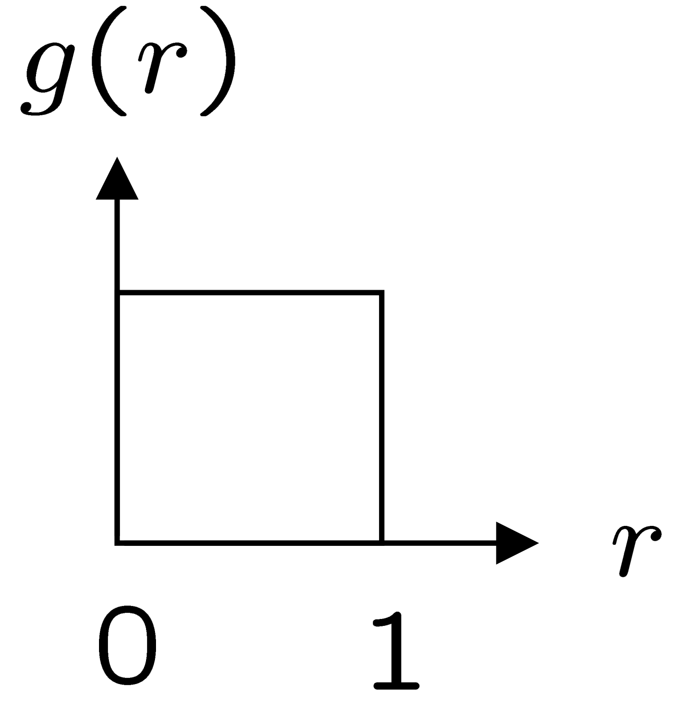
Monte Carlo Method

Hengne Li
South China Normal University

The Monte Carlo method

- What it is: a numerical technique for calculating probabilities and related quantities using sequences of random numbers.
- The usal steps:
  - (1) Generate sequence $r_1, r_2, \ldots, r_m$ uniform in [0, 1].
  - (2) Use this to produce another sequence $x_1, x_2, \ldots, x_n$ distributed according to some pdf *f(x)* in which we are interested (*x* can be a vector)



- MC generated values = “simulated data”, “simulation” ==> use for testing statistical procedures

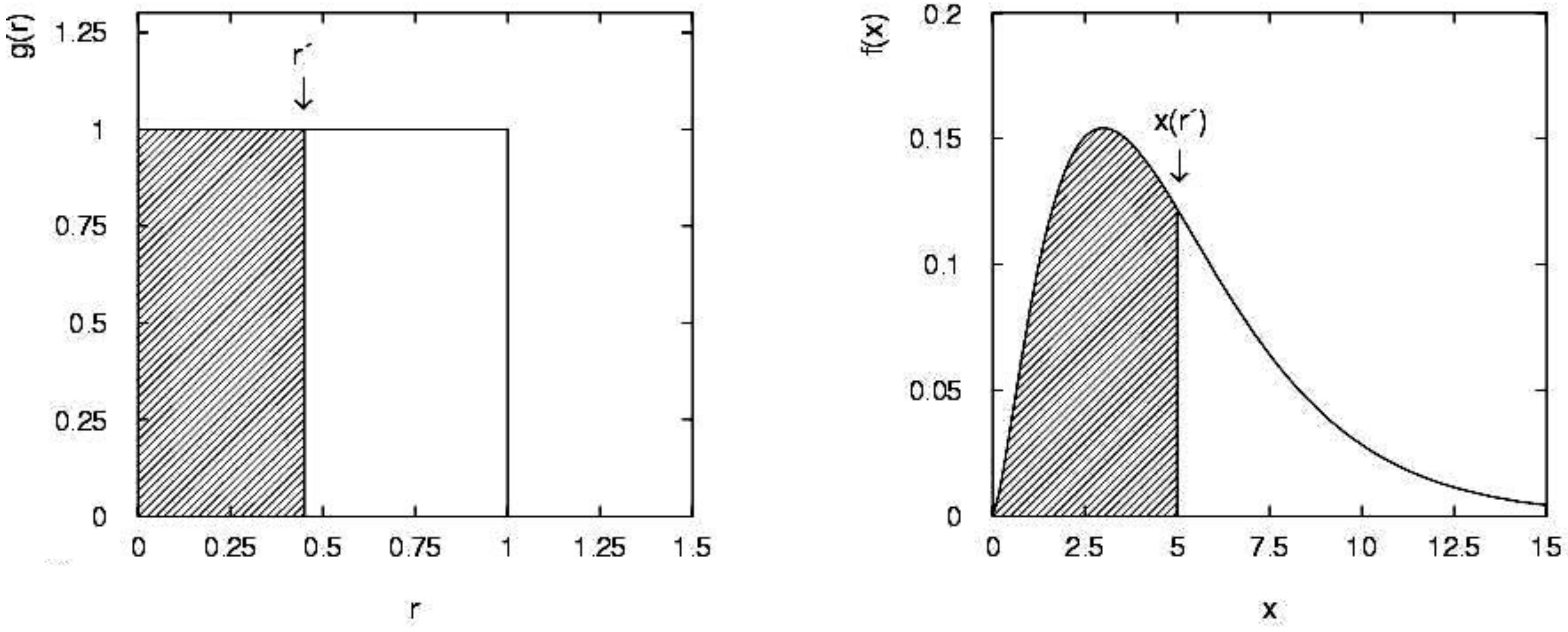
Random number generators

- Goal: generate uniformly distributed values in [0, 1].
- Tools/algorithms to generate random numbers are called “random number generator”
- There are many algorithms to do that, such as multiplicative linear congruential generator (MLCG), Mersenne twister algorithm, etc.,
- They are implemented, and we can use directly:
  - C++ standard library(stdlib): rand() function can generate random numbers.
  - ROOT: TRandom, TRandom2, TRandom3, where TRandom3 gives much better results [F. James, Comp. Phys. Comm. 60 (1990) 111]

The transformation method

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

- Remind: previously, we learned given pdf f(x) and function a(x), ask the pdf of a: g(a).



- Here, is given pdf f(x) and g(r), ask the function 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, set F(x) = r and solve for x(r)

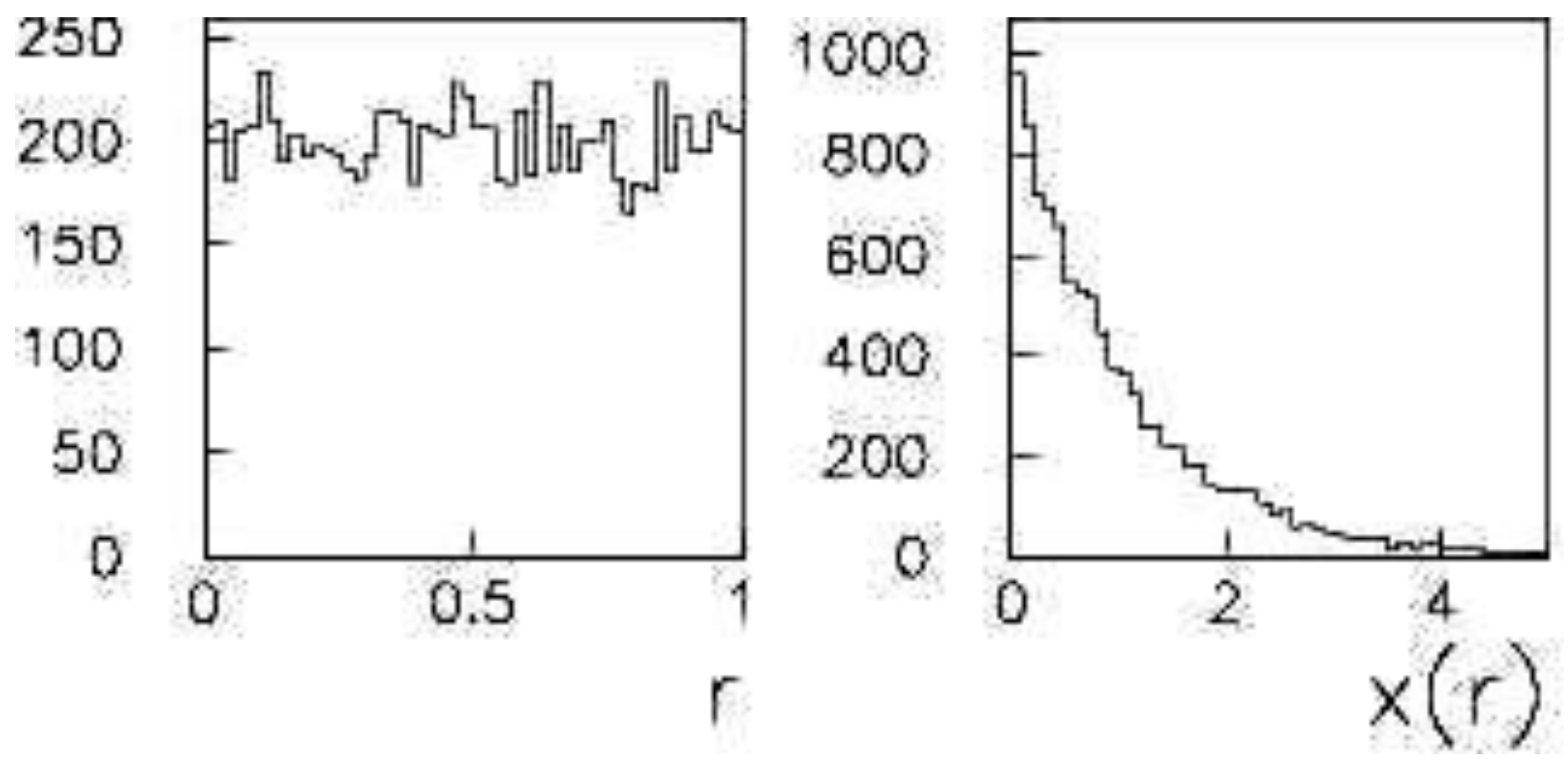
Example of the transformation method

- 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 $x\,(r)$.

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

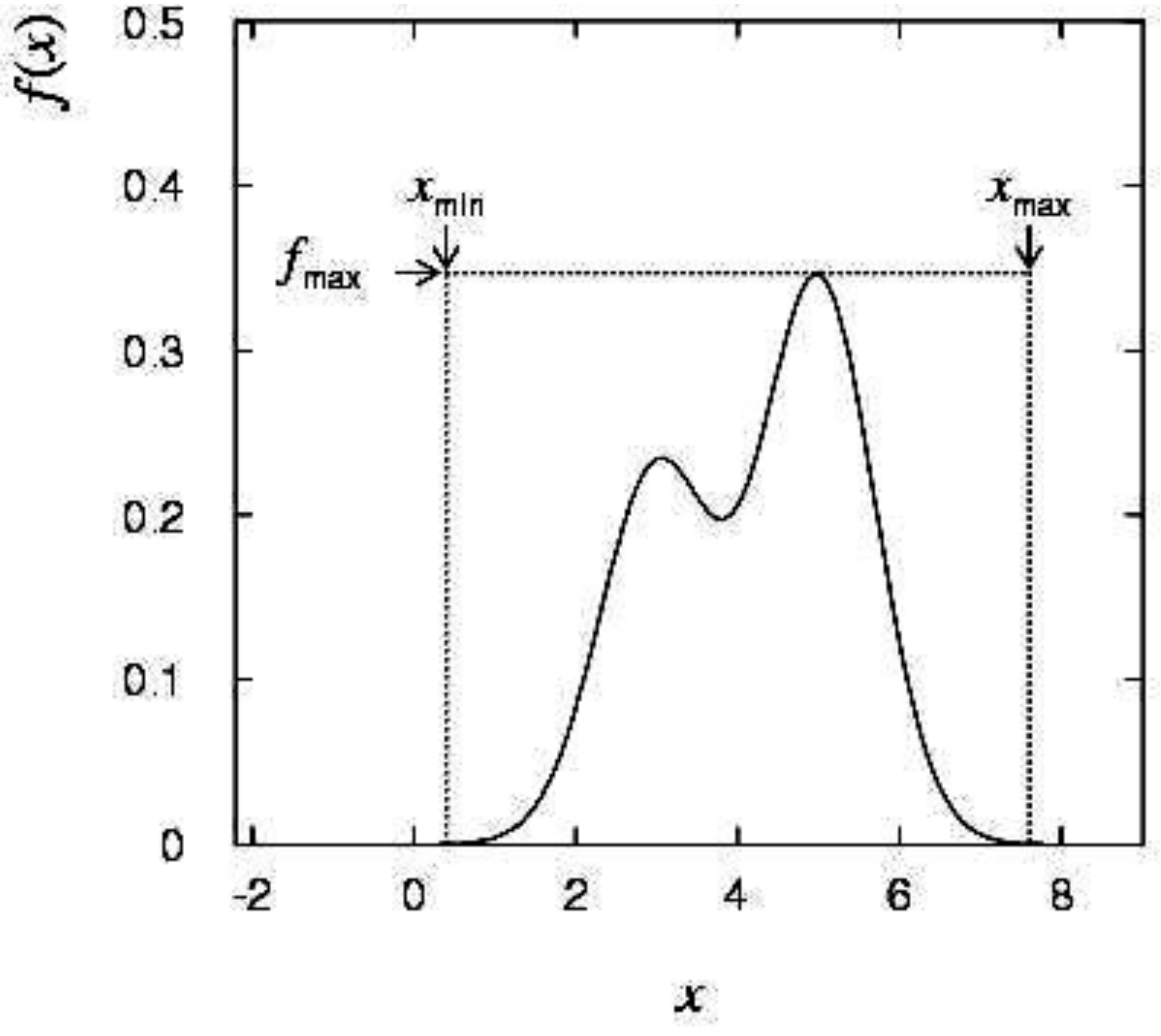
- One has to solve for x(r) analytically to use this method, not practical for arbitrary f(x)



The acceptance-rejection method

- (1) Generate a random number x, uniform in $[x_{min}, x_{max}]$, i.e. $x = x_{\min} + r_1 \left(x_{\max} - x_{\min} \right)$, r_1 is uniform in [0, 1].
- (2) Generate a 2nd independent random number u uniformly distributed between 0 and f_{max}, i.e. $u=r_2 f_{max}$
- (3) If $u<f(x)$, then accept x, if not, reject x and repeat
- Practical for f(x) with complicated shapes.

Enclose the pdf in a box

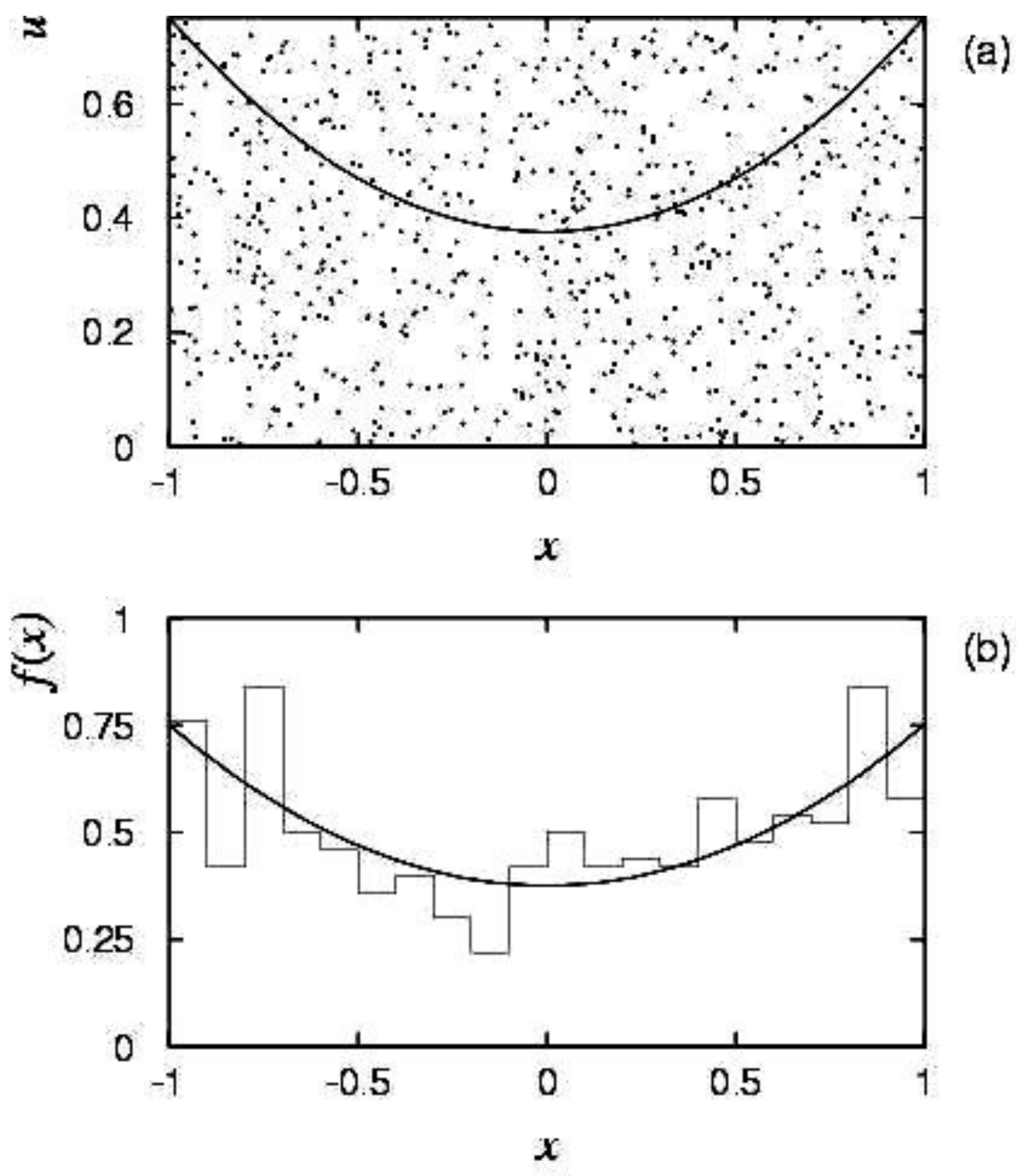


Example with acceptance-rejection method

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

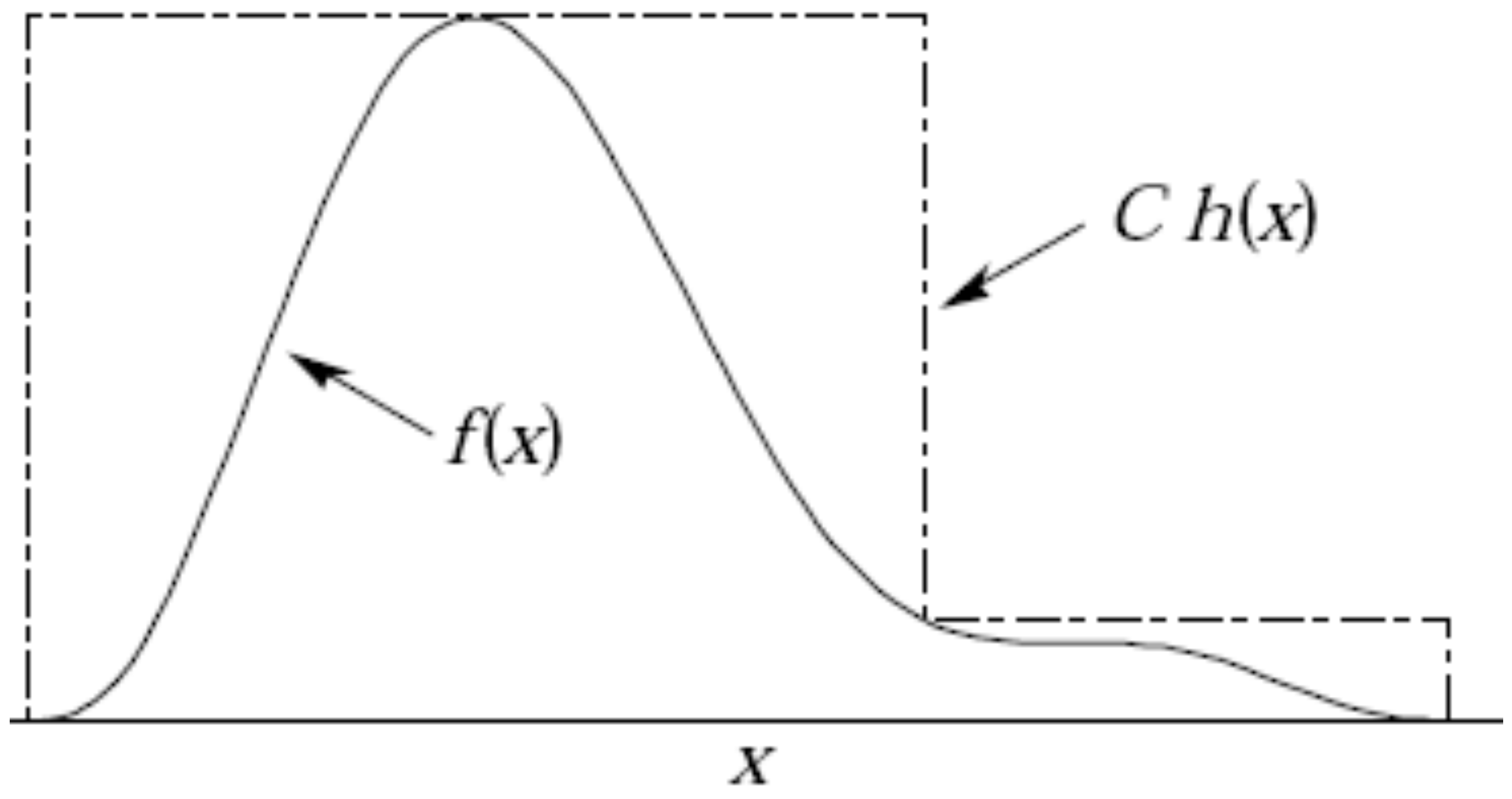
$$(-1 \leq x \leq 1)$$

- Generate random number x in [-1, 1]
- Generate another random number u in [0, 3/4]
- If u <= f(x), take the x, and fill it in the histogram. If not, drop the x.



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, the fraction maybe very low, 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:
  - (1) Generate random variable x according to C h(x) by other method, such as "transformation method".
  - (2) Generate another random number u uniformly distributted between 0 and C h(x)
  - (3) If u < f(x), take accept x.



Moving to the practical programing training.

Backup slides, Probability Density Functions.

Probability functions

- We will now introduce a short list of popular probability functions and pdfs.
- For each functions, we show expectation value, variance, a plot and discuss of some properties and applications.
- See also chapter on probability from pdg.lbl.gov

Binomial distribution

- Consider N independent experiments (Bernoulli trails):
  - outcome of each is “success” or “failure”
  - probability of “success” on any given trail is p
- Define discrete random variable n = number of successes (0 ≤ n ≤ N)
- Probability of a specific outcome (in order), e.g. “s s f s f” is:

$$pp(1-p)p(1-p) = p^n(1-p)^{N-n}$$

- But order is not important; there are $N!/[n!(N-n)!]$ ways (permutations) to get n successes in N trails, total probability for n is sum of probabilities for each permutation.

Binomial distribution

- The binomial distribution is therefore

$$f(n; N, p) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}$$

random variable (n) **parameters** (N, p)

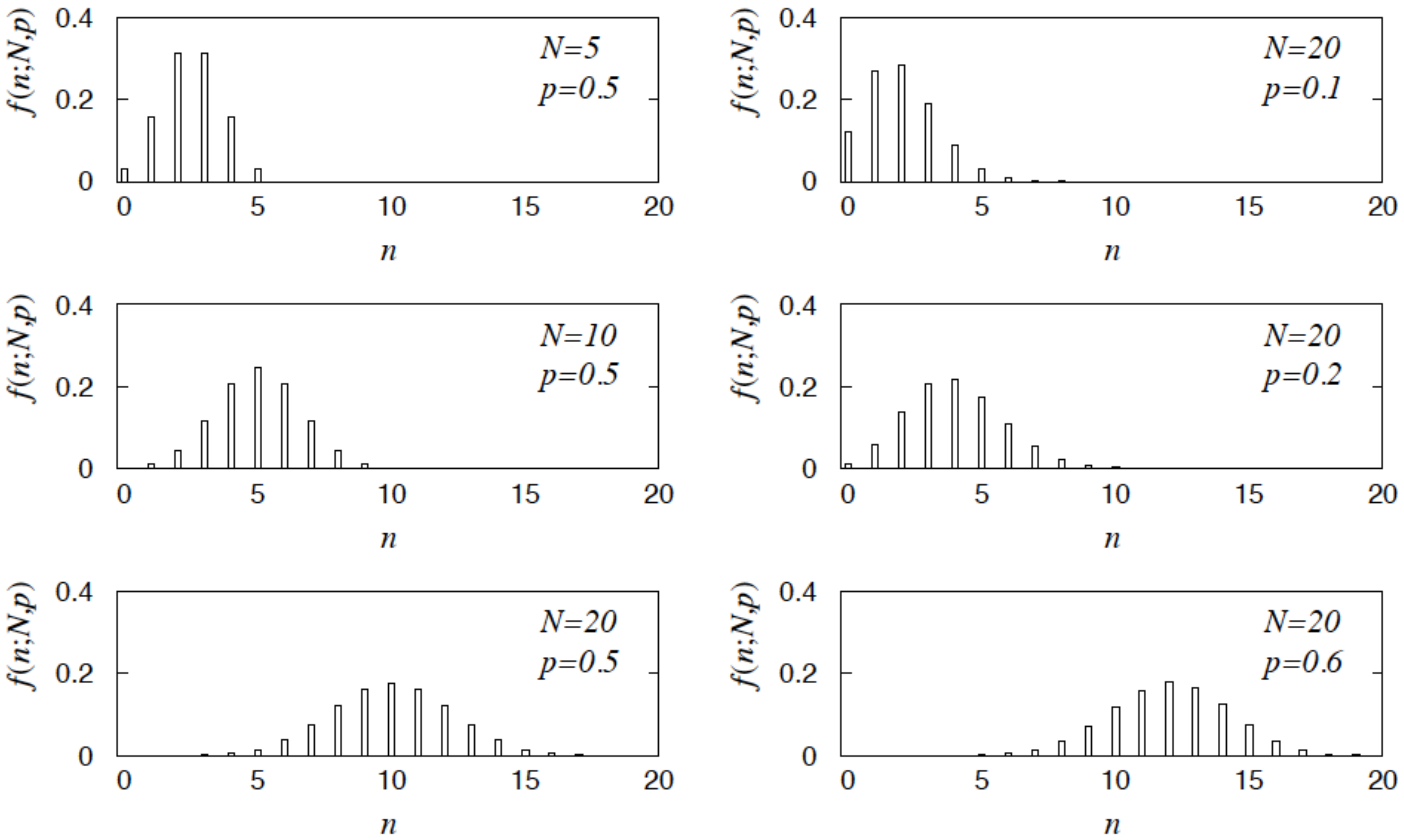
- For the expectation value and variance, we find

$$E[n] = \sum_{n=0}^{N} n f(n; N, p) = Np$$

$$V[n] = E\left[n^2\right] - (E[n])^2 = Np(1-p)$$

Binomial distribution

- Binomial distribution for different parameters values



- Example: observe N decays of W bosons, in which n is the number of $W \to \mu\nu$ decays (number of “success”), and p is the branching ratio (probability of “success”).

Multinomial distribution

- Like binomial, but now m outcomes instead of two, probabilities are

$$\vec{p} = \left(p_1, \ldots, p_m\right), \quad \textbf{with} \ \sum_{i=1}^{m} p_i = 1$$

- For N trails, we want the probability to obtain:

$$\begin{aligned} &n_1 \text{ of outcome } 1, \\ &n_2 \text{ of outcome } 2, \\ &\qquad \vdots \\ &n_m \text{ of outcome } m. \end{aligned}$$

- This is the multinomial distribution for $\vec{n} = \left(n_1, \ldots, n_m\right)$

$$f(\vec{n}; N, \vec{p}) = \frac{N!}{n_1!n_2!\cdots n_m!} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m}$$

Multinomial distribution

- Now consider outcome i as "success", all others are "failure".

$$E\left[n_i\right] = Np_i, \quad V\left[n_i\right] = Np_i\left(1 - p_i\right) \qquad \text{for all } i$$

  - Note, i (success) and NOT-i (failure), actually form a "binomial" structure

- One can also find the covariance to be

$$V_{ij} = Np_i\left(\delta_{ij} - p_j\right)$$

- Example: $\vec{n} = \left(n_1, \ldots, n_m\right)$ represents a histogram with m bins, N total entries, all entries independent.

Poisson distribution

- Consider binomial n in the limit

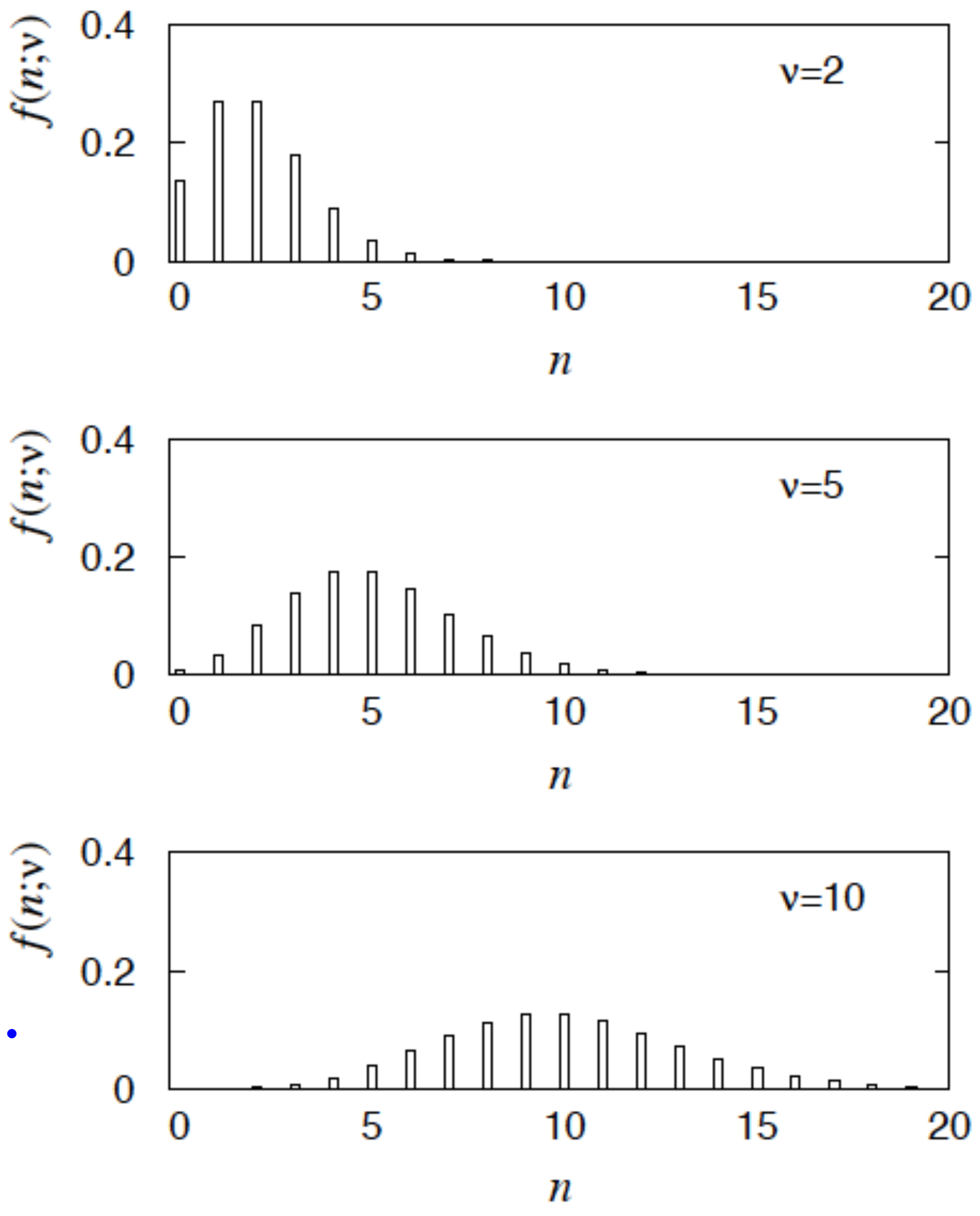
$$N \to \infty, \quad p \to 0, \quad E[n] = Np \to \nu$$

- n follows the Poisson distribution:

$$f(n;\nu) = \frac{\nu^n}{n!}e^{-\nu} \quad (n \geq 0)$$

$$E[n] = \nu, \quad V[n] = \nu$$

- Example: number of scattering events n with cross section σ found for a fixed integrated luminosity, with $\nu = \sigma \int L dt$



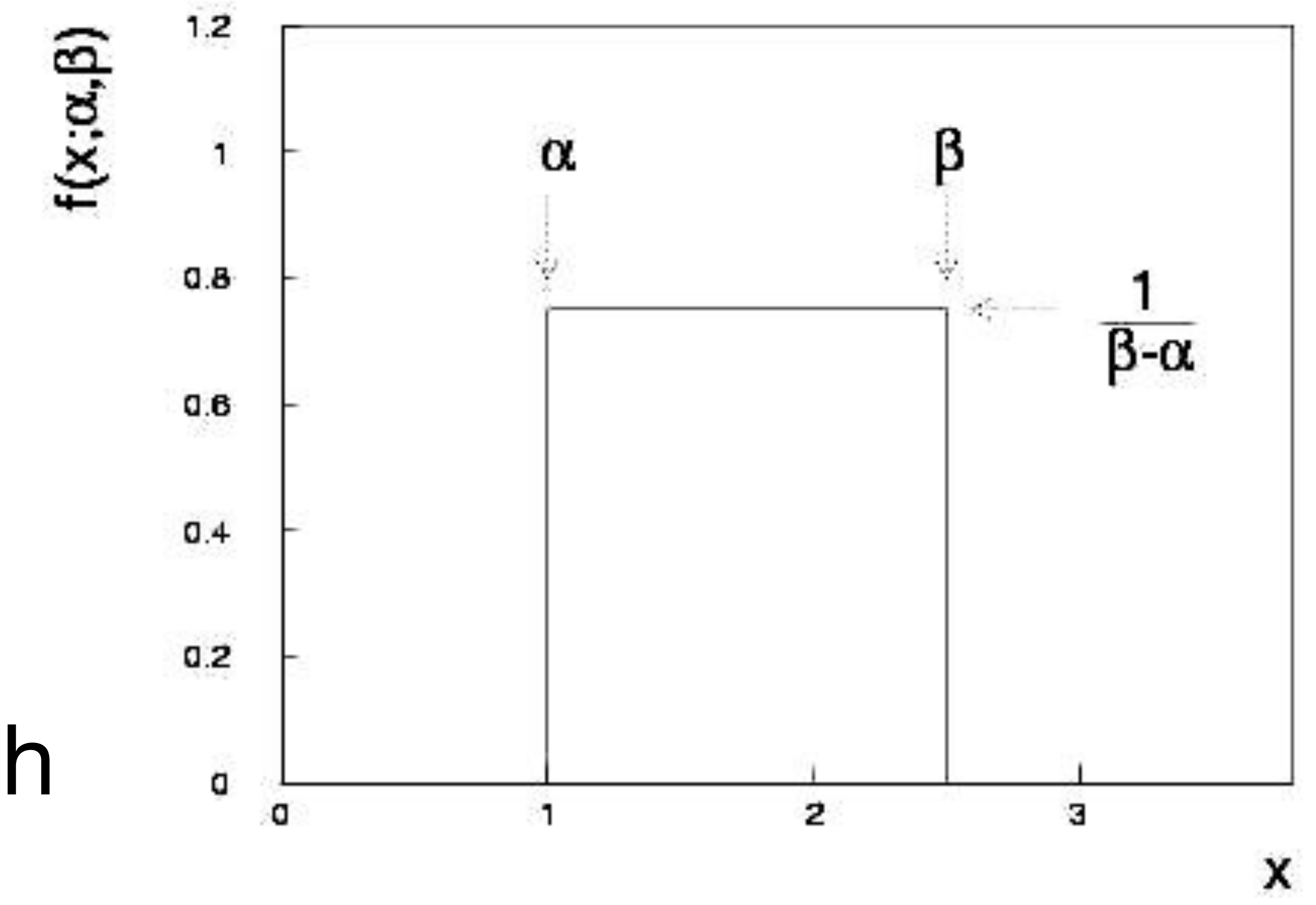
Uniform distribution

- Consider a continuous random variable with $-\infty < x < \infty$
  Uniform pdf is

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

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

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



- Note, for any random variable x with cumulative distribution F(x), y = F(x) is uniform in [0,1]
- Example: for $\pi^0 \to \gamma\gamma$, E_γ is uniform in $[E_{\min}, E_{\max}]$, with

$$E_{\min} = \frac{1}{2}E_\pi(1-\beta), \quad E_{\max} = \frac{1}{2}E_\pi(1+\beta)$$

$$\beta = v/c$$

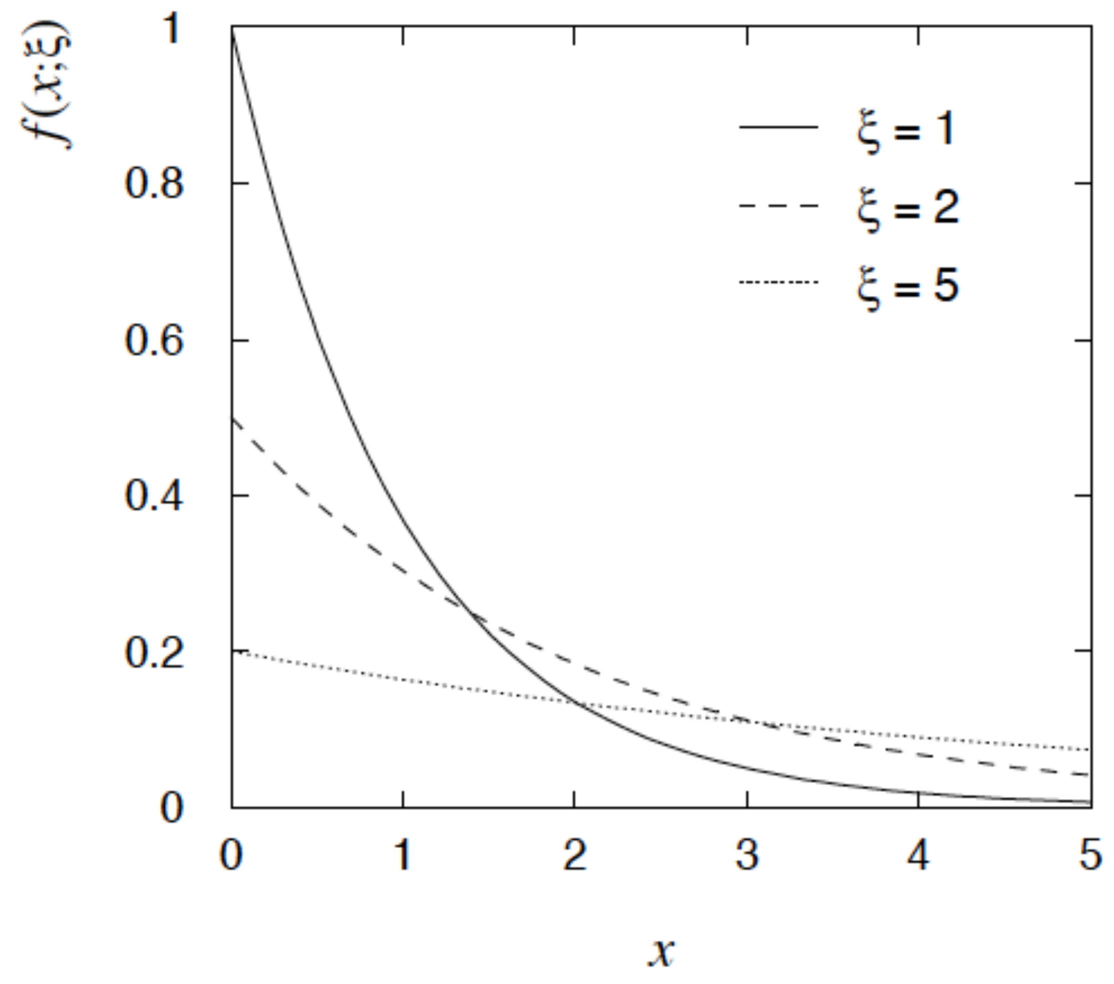
Exponential distribution

- The exponential pdf for the continuous random variable x is defined by:

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

$$E[x] = \xi$$

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



- Example: proper decay time t of an unstable particle

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

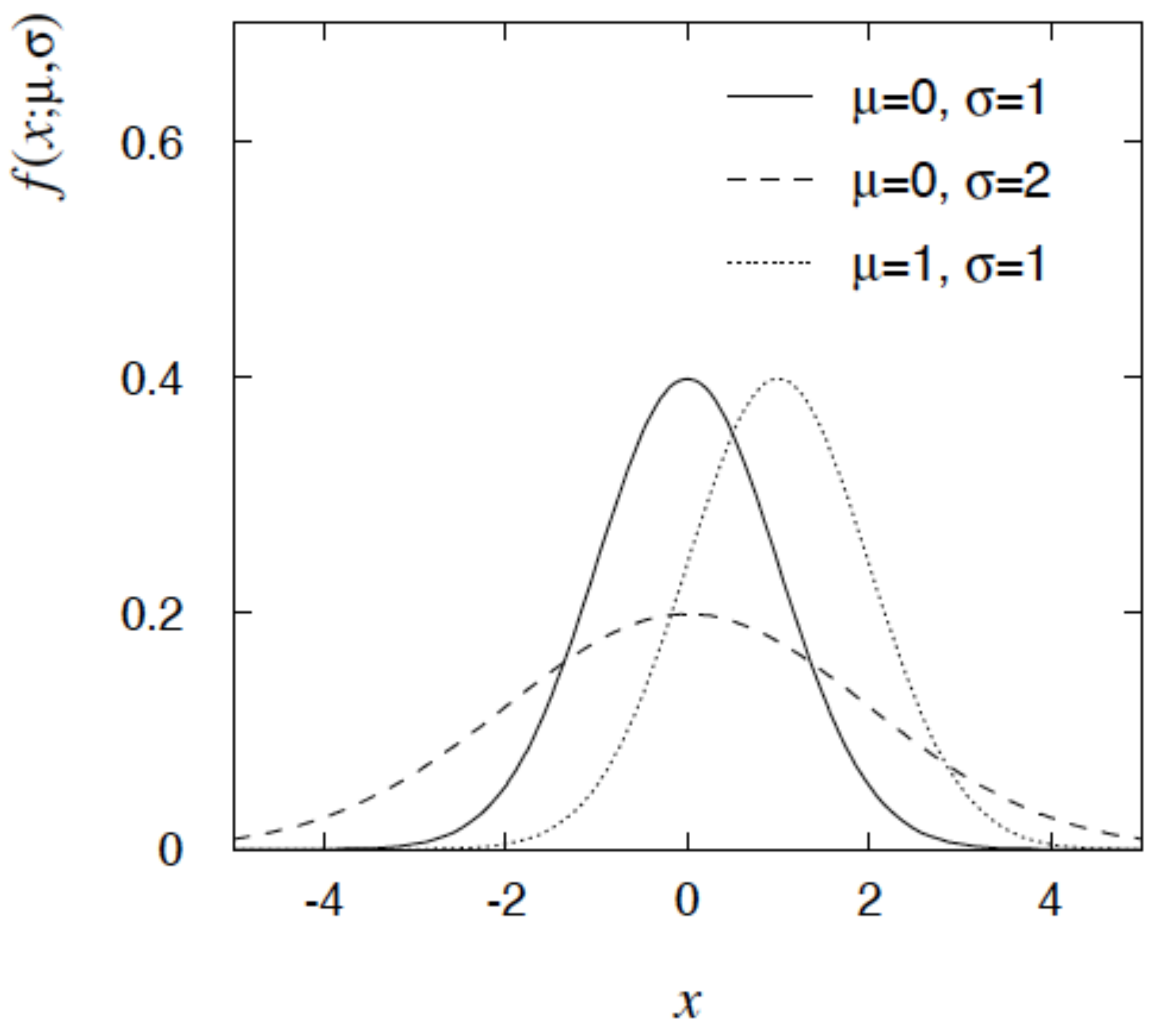
Gaussian distribution

- The Gaussian (normal) pdf for a continuous random variable x is defined by

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

$E[x] = \mu$ (N.B. often μ, σ^2 denote mean, variance of any r.v., not only Gaussian.)

$V[x] = \sigma^2$



- Special case: $\mu = 0,\ \sigma^2 = 1$ "standard Gaussian"

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

- If y ~ Gaussian with $\mu,\ \sigma^2$, then $x = (y-\mu)/\sigma$ follows $\varphi(x)$

Gaussian pdf and the Central Limit Theorem

- The Gaussian pdf is very 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 random variables x_i with finite variances σ_i^2, otherwise arbitrary pdfs, consider the sum

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

- In the limit $n \to \infty$, y is a Gaussian random variable 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 random variables.

Central Limit Theorem (CLT)

- The CLT can be proved using characteristic functions (Fourier transforms).
- For finite n, the theorem is approximately valid to the extent that the fluctuation of the sum is not dominated by one (or few) terms.
  => beware of measurement errors with non-Gaussian tails.
- Good example: velocity components v_x of air molecules.
- OK example: total deflection due to multiple Coulomb scattering.
  (Rare large angle deflections give non-Gaussian tail.)
- Bad example: energy loss of charged particle traversing thin gas layer.

Multivariate Gaussian distribution

- Multivariate Gaussian pdf for the vector $\overrightarrow{x} = \left(x_1, \ldots, x_n\right)$

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

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

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

- For n=2 this is

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

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

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

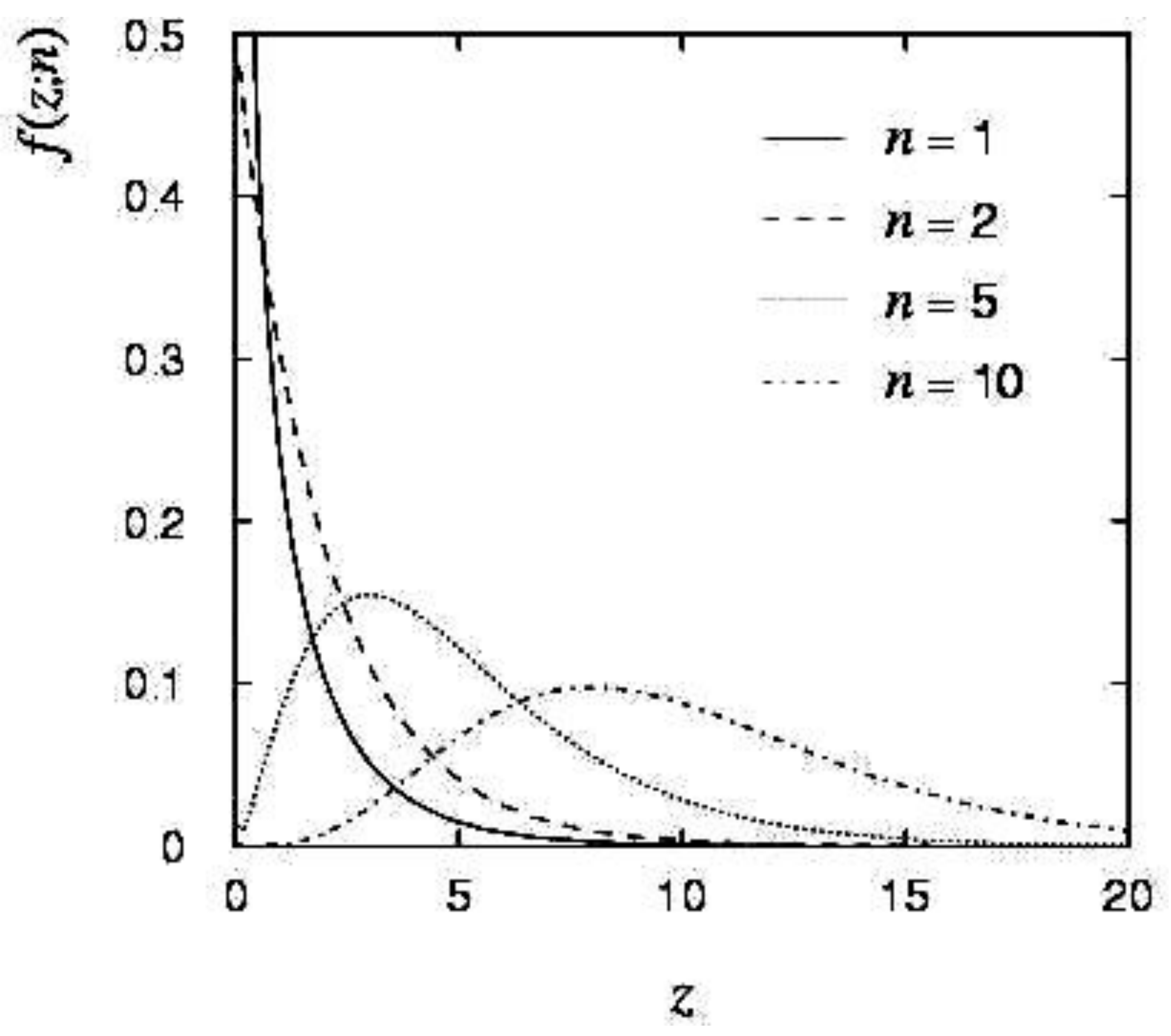
Chi-square (χ^2) distribution

- The chi-square pdf for the continuous random variable z (z>=0) is defined by

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

$n = 1, 2, \ldots =$ number of 'degrees of freedom' (dof)

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



- 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.

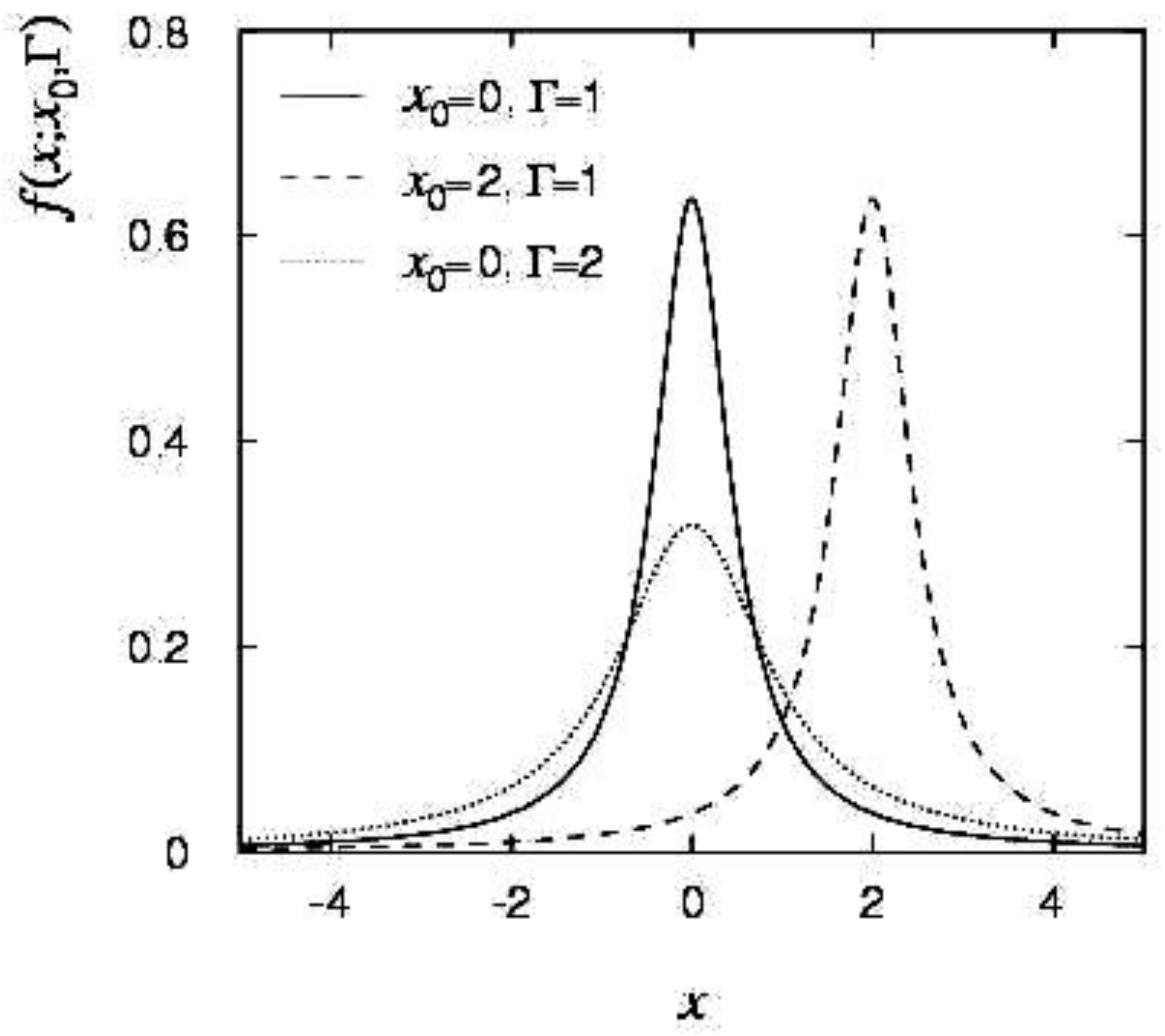
- Example: goodness-of-fit test variable especially in conjunction with method of least squares.

Cauchy (Breit-Wigner) distribution

- The Breit-Wigner pdf for the continuous random variable 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] is not well defined, $V[x] \to \infty$
  x_0 = mode (most probable value)
  Γ = full width at half maximum
- Example: mass of resonance particles, e.g. $Z^0, \rho, \mathrm{K}^*, \phi^0, \ldots$
- Γ = decay rate (inverse of mean lifetime)

Landau distribution

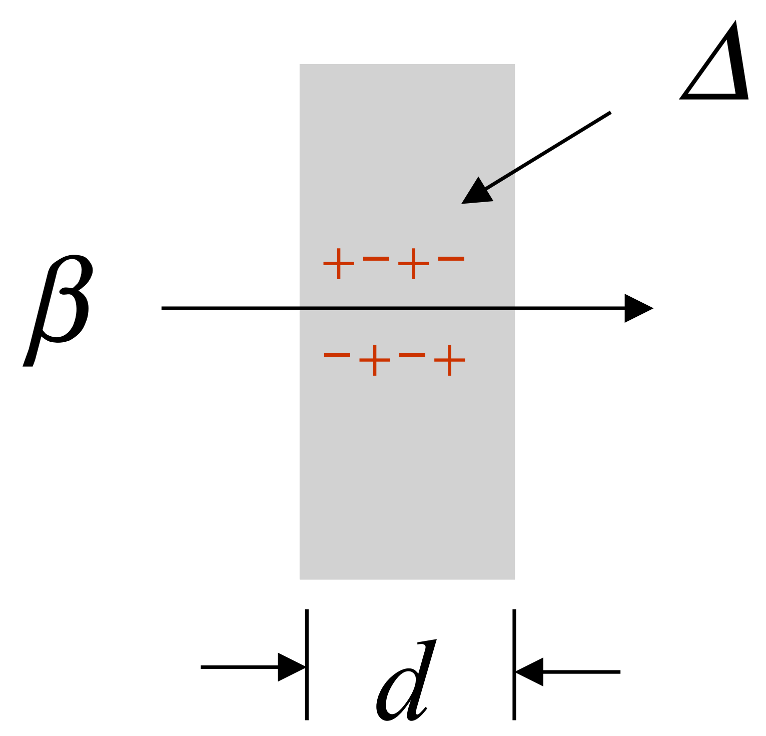
- For a charged particle with β = v/c traversing a layer of matter of thickness d, the energy loss Δ follows the Landau pdf:

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

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

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

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



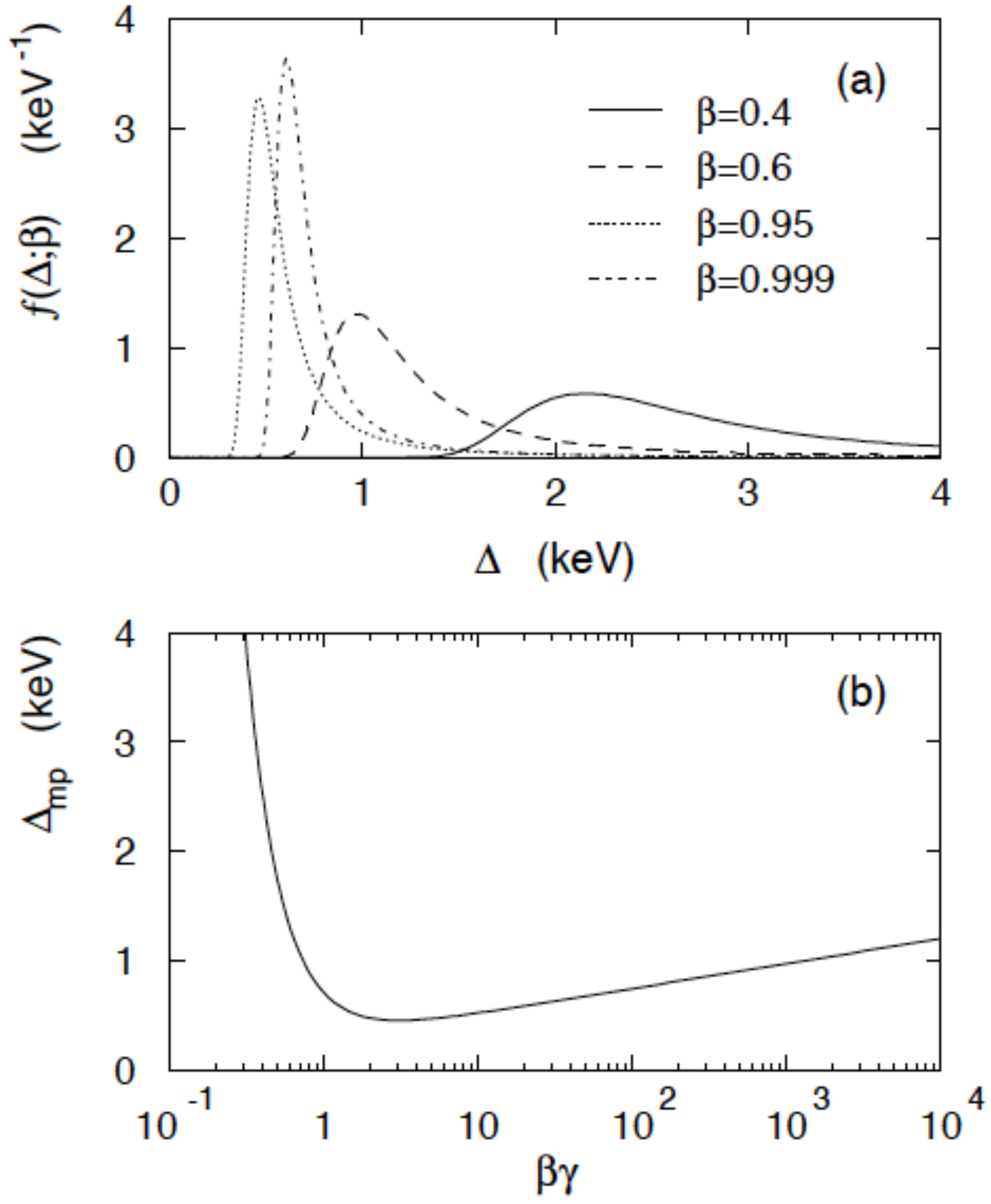
Landau distribution

Long ‘Landau tail’

→ all moments ∞

Mode (most probable value) sensitive to β,

→ particle i.d.



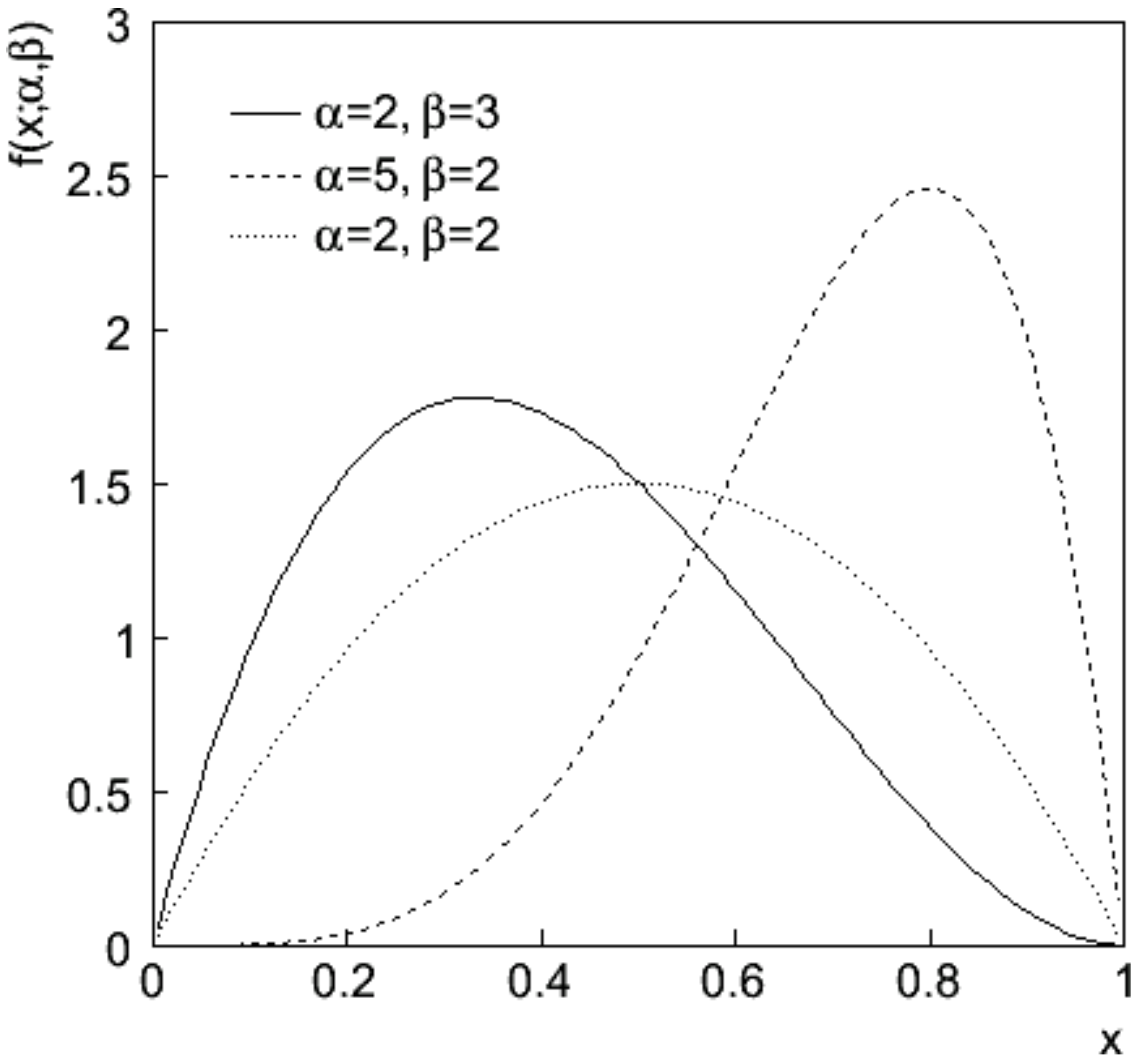
Beta distribution

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

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

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

- Often used to respresent pdf of continuous non-zero-only random variable between finite limits



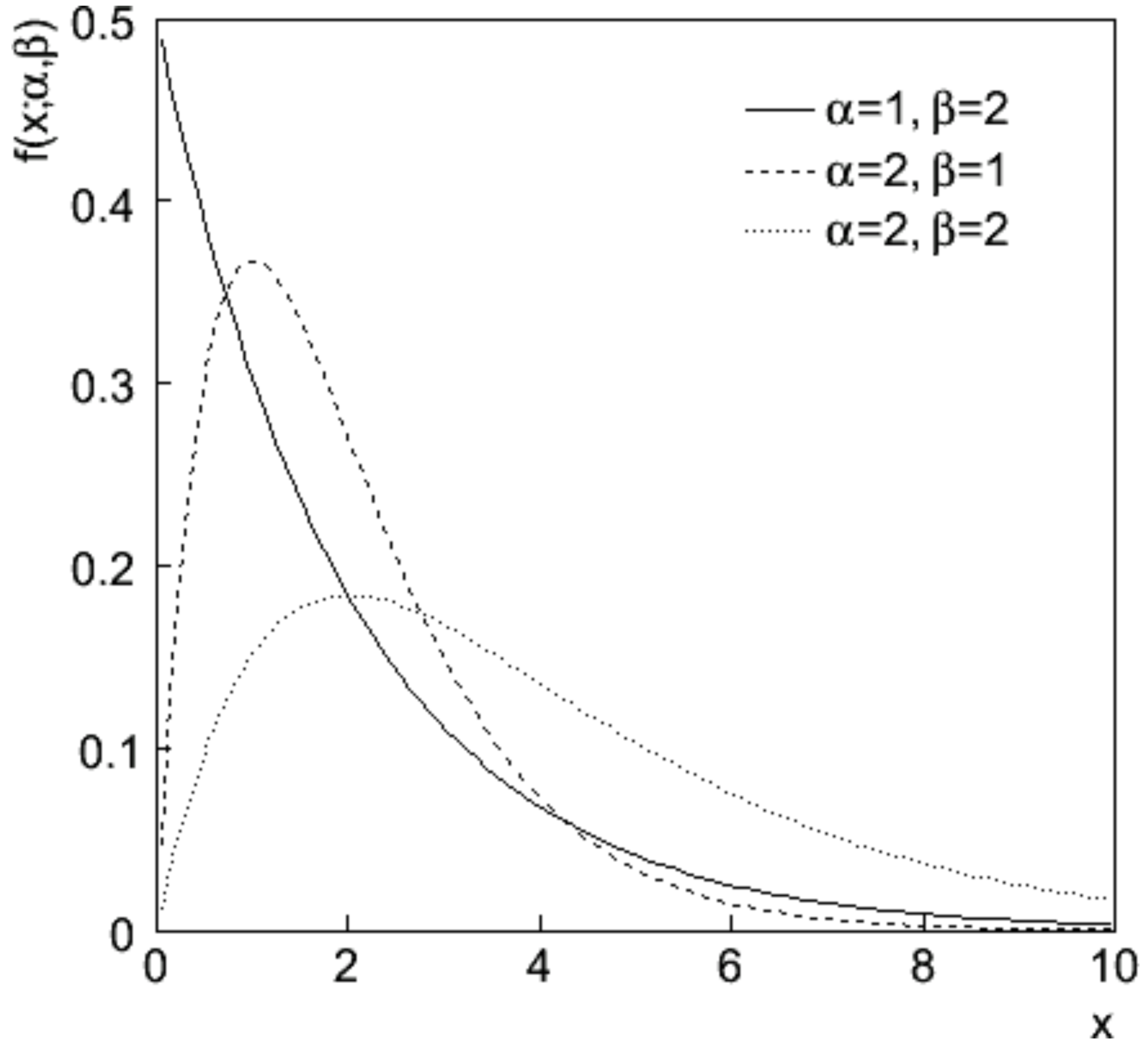
Gamma distribution

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

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

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

- Often used to represent pdf of continuous r.v. nonzero only in [0,∞].
- Also e.g. sum of n exponential random variables or time until nth event in Poisson process ~ Gamma

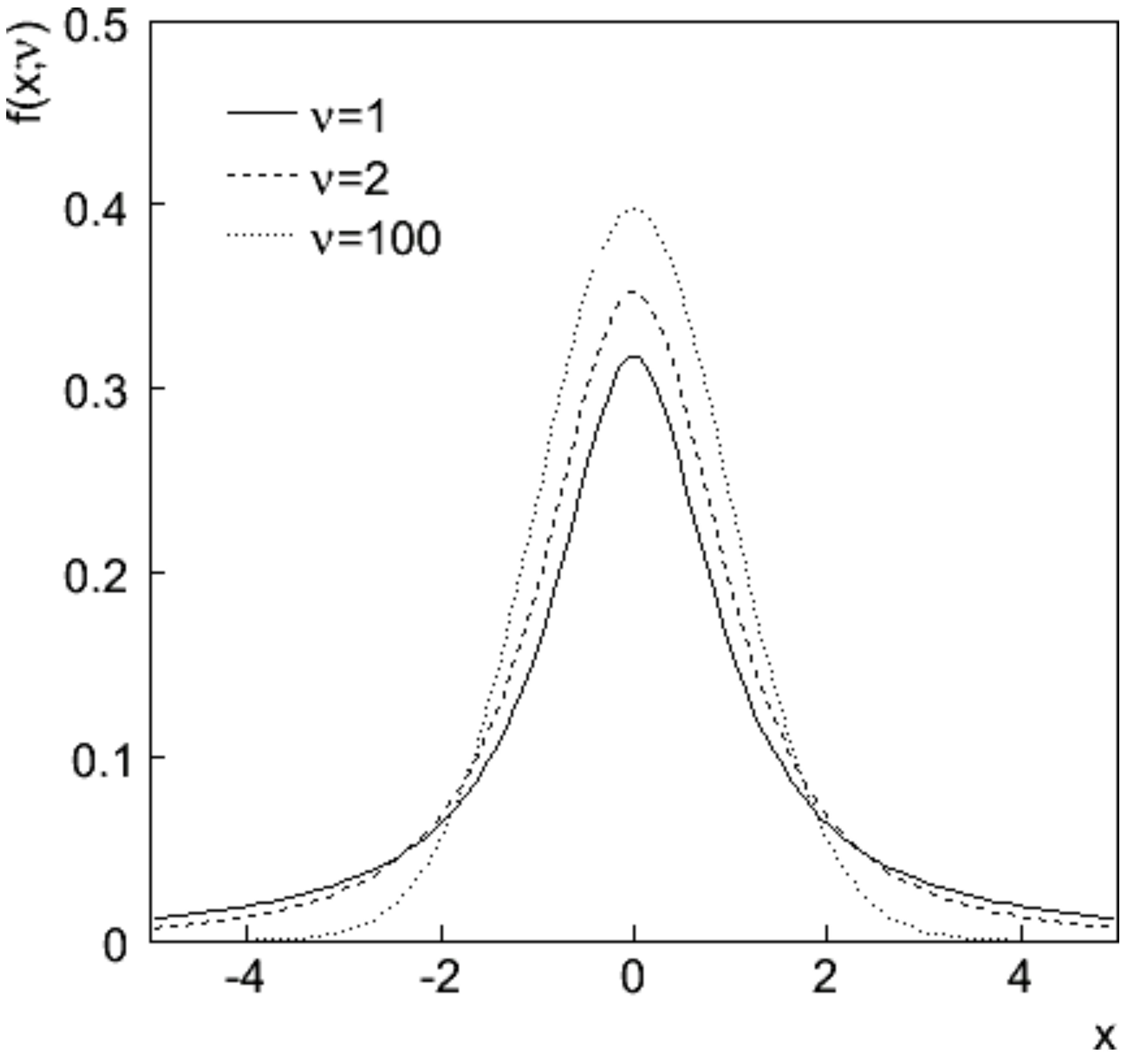


Student's t distribution

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

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

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



- ν = number of degrees of freedom (not necessarily integer)
- ν = 1 gives Cauchy,
- ν → ∞ gives Gaussian.

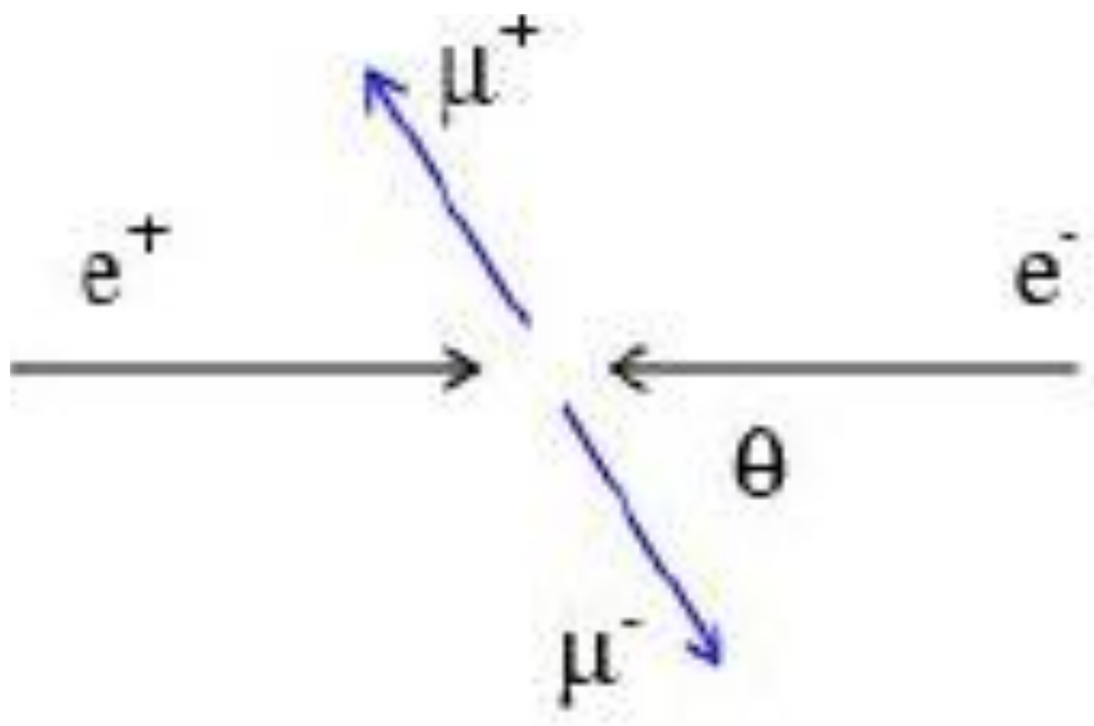
Student's t distribution

- If x ~ Gaussian with μ= 0, σ^2 = 1, and
  z ~ χ^2 with n degrees of freedom, then
  t = x / $(z/n)^{1/2}$ follows Student's t with v = n.
- This arises in problems where one forms the ratio of a sample mean to the sample standard deviation of Gaussian random variables.
- The Student's t provides a bell-shaped pdf with adjustable tails, ranging from those of a Gaussian, which fall off very
- quickly, (v → ∞, but in fact already very Gauss-like for v = two dozen), to the very long-tailed Cauchy (v = 1).

Monte Carlo event generators

- Simple example: $e^+e^- \rightarrow \mu^+\mu^-$
- Generate $\cos\theta$ and ϕ

$$f(\cos\theta; A_{FB}) \propto (1 + \frac{8}{3}A_{FB}\cos\theta + \cos^2\theta)\ ,$$
$$g(\phi) = \frac{1}{2\pi} \quad (0 \leq \phi \leq 2\pi)$$

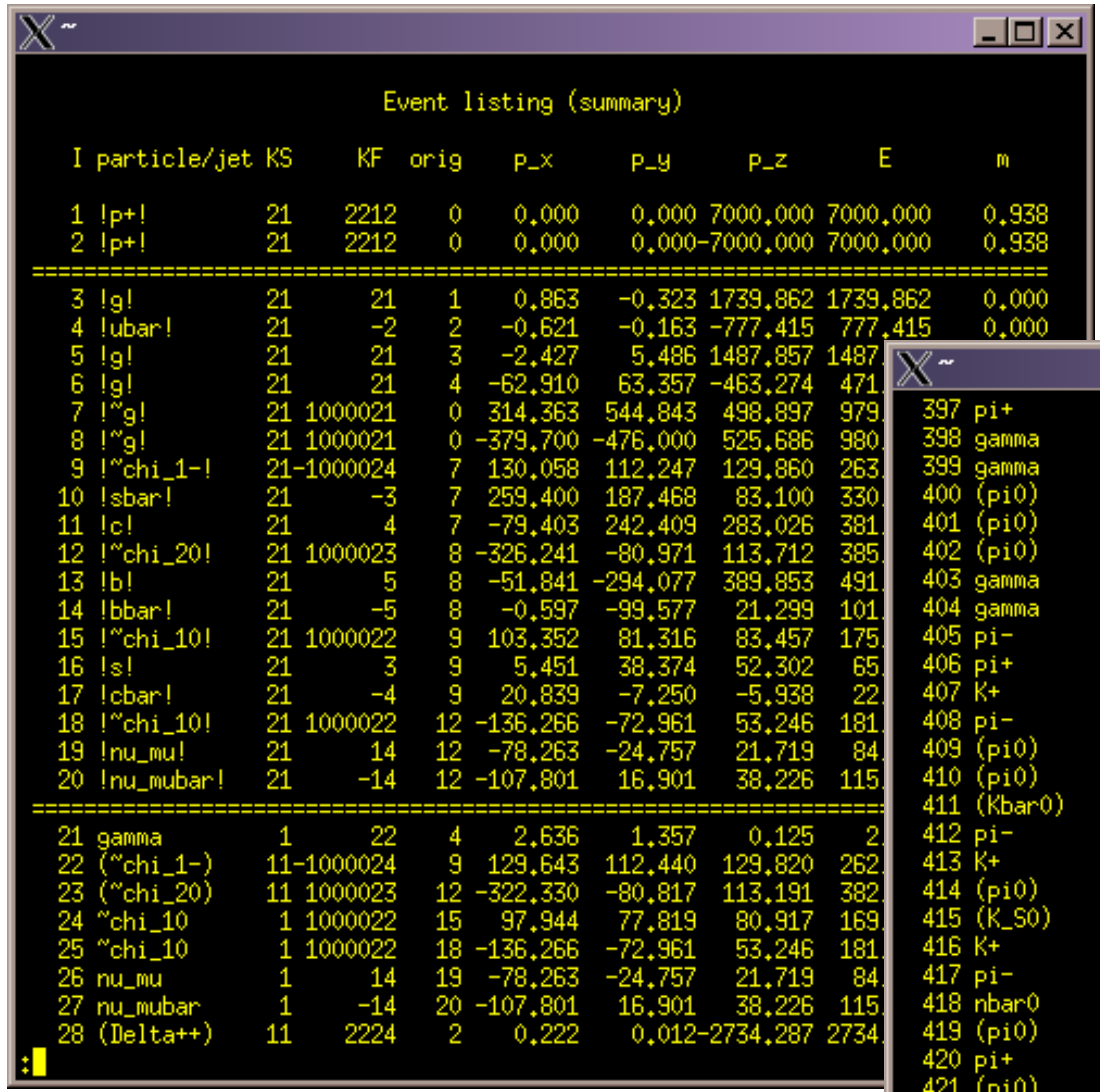


- Less simple: "event generators" for a variety of reactions:

$e^+e^- \rightarrow \mu^+\mu^-$, hadrons, ...
pp $\rightarrow$ hadrons, D-Y, SUSY,...

- e.g. PYTHIA, POWHEG, HERWIG, ISAJET …
- Output = "events", and for each event we get a list of generated particles and their momentum vectors, types, etc..

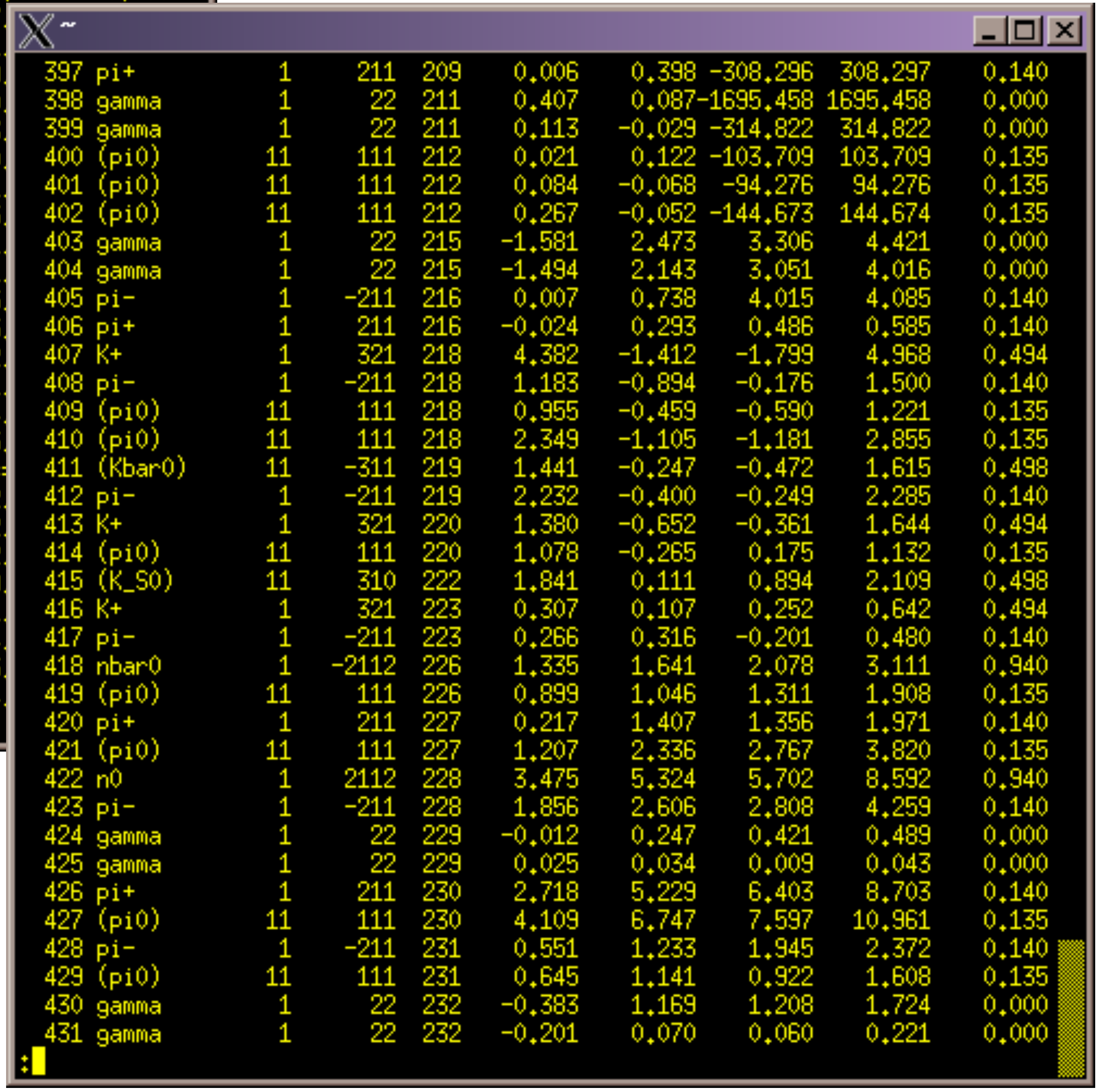
Example of a simulated event



```
                      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
:
```

A simulated event

⋮

⋮

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 -> input to reconstruction software: track finding, fitting, shower clustering, etc.
- Predict what you should see at "detector level" of events generated by the "event generator". Compare with "real data", e.g. can be used to estimate "efficiency" = [N events found] / [N events generated]
- Software: GEANT4