a set of random variates drawn from a Cauchy distribution is no better than a single observation, because the chance of including extreme values is high.

Student's t-distribution

(più generale)



$$f(t_n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\,\Gamma\left(\frac{n}{2}\right)} \cdot \left(1 + \frac{t^2}{n}\right)^{-\frac{n-1}{2}}$$

Notiamo che il limite di questa successione di funzioni per $n \to \infty$ è

$$\lim_{n\to\infty} f(t_n) = \frac{1}{\sqrt{\pi}} \lim_{n\to\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n}\,\Gamma\left(\frac{n}{2}\right)} \lim_{n\to\infty} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

Sapendo che il primo limite ha come risultato $\frac{1}{\sqrt{2}}$ e il secondo tende a $e^{-\frac{t^2}{2}}$.

$$\lim_{n\to\infty} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \sim \frac{1}{t^{-(n+1)}}$$

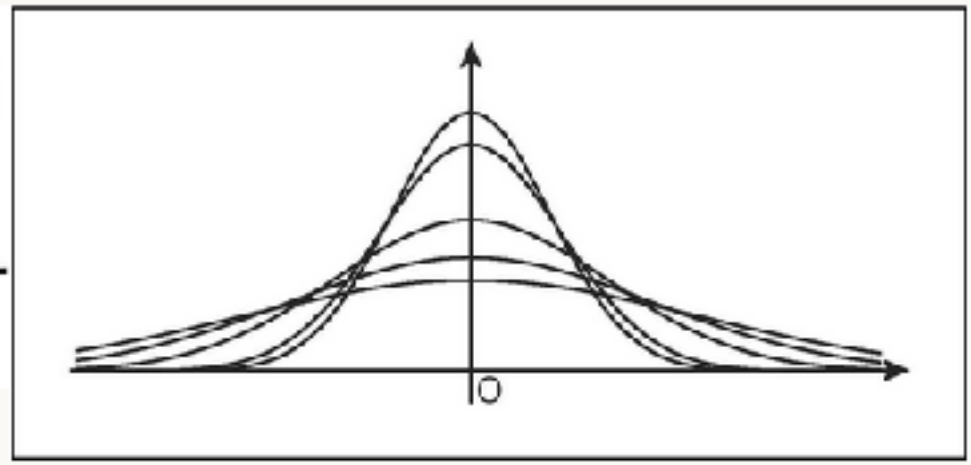
Varianza	$\frac{n}{n-2}$ se $n > 2$ infinita altrimenti

In pratica, prendendo una popolazione di numerosità N molto grande, la variabile aleatoria t tende ad essere una normale standard.

Curtosi (dal greco *kurtós*, gobba) in statistica, termine che indica quanto una distribuzione di dati si allontani da una curva normale standardizzata (cioè se è, rispetto a questa, per la quale l'indice è 0, più "schiacciata" o meno "schiacciata"). L'*indice di curtosi* per una distribuzione discreta *X* di *n* elementi è dato da:

$$Curt(x) = \frac{\sum_{1}^{n}(x_i - \mu)^4}{n\sigma^4}$$

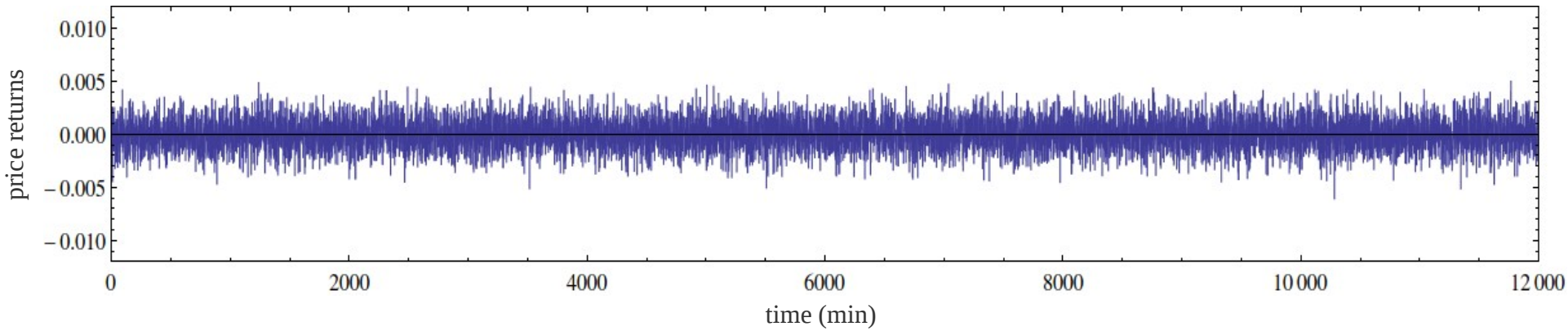
formula



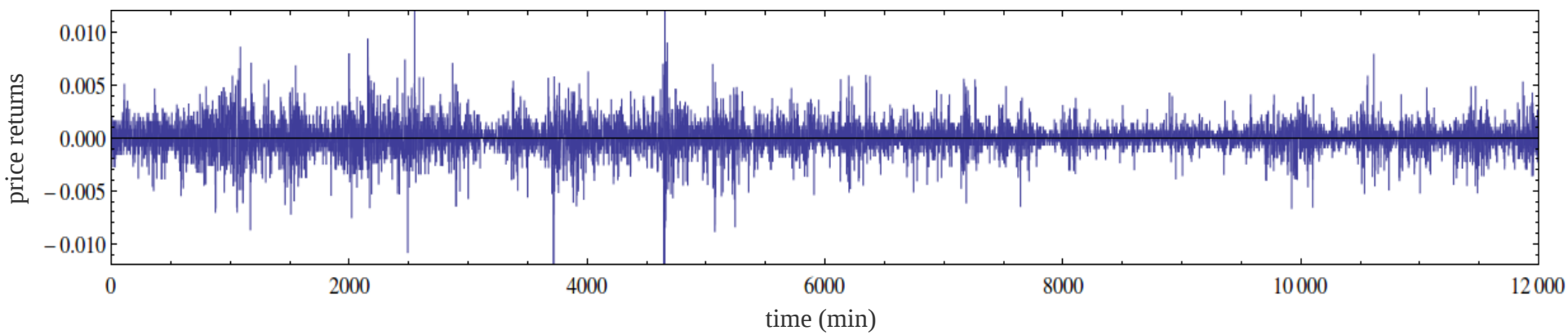
CURTOSI

Statistical Properties of Price Returns

Simulated Returns (Geometric Brownian Motion)



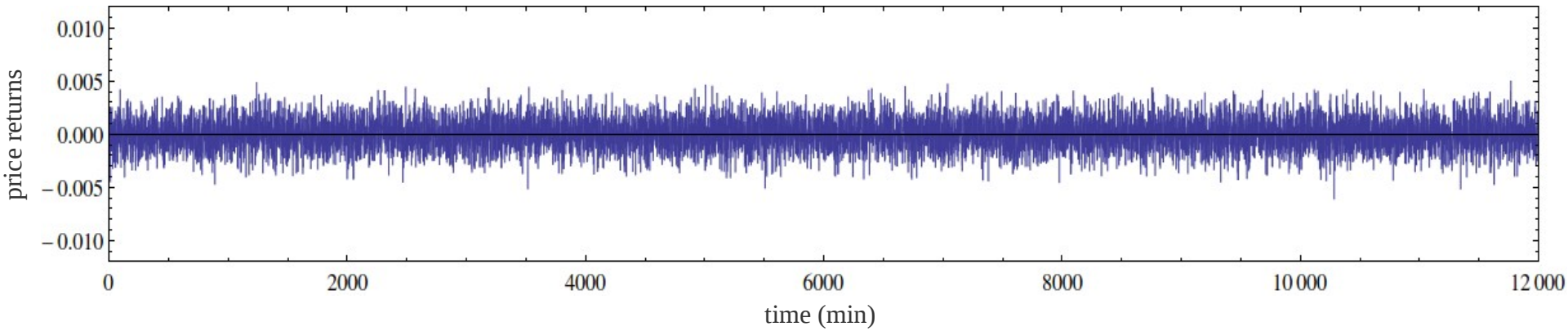
Real Returns (Financial Time-Series)



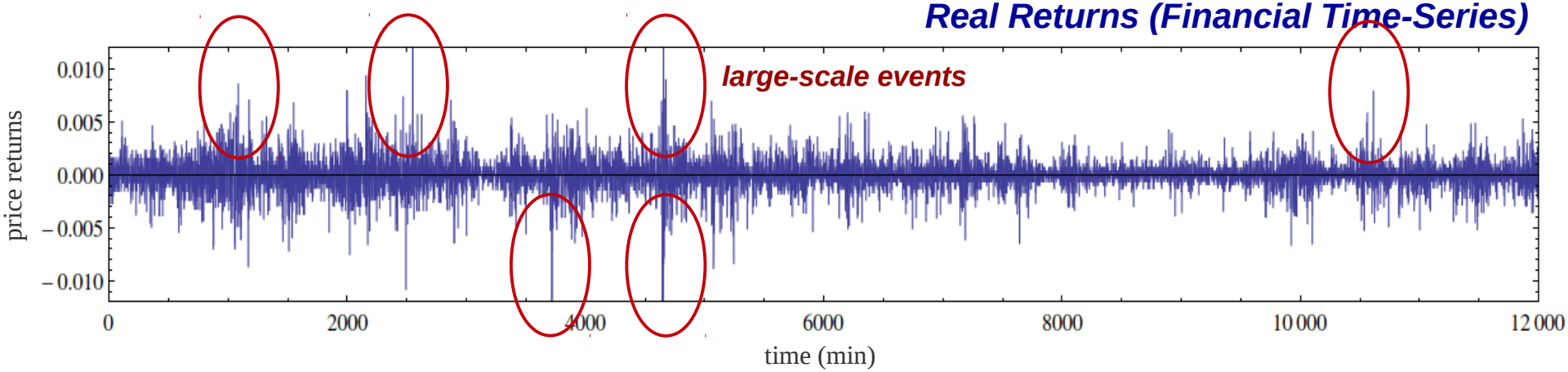
Cont, *Empirical properties of asset returns, stylized facts and statistical issues*, 2001

Statistical Properties of Price Returns

Simulated Returns (Geometric Brownian Motion)



Real Returns (Financial Time-Series)



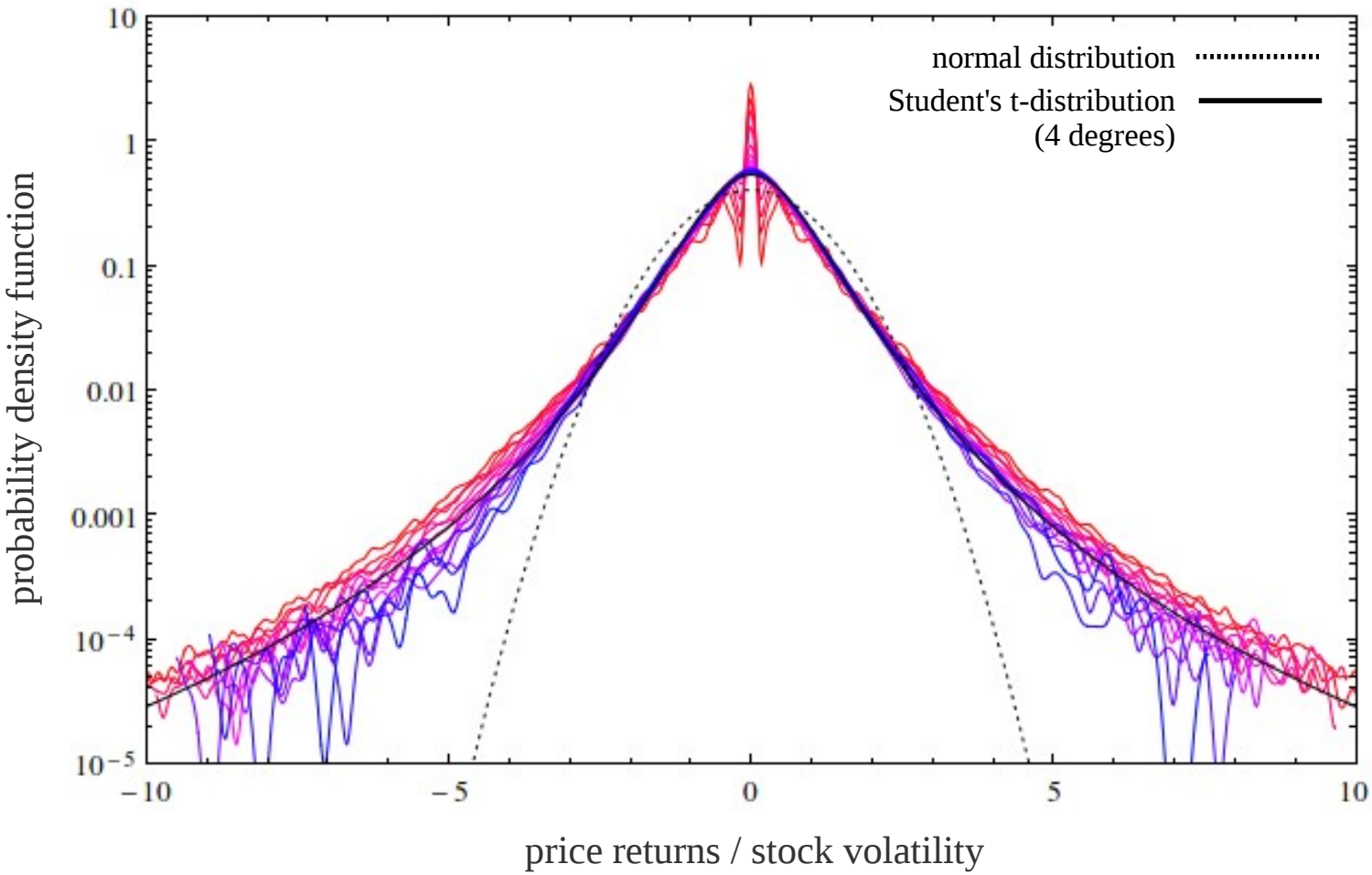
1st issue

Price returns are not normal

Cont, *Empirical properties of asset returns, stylized facts and statistical issues*, 2001

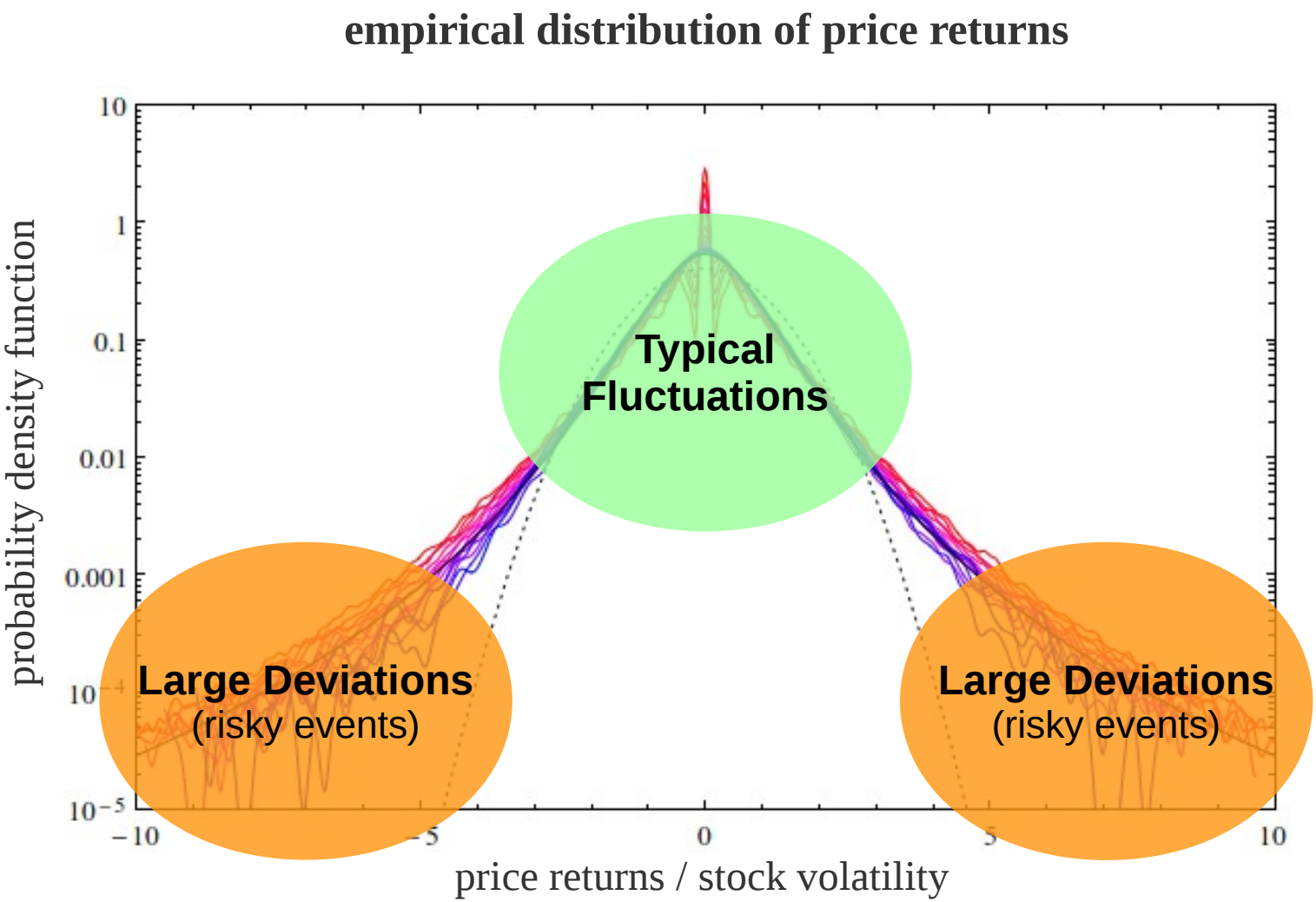
Empirical Distribution of Price Returns

Empirical Distribution of Returns (superposition of all stocks)
different time-scales – from **1 minute** to **2 hours**



Filiasi, *PhD Thesis*
Cont, Empirical properties of asset returns, stylized facts and statistical issues, 2001

Empirical Distribution of Price Returns



Filiasi, *PhD Thesis*

Cont, Empirical properties of asset returns, stylized facts and statistical issues, 2001