Statistical Data Analysis 2024/25
Lecture Week 3

London Postgraduate Lectures on Particle Physics

University of London MSc/MSci course PH4515

Glen Cowan
Physics Department
Royal Holloway, University of London
g.cowan@rhul.ac.uk
www.pp.rhul.ac.uk/~cowan

Course web page via RHUL moodle (PH4515) and also

www.pp.rhul.ac.uk/~cowan/stat_course.html

Some distributions

Distribution/pdf	Example use in Particle Physics
Binomial	Branching ratio
Multinomial	Histogram with fixed N
Poisson	Number of events found
Uniform	Monte Carlo method
Exponential	Decay time
Gaussian	Measurement error
Chi-square	Goodness-of-fit
Cauchy	Mass of resonance
Landau	Ionization energy loss
Beta	Prior pdf for efficiency
Gamma	Sum of exponential variables
Student's t	Resolution function with adjustable tails

Statistical Data Analysis
Lecture 3-1

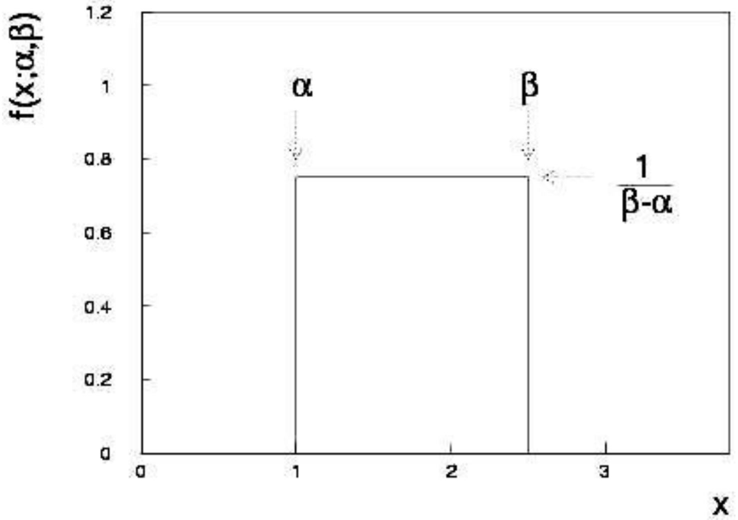
- Continuous probability density functions
  - Uniform
  - Exponential

Uniform distribution

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\beta-\alpha} & \alpha \le x \le \beta \\ 0 & \text{otherwise} \end{cases}$$

$$E[x] = \frac{1}{2}(\alpha+\beta)$$

$$V[x] = \frac{1}{12}(\beta-\alpha)^2$$



Notation: x follows a uniform distribution between α and β

write as: $x \sim \mathrm{U}[\alpha,\beta]$

Uniform distribution (2)

Very often used with $\alpha = 0$, $\beta = 1$ (e.g., Monte Carlo method).

For any r.v. x with pdf $f(x)$, cumulative distribution $F(x)$, the function $y = F(x)$ is uniform in $[0,1]$:

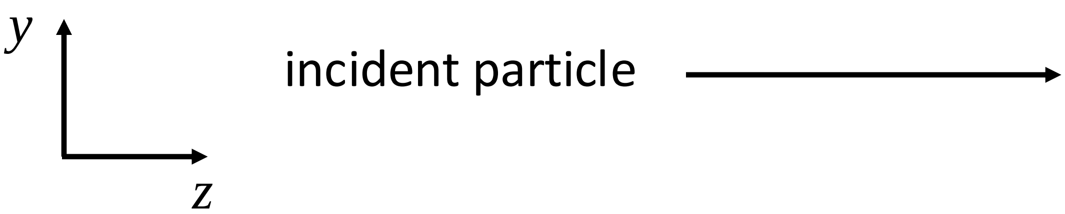
$$g(y) = f(x) \left| \frac{dx}{dy} \right| = \frac{f(x)}{|dy/dx|}$$

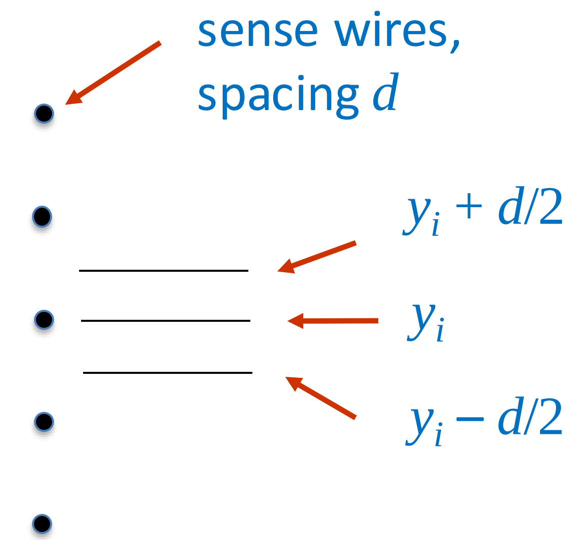
$$= \frac{f(x)}{|dF/dx|} = \frac{f(x)}{f(x)} = 1 \,, \quad 0 \le y \le 1$$

because $f(x) = dF/dx = dy/dx$

Uniform distribution: particle detector example

Vertical (y) position of particle's trajectory uniformly distributed over perpendicular plane of sense wires.





Sense wire closest to passage of particle gives signal.

If i-th wire gives signal,

estimated y position is y_i,

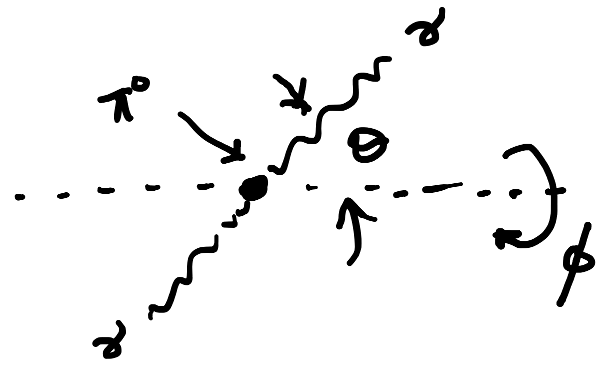
actual y position ~ $\mathrm{U}[y_i - d/2, y_i + d/2]$,

$V[y] = (y_i + d/2 - (y_i - d/2))^2 / 12 = d^2 / 12$,

position resolution = $\sigma_y = d/\sqrt{12}$

Uniform distribution: particle decay example

Decay $\pi^0 \to \gamma\gamma$ in π^0 rest frame:

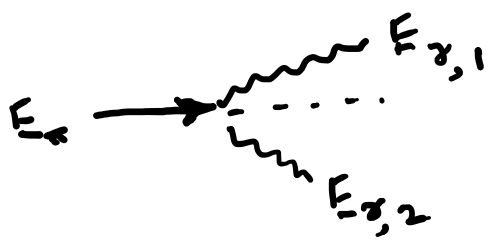


π^0 decay isotropic:

$$\cos\theta \sim U[-1, 1]$$

$$\phi \sim U[0, 2\pi]$$

In lab frame:



$$E_{\gamma,i} \sim U[E_{min}, E_{max}]$$

$$E_{min} = \frac{1}{2} E_\pi (1-\beta)$$

$$E_{max} = \frac{1}{2} E_\pi (1+\beta)$$

$$\beta = v_\pi / c$$

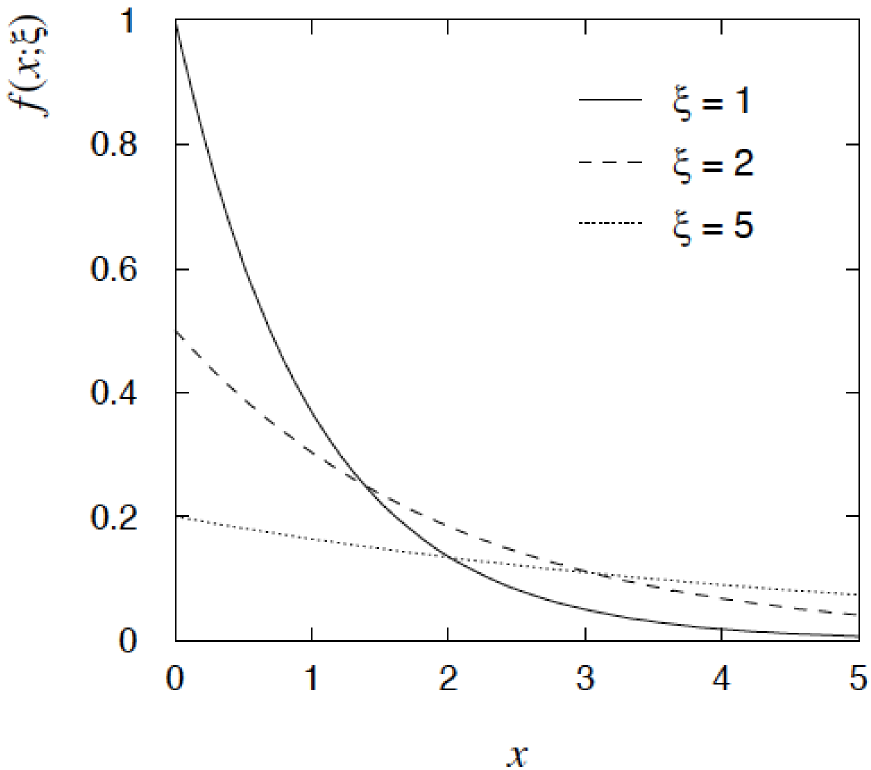
Exponential distribution

The exponential pdf for the continuous r.v. x is defined by:

$$f(x;\xi) = \begin{cases} \frac{1}{\xi} e^{-x/\xi} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$E[x] = \xi$$

$$V[x] = \xi^2$$



Exponential distribution (2)

Example: proper decay time t of an unstable particle

$$f(t;\tau) = \frac{1}{\tau} e^{-t/\tau}$$ (/ = mean lifetime)

Lack of memory (unique to exponential): $f(t - t_0 | t \geq t_0) = f(t)$

Question for discussion:

A cosmic ray muon is created 30 km high in the atmosphere, travels to sea level and is stopped in a block of scintillator, giving a start signal at t_0. At a time t it decays to an electron giving a stop signal. What is distribution of the difference between stop and start times, i.e., the pdf of $t - t_0$ given $t > t_0$?

Statistical Data Analysis
Lecture 3-2

- The Gaussian (normal) distribution
  - Univariate Gaussian
  - Standardized random variables
  - Location and scale parameters
  - Central Limit Theorem
  - Multivariate Gaussian

Gaussian (normal) distribution

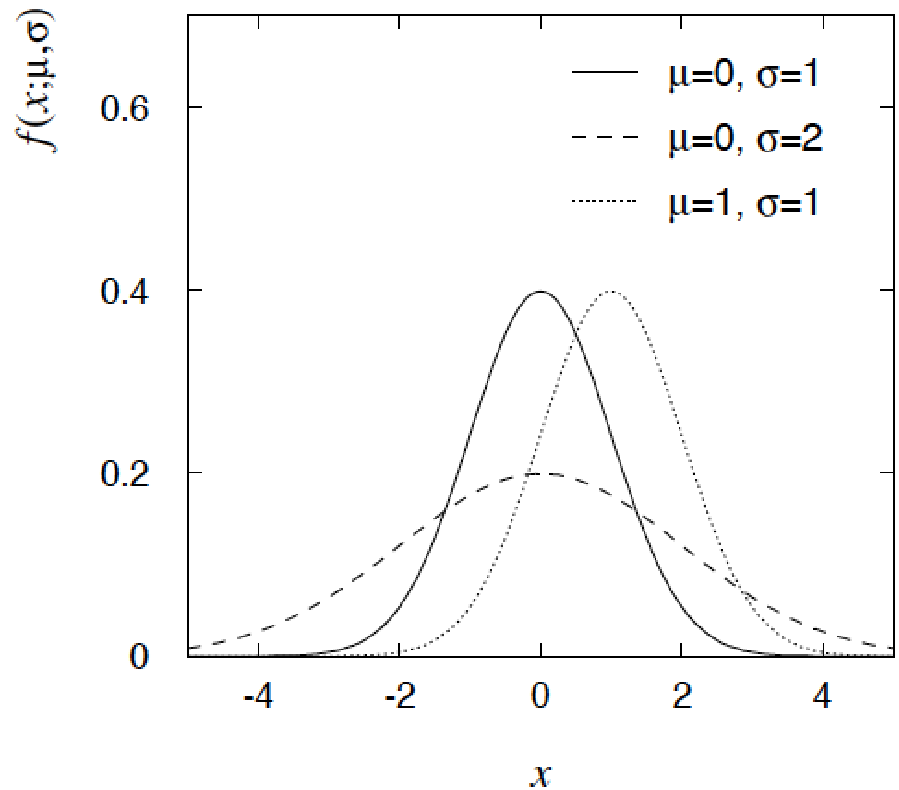
The Gaussian (normal) pdf for a continuous r.v. x is defined by:

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

$$E[x] = \mu$$

$$V[x] = \sigma^2$$

N.B. often μ, σ^2 denote mean, variance of any r.v., not only Gaussian.



Standardized random variables

If a random variable y has pdf $f(y)$ with mean μ and std. dev. σ, then the *standardized* variable

$$x = \frac{y-\mu}{\sigma} \quad \text{has the pdf} \quad g(x) = f(y(x)) \left| \frac{dy}{dx} \right| = \sigma f(\mu + \sigma x)$$

has mean of zero and standard deviation of 1.

Often work with the *standard* Gaussian distribution (μ = 0. σ = 1) using notation:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \,, \qquad \Phi(x) = \int_{-\infty}^{x} \varphi(x')\, dx'$$

Then e.g. $y = \mu + \sigma x$ follows

$$f(y) = \frac{1}{\sigma} \varphi\left(\frac{y-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-\mu)^2/2\sigma^2}$$

Digression: location/scale parameters

If a pdf $f(x; a)$ depending on a parameter a can be written as

$$f(x; a) = f(x - a; 0)$$

then a is called a location parameter. Adjusting a shifts the pdf to the right/left without changing its shape.

The parameter μ of the Gaussian is an example of a location parameter.

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

Digression: location/scale parameters (2)

If a pdf $f(x; b)$ depending on a parameter b can be written as

$$f(x;b) = \frac{1}{b} f(x/b;1)$$

then b is called a scale parameter. Adjusting b changes the ”units” of the random variable.

The parameter ξ of the exponential is an example of a scale parameter.

$$f(x;\xi) = \frac{1}{\xi} e^{-x/\xi}$$

Or if a pdf $f(x; a, b)$ has a location parameter a and can be written

$$f(x;a,b) = \frac{1}{b} f\left(\frac{x-a}{b};0,1\right)$$

then a and b are said to be location and scale parameters.

Example: μ and σ of Gaussian.

Gaussian pdf and the Central Limit Theorem

The Gaussian pdf is so useful because almost any random variable that is a sum of a large number of small contributions follows it. This follows from the Central Limit Theorem:

For n independent r.v.s x_i with finite variances σ_i^2, mean values μ_i, otherwise arbitrary pdfs, consider the sum

$$y = \sum_{i=1}^{n} x_i$$

In the limit $n \to \infty$, y is a Gaussian r.v. with

$$E[y] = \sum_{i=1}^{n} \mu_i \qquad V[y] = \sum_{i=1}^{n} \sigma_i^2$$

Measurement errors are often the sum of many contributions, so frequently measured values can be treated as Gaussian r.v.s.

Central Limit Theorem (2)

Versions of CLT differ in criteria for convergence and requirement (or not) of same pdf for all x_i.

See e.g. **`en.wikipedia.org/wiki/Central_limit_theorem`**

Classical CLT: all x_i independent and have same pdf with finite variance, can be proved using characteristic functions (Fourier transforms), see, e.g., SDA Chapter 10.

Physicist's CLT: for finite n, the sum $\Sigma_{i=1}^{n} x_i$ becomes approximately Gaussian to the extent that the fluctuation of the sum is not dominated by one (or few) terms.

Far enough in the tails the approximation generally breaks down.

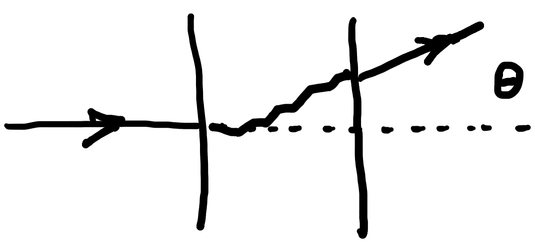
Central Limit Theorem (3)

Good example: velocity component of air molecule $v_x = \sum_i \delta v_{xi}$

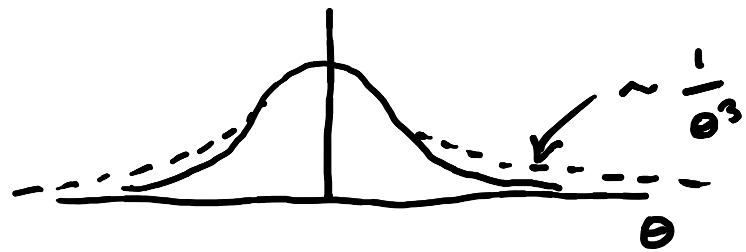
If v_x, v_y, v_z ~ Gaussian, then

$v = (v_x^{\,2} + v_y^{\,2} + v_z^{\,2})^{1/2}$ ~ Maxwell-Boltzmann

OK example: total deflection of charged particle from multiple Coulomb scattering. (Rare large-angle scatters $\rightarrow$ non-Gaussian tail.)



$\theta = \sum_i \delta\theta_i$



Bad example: energy loss of charged particle traversing thin gas layer. Rare collisions make up large fraction of energy loss, cf. Landau pdf.

Multivariate Gaussian distribution

Multivariate Gaussian pdf for the vector $\vec{x} = (x_1, \ldots, x_n)$:

$$f(\vec{x}; \vec{\mu}, V) = \frac{1}{(2\pi)^{n/2}|V|^{1/2}} \exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T V^{-1} (\vec{x}-\vec{\mu})\right]$$

$\vec{x}$, $\vec{\mu}$ are column vectors, $\vec{x}^T$, $\vec{\mu}^T$ are transpose (row) vectors,

$$E[x_i] = \mu_i, \, , \qquad \mathrm{cov}[x_i, x_j] = V_{ij} \; .$$

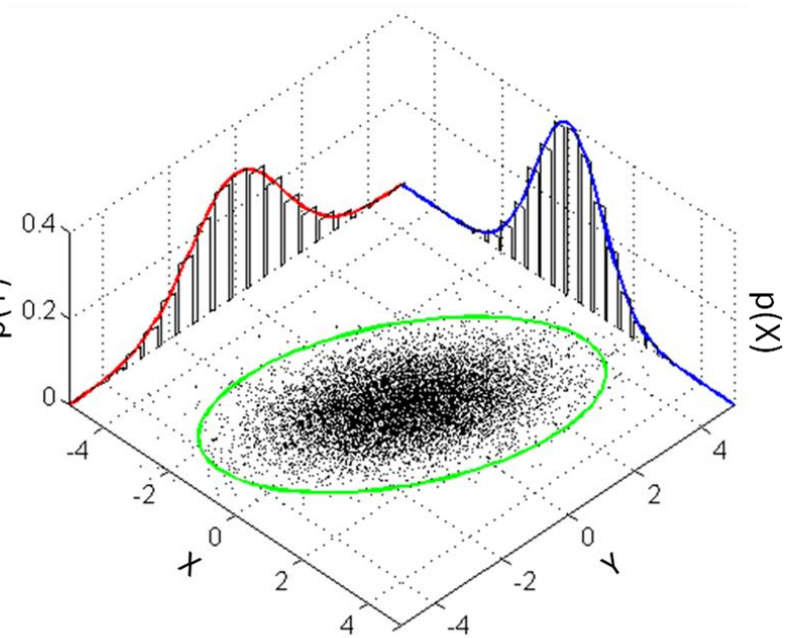
Marginal pdf of each x_i is Gaussian with mean μ_i, standard deviation $\sigma_i = \sqrt{V_{ii}}$.

Two-dimensional Gaussian distribution

$$f(x_1, x_2, ; \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

$$\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\}$$

where $\rho = \mathrm{cov}[x_1, x_2]/(\sigma_1 \sigma_2)$ is the correlation coefficient.



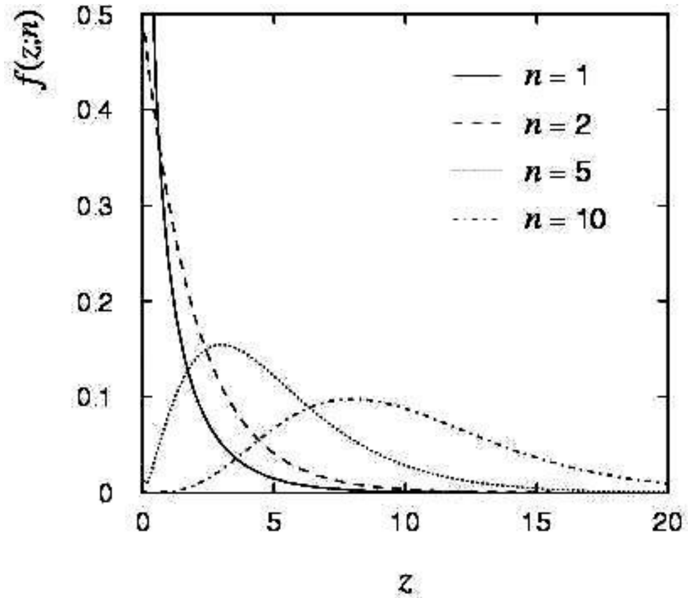
Statistical Data Analysis
Lecture 3-3

- More continuous probability density functions
  - Chi-square
  - Cauchy
  - Landau
  - Beta
  - Gamma
  - Student's t

Chi-square (χ^2) distribution

The chi-square pdf for the continuous r.v. z ($z \geq 0$) is defined by

$$f(z;n) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1} e^{-z/2}$$



n = 1, 2, ... = number of 'degrees of freedom' (dof)

$$E[z] = n\,, \qquad V[z] = 2n\,.$$

For independent Gaussian x_i, i = 1, ..., n, means μ_i, variances σ_i^2,

$$z = \sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{\sigma_i^2}$$

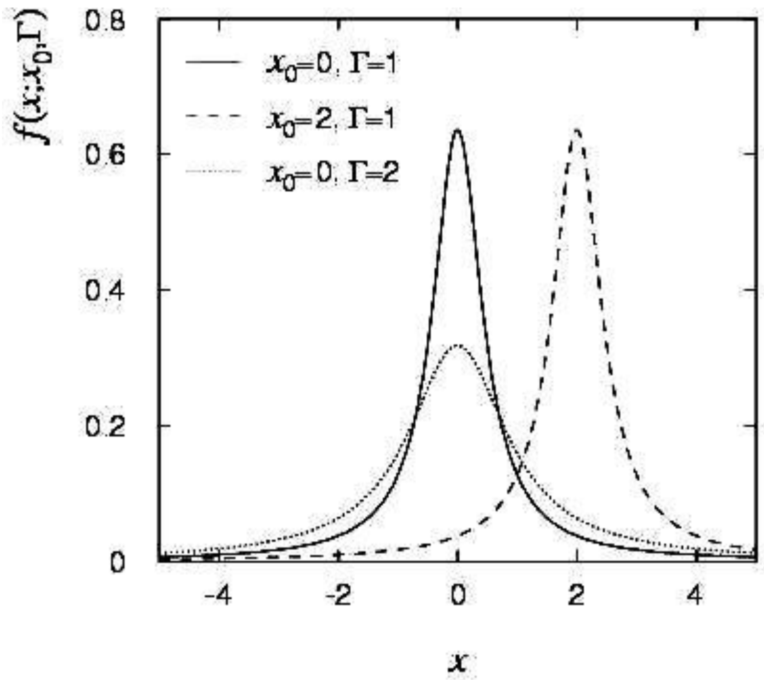
follows χ^2 pdf with n dof.

Example: goodness-of-fit test variable especially in conjunction with method of least squares.

Cauchy (Breit-Wigner) distribution

The Breit-Wigner pdf for the continuous r.v. x is defined by

$$f(x;\Gamma,x_0) = \frac{1}{\pi}\frac{\Gamma/2}{\Gamma^2/4+(x-x_0)^2}$$



($\Gamma = 2$, $x_0 = 0$ is the Cauchy pdf.)

$E[x]$ not well defined, $V[x] \to \infty$.

x_0 = mode (most probable value)

Γ = full width at half maximum

Example: mass of resonance particle, e.g. ρ, K^*, φ^0, ...

Γ = decay rate (inverse of mean lifetime)

Landau distribution

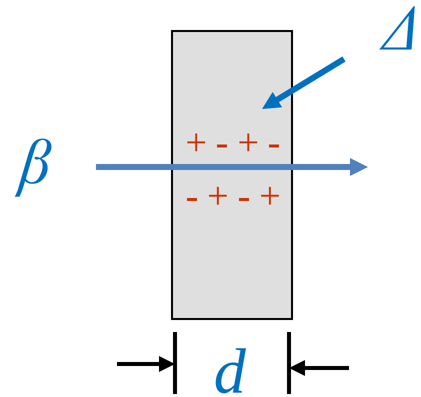
For a charged particle with $\beta = v/c$ traversing a layer of matter of thickness d, the energy loss Δ follows the Landau pdf:

$$f(\Delta;\beta) = \frac{1}{\xi}\phi(\lambda) \, ,$$

$$\phi(\lambda) = \frac{1}{\pi}\int_0^\infty \exp(-u\ln u - \lambda u)\sin \pi u \, du \, ,$$

$$\lambda = \frac{1}{\xi}\left[\Delta - \xi\left(\ln\frac{\xi}{\epsilon'} + 1 - \gamma_{\mathrm{E}}\right)\right] \, ,$$

$$\xi = \frac{2\pi N_{\mathrm{A}} e^4 z^2 \rho \sum Z}{m_{\mathrm{e}} c^2 \sum A}\frac{d}{\beta^2} \, , \qquad \epsilon' = \frac{I^2 \exp\beta^2}{2 m_{\mathrm{e}} c^2 \beta^2 \gamma^2} \, .$$



L. Landau, J. Phys. USSR **8** (1944) 201; see also
W. Allison and J. Cobb, Ann. Rev. Nucl. Part. Sci. **30** (1980) 253.

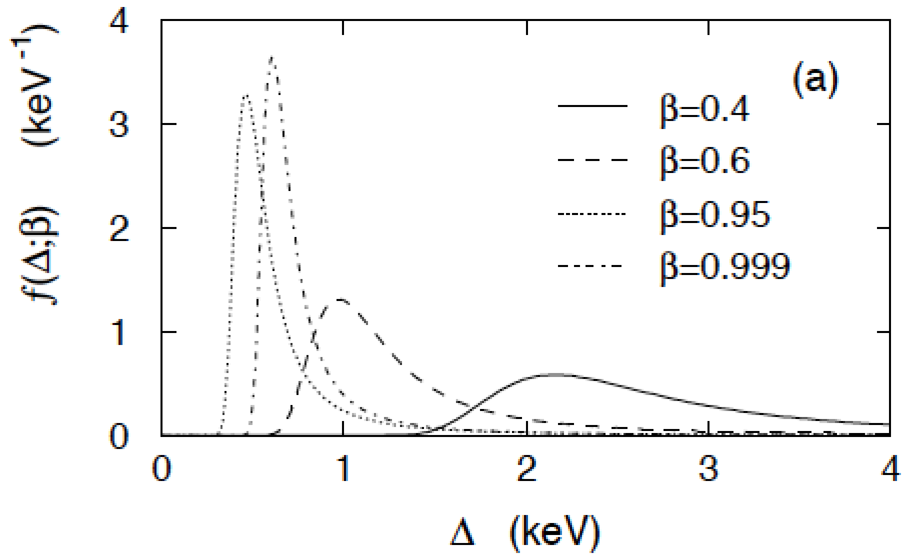
Landau distribution (2)

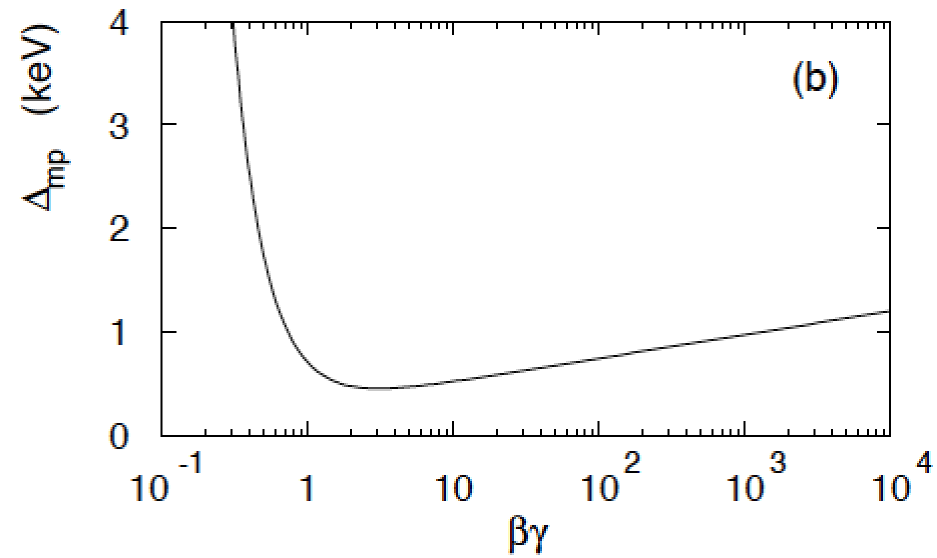
Long 'Landau tail'

→ all moments ∞

Mode (most probable value) sensitive to β ,

→ particle i.d.



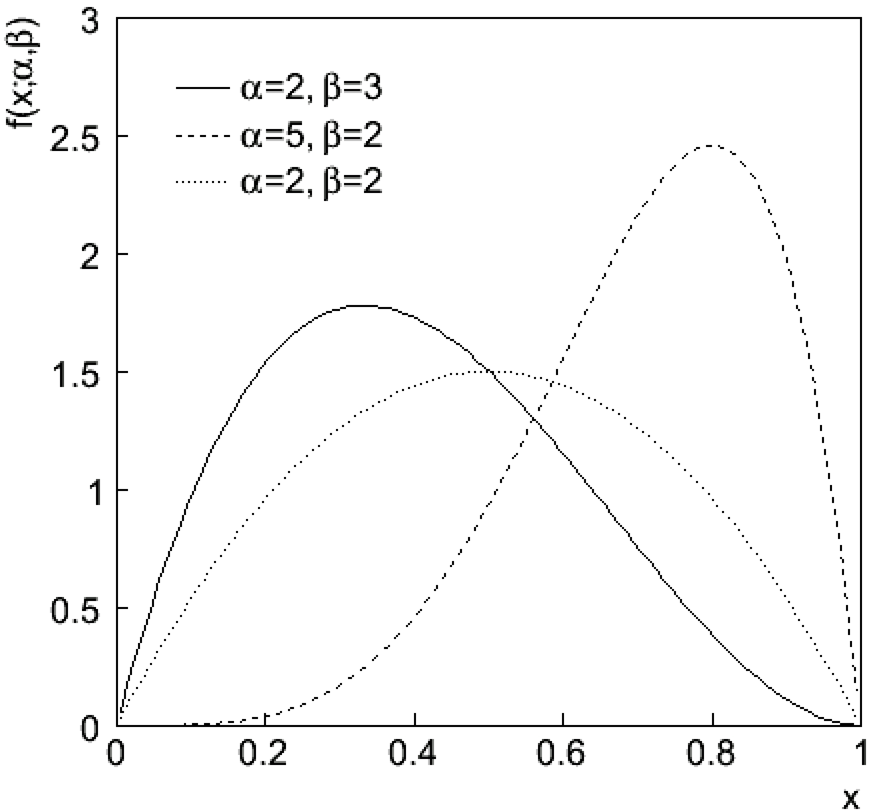


Beta distribution

$$f(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \;, \qquad 0 \le x \le 1$$

$$E[x] = \frac{\alpha}{\alpha+\beta}$$

$$V[x] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$



Often used to represent pdf of continuous r.v. nonzero only between finite limits, e.g.,

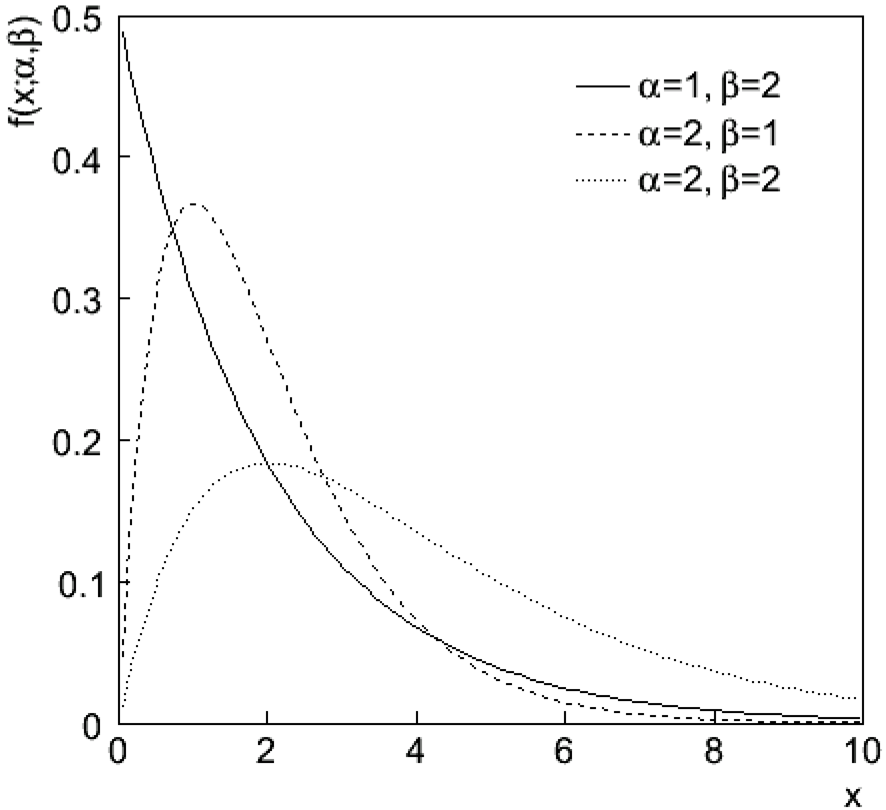
$y = a_0 + a_1 x, \quad a_0 \le y \le a_0 + a_1$

Gamma distribution

$$f(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}}\, x^{\alpha-1}\, e^{-x/\beta}\,, \qquad x \geq 0$$

$$E[x] = \alpha\beta$$

$$V[x] = \alpha\beta^2$$



Often used to represent pdf of continuous r.v. nonzero only in $[0,\infty]$.

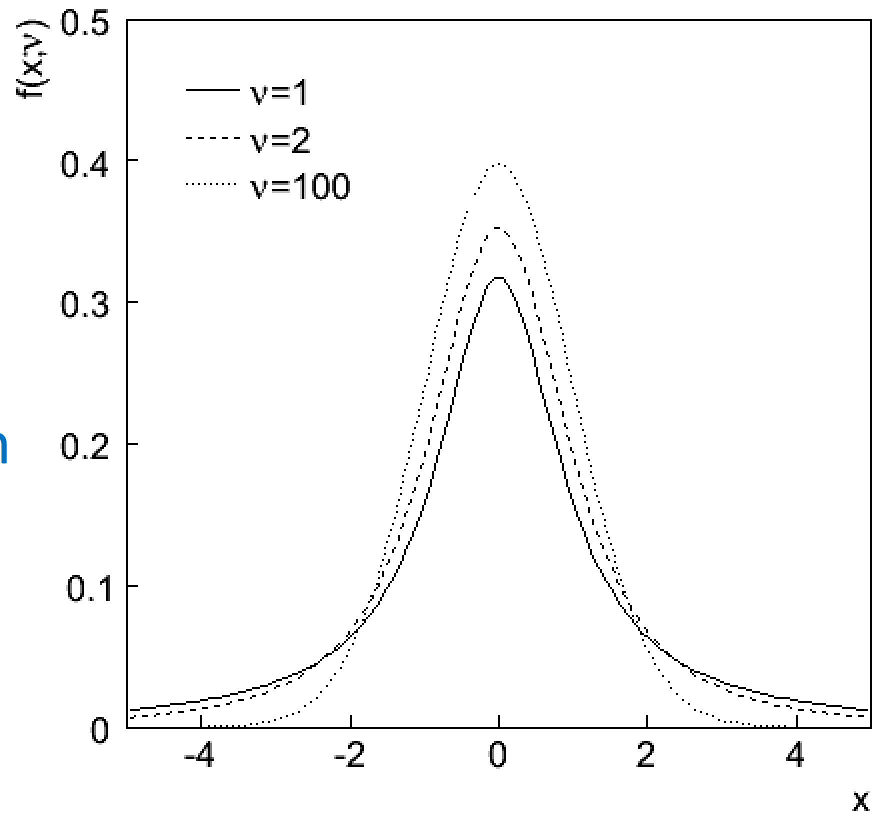
Also e.g. sum of n exponential r.v.s or time until nth event in Poisson process ~ Gamma

Student's t distribution

$$f(x;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1+\frac{x^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

$$E[x] = 0 \quad (\nu > 1)$$

$$V[x] = \frac{\nu}{\nu-2} \quad (\nu > 2)$$



ν = number of degrees of freedom (not necessarily integer)

$\nu = 1$ gives Cauchy,

$\nu \rightarrow \infty$ gives Gaussian.

Student's *t* distribution (2)

If x ~ Gaussian with $\mu = 0$, $\sigma^2 = 1$, and

$z \sim \chi^2$ with n degrees of freedom, then

$t = x \,/\, (z/n)^{1/2}$ follows Student's t with $\nu = n$.

This arises in problems where one forms the ratio of a sample mean to the sample standard deviation of Gaussian r.v.s.

The Student's t provides a bell-shaped pdf with adjustable tails, ranging from those of a Gaussian, which fall off very quickly, ($\nu \to \infty$, but in fact already very Gauss-like for $\nu =$ two dozen), to the very long-tailed Cauchy ($\nu = 1$).

Developed in 1908 by William Gosset, who worked for the Guinness Brewery and published under the pseudonym "Student".

VOLUME VI MARCH, 1908 No. 1

BIOMETRIKA.

THE PROBABLE ERROR OF A MEAN.

BY STUDENT.

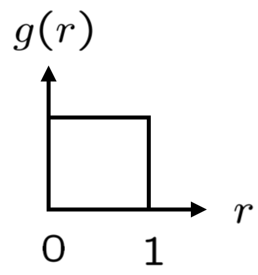
Statistical Data Analysis
Lecture 3-4

- The Monte Carlo method
  - basic ingredients
  - random number generators
  - transformation method
  - acceptance-rejection method
  - example uses

The Monte Carlo method

What it is: a numerical technique for calculating probabilities and related quantities using sequences of random numbers.

The usual steps:



(1) Generate sequence $r_1, r_2, \ldots, r_m$ independent and uniform on $[0, 1]$.

(2) Use this to produce another sequence $x_1, x_2, \ldots, x_n$ independent and distributed according to some pdf $f(x)$ in which we're interested (x can be a vector).

(3) Use the x values to estimate some property of $f(x)$, e.g., fraction of x values with $a < x < b$ gives $\int_a^b f(x)\,dx$.

$\rightarrow$ MC calculation = integration (at least formally)

MC generated values = 'simulated data'

$\rightarrow$ use for testing statistical procedures

Random number generators

Goal: generate uniformly distributed values in $[0, 1]$.

Toss coin for e.g. 32 bit number... (too tiring).

$\rightarrow$ 'random number generator'

= computer algorithm to generate $r_1, r_2, \ldots, r_n$.

Example: multiplicative linear congruential generator (MLCG)

$n_{i+1} = (a\, n_i) \bmod m$, where

n_i = integer

a = multiplier

m = modulus

n_0 = seed (initial value)

N.B. mod = modulus (remainder), e.g. $27 \bmod 5 = 2$.

This rule produces a sequence of numbers $n_0, n_1, \ldots$

Random number generators (2)

The sequence is (unfortunately) periodic!

Example (see Brandt Ch 4): $a = 3$, $m = 7$, $n_0 = 1$

$$
\begin{aligned}
n_1 &= (3 \cdot 1)\mathrm{mod}7 = 3 \\
n_2 &= (3 \cdot 3)\mathrm{mod}7 = 2 \\
n_3 &= (3 \cdot 2)\mathrm{mod}7 = 6 \\
n_4 &= (3 \cdot 6)\mathrm{mod}7 = 4 \\
n_5 &= (3 \cdot 4)\mathrm{mod}7 = 5 \\
n_6 &= (3 \cdot 5)\mathrm{mod}7 = 1 \qquad \leftarrow \text{ sequence repeats}
\end{aligned}
$$

Choose a, m to obtain long period (maximum = $m - 1$); m usually close to the largest integer that can represented in the computer.

Only use a subset of a single period of the sequence.

Random number generators (3)

$r_i = n_i/n_{\max}$ are in $[0, 1]$ but are they independent and uniform?

Choose a, m so that the r_i pass various tests of randomness:

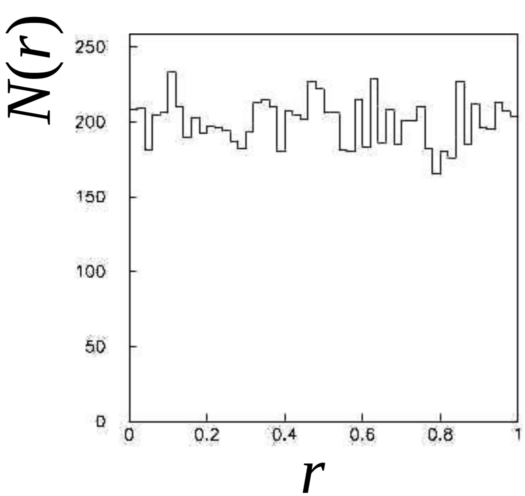
uniform distribution in $[0, 1]$,

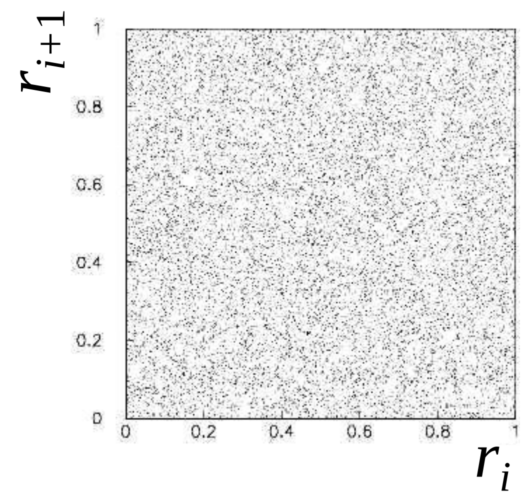
all values independent (no correlations between pairs),

e.g. L'Ecuyer, Commun. ACM **31** (1988) 742 suggests

$a = 40692$

$m = 2147483399$



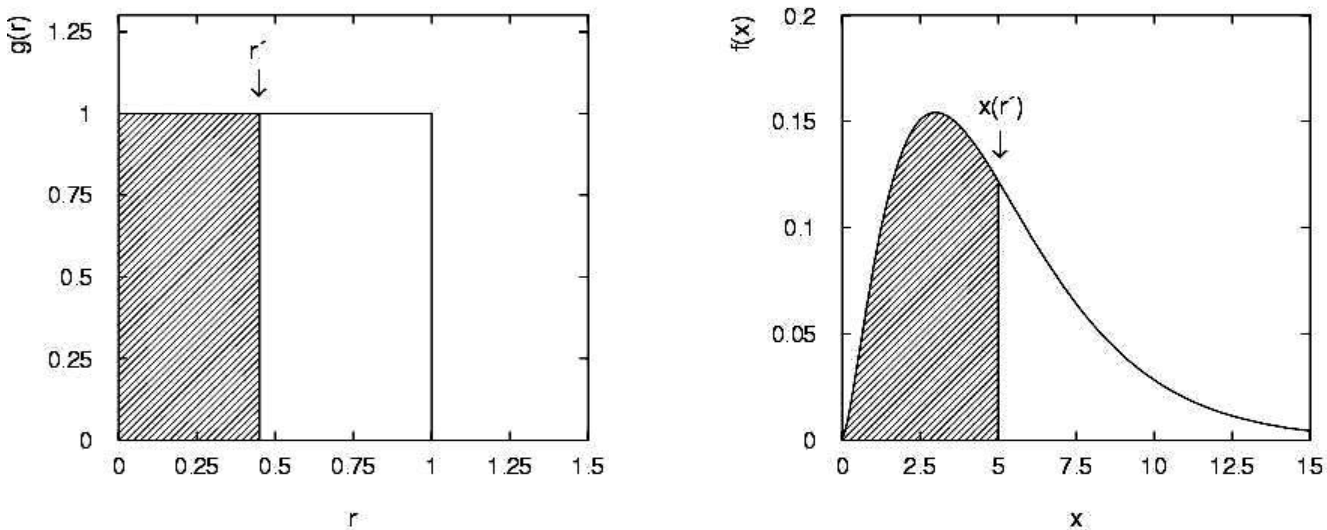


Far better generators available, e.g. `TRandom3`, based on Mersenne twister algorithm, period $= 2^{19937} - 1$ (a "Mersenne prime").

See F. James, Comp. Phys. Comm. 60 (1990) 111; Brandt Ch. 4.

The transformation method

Given $r_1, r_2, \ldots, r_n$ uniform in $[0, 1]$, find $x_1, x_2, \ldots, x_n$ that follow $f(x)$ by finding a suitable transformation $x(r)$.



Require: $P(r \leq r') = P(x \leq x(r'))$

i.e. $$\int_{-\infty}^{r'} g(r)\,dr = r' = \int_{-\infty}^{x(r')} f(x')\,dx' = F(x(r'))$$

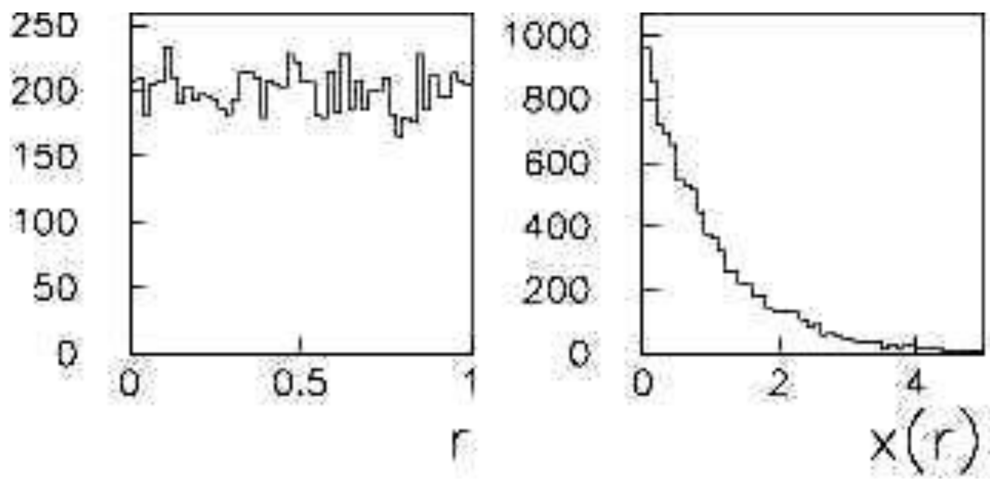
That is, $\boxed{\text{set } F(x) = r \text{ and solve for } x(r).}$

Example of the transformation method

Exponential pdf: $$f(x;\xi) = \frac{1}{\xi} e^{-x/\xi} \quad (x \geq 0)$$

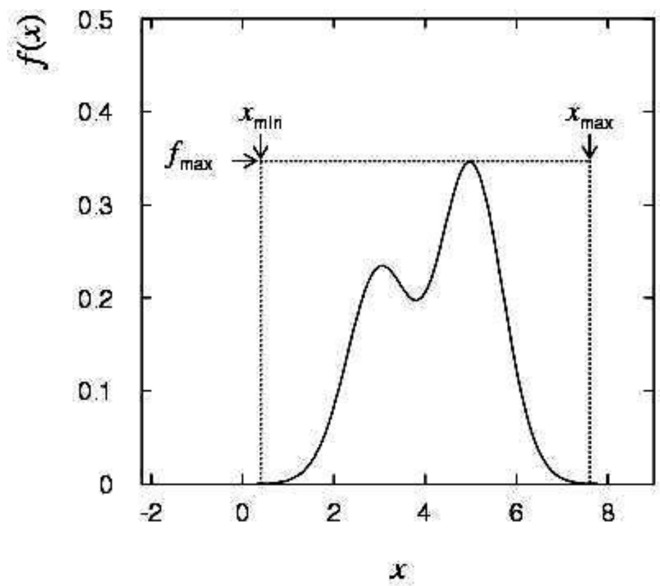
Set $\int_0^x \frac{1}{\xi} e^{-x'/\xi}\, dx' = r$ and solve for $x(r)$.

$\rightarrow \quad x(r) = -\xi \ln(1-r) \qquad (\, x(r) = -\xi \ln r$ works too.)



The acceptance-rejection method

Enclose the pdf in a box:



(1) Generate a random number x, uniform in $[x_{min}, x_{max}]$, i.e.

$$x = x_{\text{min}} + r_1(x_{\text{max}} - x_{\text{min}}) \ , \quad r_1 \text{ is uniform in } [0,1].$$

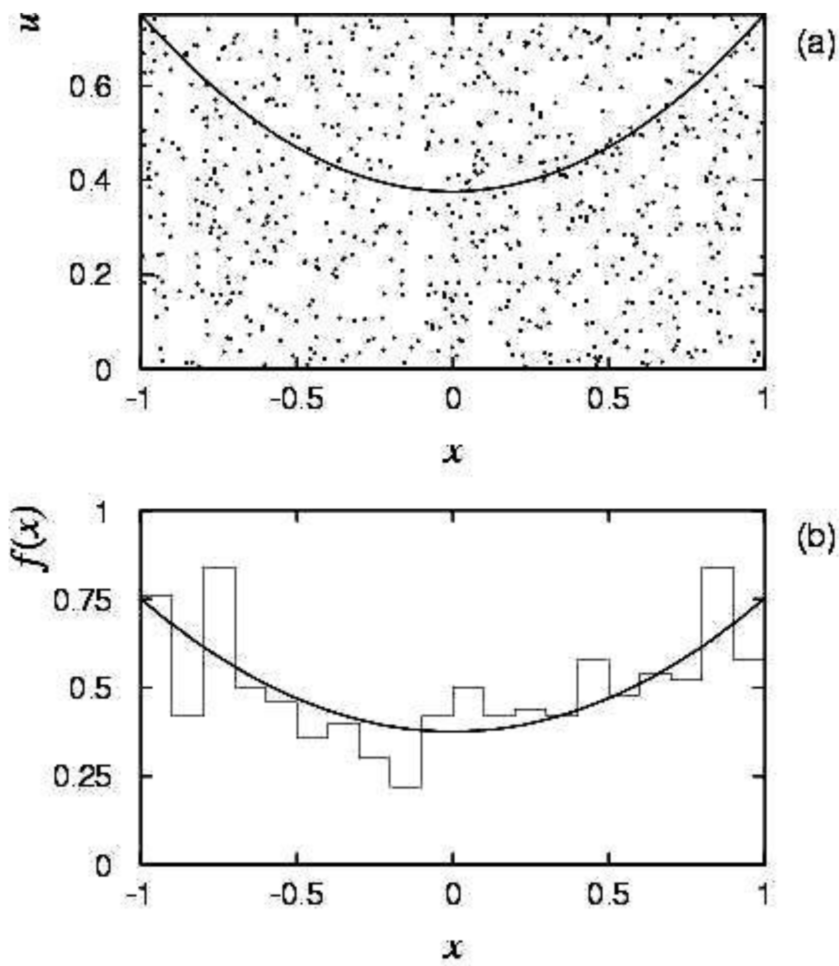
(2) Generate a 2nd independent random number u uniformly distributed between 0 and f_{max}, i.e. $u = r_2 f_{\text{max}}$.

(3) If $u < f(x)$, then accept x. If not, reject x and repeat.

Example with acceptance-rejection method

$$f(x) = \frac{3}{8}(1 + x^2)$$

$$(-1 \leq x \leq 1)$$



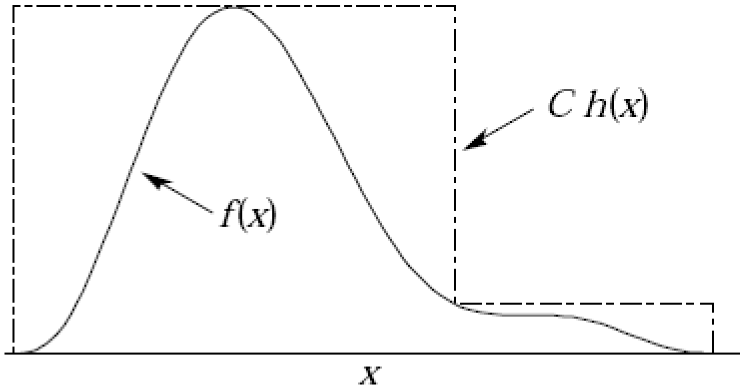
If dot below curve, use x value in histogram.

Improving efficiency of the acceptance-rejection method

The fraction of accepted points is equal to the fraction of the box's area under the curve.

For very peaked distributions, this may be very low and thus the algorithm may be slow.

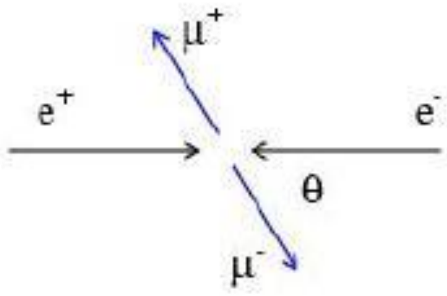
Improve by enclosing the pdf $f(x)$ in a curve $C\,h(x)$ that conforms to $f(x)$ more closely, where $h(x)$ is a pdf from which we can generate random values and C is a constant.



Generate points uniformly over $C\,h(x)$.

If point is below $f(x)$, accept x.

Monte Carlo event generators



Simple example: $e^+e^- \to \mu^+\mu^-$

Generate $\cos\theta$ and φ:

$$f(\cos\theta; A_{\rm FB}) \propto \left(1 + \frac{8}{3} A_{\rm FB} \cos\theta + \cos^2\theta\right) ,$$

$$g(\phi) = \frac{1}{2\pi} \quad (0 \leq \phi \leq 2\pi)$$

Less simple: ‘event generators’ for a variety of reactions:

$e^+e^- \to \mu^+ \mu^-$, hadrons, ...

$pp \to$ hadrons, D-Y, SUSY,...

e.g. PYTHIA, HERWIG, ISAJET...

Output = ‘events’, i.e., for each event we get a list of generated particles and their momentum vectors, types, etc.

A simulated event

Event listing (summary)

I	particle/jet	KS	KF	orig	p_x	p_y	p_z	E	m
1	!p+!	21	2212	0	0.000	0.000	7000.000	7000.000	0.938
2	!p+!	21	2212	0	0.000	0.000	-7000.000	7000.000	0.938
3	!g!	21	21	1	0.863	-0.323	1739.862	1739.862	0.000
4	!ubar!	21	-2	2	-0.621	-0.163	-777.415	777.415	0.000
5	!g!	21	21	3	-2.427	5.486	1487.857	1487.	
6	!g!	21	21	4	-62.910	63.357	-463.274	471.	
7	!~g!	21	1000021	0	314.363	544.843	498.897	979.	
8	!~g!	21	1000021	0	-379.700	-476.000	525.686	980.	
9	!~chi_1-!	21	-1000024	7	130.058	112.247	129.860	263.	
10	!sbar!	21	-3	7	259.400	187.468	83.100	330.	
11	!c!	21	4	7	-79.403	242.409	283.026	381.	
12	!~chi_20!	21	1000023	8	-326.241	-80.971	113.712	385.	
13	!b!	21	5	8	-51.841	-294.077	389.853	491.	
14	!bbar!	21	-5	8	-0.597	-99.577	21.299	101.	
15	!~chi_10!	21	1000022	9	103.352	81.316	83.457	175.	
16	!s!	21	3	9	5.451	38.374	52.302	65.	
17	!cbar!	21	-4	9	20.839	-7.250	-5.938	22.	
18	!~chi_10!	21	1000022	12	-136.266	-72.961	53.246	181.	
19	!nu_mu!	21	14	12	-78.263	-24.757	21.719	84.	
20	!nu_mubar!	21	-14	12	-107.801	16.901	38.226	115.	
21	gamma	1	22	4	2.636	1.357	0.125	2.	
22	(~chi_1-)	11	-1000024	9	129.643	112.440	129.820	262.	
23	(~chi_20)	11	1000023	12	-322.330	-80.817	113.191	382.	
24	~chi_10	1	1000022	15	97.944	77.819	80.917	169.	
25	~chi_10	1	1000022	18	-136.266	-72.961	53.246	181.	
26	nu_mu	1	14	19	-78.263	-24.757	21.719	84.	
27	nu_mubar	1	-14	20	-107.801	16.901	38.226	115.	
28	(Delta++)	11	2224	2	0.222	0.012	-2734.287	2734.	

⋮

397	pi+	1	211	209	0.006	0.398	-308.296	308.297	0.140
398	gamma	1	22	211	0.407	0.087	-1695.458	1695.458	0.000
399	gamma	1	22	211	0.113	-0.029	-314.822	314.822	0.000
400	(pi0)	11	111	212	0.021	0.122	-103.709	103.709	0.135
401	(pi0)	11	111	212	0.084	-0.068	-94.276	94.276	0.135
402	(pi0)	11	111	212	0.267	-0.052	-144.673	144.674	0.135
403	gamma	1	22	215	-1.581	2.473	3.306	4.421	0.000
404	gamma	1	22	215	-1.494	2.143	3.051	4.016	0.000
405	pi-	1	-211	216	0.007	0.738	4.015	4.085	0.140
406	pi+	1	211	216	-0.024	0.293	0.486	0.585	0.140
407	K+	1	321	218	4.382	-1.412	-1.799	4.968	0.494
408	pi-	1	-211	218	1.183	-0.894	-0.176	1.500	0.140
409	(pi0)	11	111	218	0.955	-0.459	-0.590	1.221	0.135
410	(pi0)	11	111	218	2.349	-1.105	-1.181	2.855	0.135
411	(Kbar0)	11	-311	219	1.441	-0.247	-0.472	1.615	0.498
412	pi-	1	-211	219	2.232	-0.400	-0.249	2.285	0.140
413	K+	1	321	220	1.380	-0.652	-0.361	1.644	0.494
414	(pi0)	11	111	220	1.078	-0.265	0.175	1.132	0.135
415	(K_S0)	11	310	222	1.841	0.111	0.894	2.109	0.498
416	K+	1	321	223	0.307	0.107	0.252	0.642	0.494
417	pi-	1	-211	223	0.266	0.316	-0.201	0.480	0.140
418	nbar0	1	-2112	226	1.335	1.641	2.078	3.111	0.940
419	(pi0)	11	111	226	0.899	1.046	1.311	1.908	0.135
420	pi+	1	211	227	0.217	1.407	1.356	1.971	0.140
421	(pi0)	11	111	227	1.207	2.336	2.767	3.820	0.135
422	n0	1	2112	228	3.475	5.324	5.702	8.592	0.940
423	pi-	1	-211	228	1.856	2.606	2.808	4.259	0.140
424	gamma	1	22	229	-0.012	0.247	0.421	0.489	0.000
425	gamma	1	22	229	0.025	0.034	0.009	0.043	0.000
426	pi+	1	211	230	2.718	5.229	6.403	8.703	0.140
427	(pi0)	11	111	230	4.109	6.747	7.597	10.961	0.135
428	pi-	1	-211	231	0.551	1.233	1.945	2.372	0.140
429	(pi0)	11	111	231	0.645	1.141	0.922	1.608	0.135
430	gamma	1	22	232	-0.383	1.169	1.208	1.724	0.000
431	gamma	1	22	232	-0.201	0.070	0.060	0.221	0.000

⋮

PYTHIA Monte Carlo

pp → gluino-gluino

Monte Carlo detector simulation

Takes as input the particle list and momenta from generator.

Simulates detector response:

multiple Coulomb scattering (generate scattering angle),
particle decays (generate lifetime),
ionization energy loss (generate Δ),
electromagnetic, hadronic showers,
production of signals, electronics response, ...

Output = simulated raw data $\rightarrow$ input to reconstruction software:
track finding, fitting, etc.

Predict what you should see at 'detector level' given a certain hypothesis for 'generator level'. Compare with the real data.

Programming package: **GEANT**

Extra slides

Importance Sampling

Often the goal of a Monte Carlo calculation is to determine an expectation value of a function $h(x)$, where x is a random variable that follows a pdf $f(x)$,

$$E_f[h(x)] = \int h(x) f(x)\, dx \equiv \mu \ . \tag{1}$$

A Monte-Carlo estimator $\hat{\mu}_{\rm MC}$ for μ is the average of N values of $h(x)$ where x is sampled (generated) from $f(x)$:

$$\hat{\mu}_{\rm MC} = \frac{1}{N} \sum_{i=1}^{N} h(x_i) \ , \tag{2}$$

which has a variance

$$V[\hat{\mu}_{\rm MC}] = \frac{1}{N} V_f[h(x)] = \frac{1}{N} \left(E_f[h^2(x)] - \mu^2 \right) \ . \tag{3}$$

Importance Sampling (2)

By using the method of *importance sampling*, one can achieve a reduction in this variance and thus a more accurate determination of the expectation value for a given number of random values generated. The key idea is to rewrite the expectation value in Eq. (1) as

$$\mu = \int h(x) f(x)\, dx = \int \frac{h(x) f(x)}{g(x)} g(x)\, dx = E_g \left[\frac{h(x) f(x)}{g(x)} \right] , \tag{4}$$

where $g(x)$ is any other pdf of x with the same support as $f(x)$ (i.e., nonzero for the same region of x). Thus the desired quantity μ is the expectation value with respect to g of $h(x)f(x)/g(x)$. It can be estimated by generating N values of x sampled from g and computing

$$\hat{\mu}_{\rm IS} = \frac{1}{N} \sum_{i=1}^{N} \frac{h(x_i) f(x_i)}{g(x_i)} . \tag{5}$$

Importance Sampling (3)

The variance of $\hat{\mu}_{\rm IS}$ is given by

$$V[\hat{\mu}_{\rm IS}] = \frac{1}{N} V_g \left[\frac{h(x)f(x)}{g(x)} \right] = \frac{1}{N} \left(E_g \left[\frac{h^2(x)f^2(x)}{g^2(x)} \right] - \mu^2 \right) , \qquad (6)$$

By choosing $g(x)$ such that $h(x)f(x)/g(x)$ is as constant as possible, one can minimize the variance of $\hat{\mu}_{\rm IS}$. One can show (see, e.g., Refs. [1, 2]) that the variance is minimized when $g(x) \propto |h(x)|f(x)$.

An alternative importance sampling estimator can be constructed by replacing the number of generated values N in Eq. (5) by the sum $\sum_{i=1}^N f(x_i)/g(x_i)$. This can given an even smaller variance at the price of a small bias. It can be of further advantage in problems where the pdf $f(x)$ is known only up to a normalization constant, which then cancels (see Refs. [1, 2]).

References

[1] C.P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed., (Springer, New York, 2004).

[2] J.S. Liu, *Monte Carlo Strategies in Scientific Computing*, (Springer, New York, 2001).