Computing and Statistical Data Analysis
Stat 2: Catalogue of pdfs

London Postgraduate Lectures on Particle Physics;

University of London 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:

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

Some distributions

Distribution/pdf	Example use in HEP
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

Binomial distribution

Consider N independent experiments (Bernoulli trials):

outcome of each is ‘success’ or ‘failure’,

probability of success on any given trial is p.

Define discrete r.v. n = number of successes ($0 \leq n \leq N$).

Probability of a specific outcome (in order), e.g. ‘ssfsf’ is

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

But order not important; there are $$\frac{N!}{n!(N-n)!}$$

ways (permutations) to get n successes in N trials, total probability for n is sum of probabilities for each permutation.

Binomial distribution (2)

The binomial distribution is therefore

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

random variable

parameters

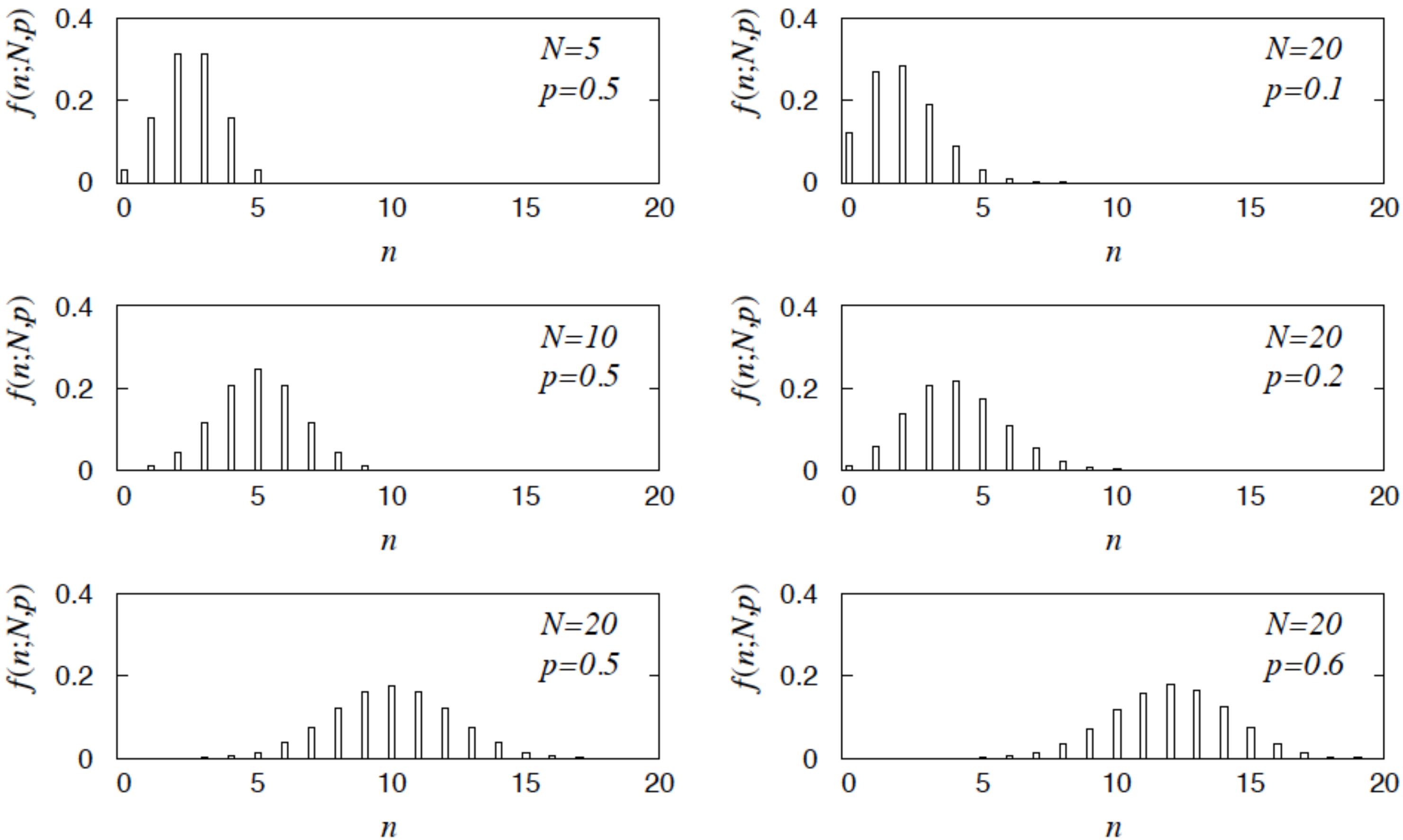
For the expectation value and variance we find:

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

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

Binomial distribution (3)

Binomial distribution for several values of the parameters:



Example: observe N decays of W$^{\pm}$, the number n of which are W→μν is a binomial r.v., p = branching ratio.

Multinomial distribution

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

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

For N trials we want the probability to obtain:

n_1 of outcome 1,
n_2 of outcome 2,
…
n_m of outcome m.

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

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

Multinomial distribution (2)

Now consider outcome i as 'success', all others as 'failure'.

→ all n_i individually binomial with parameters N, p_i

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

One can also find the covariance to be

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

Example: $\vec{n} = (n_1, \ldots, n_m)$ represents a histogram

with m bins, N total entries, all entries independent.

Poisson distribution

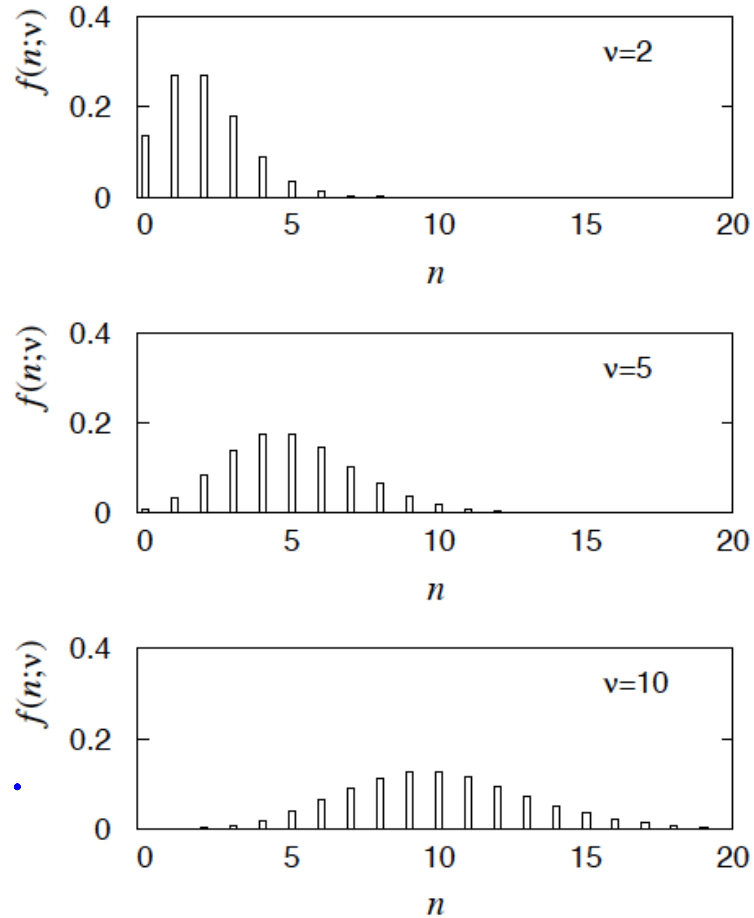
Consider binomial n in the limit

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

→ n follows the Poisson distribution:

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

$$E[n] = \nu \, , \qquad V[n] = \nu \, .$$



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

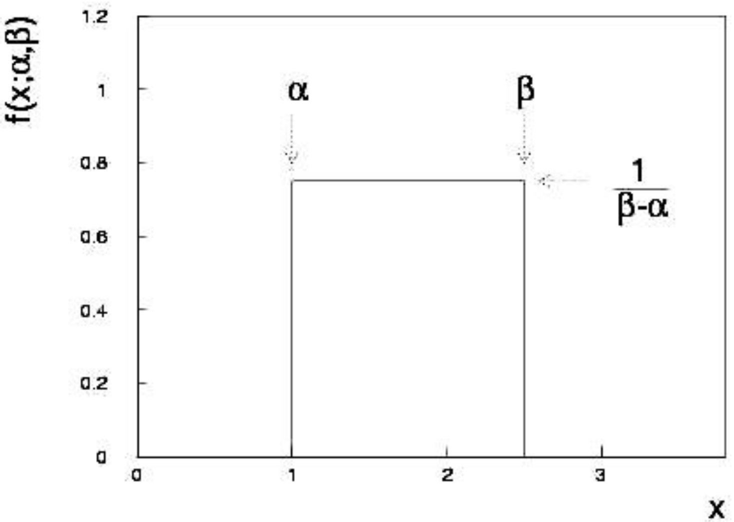
Uniform distribution

Consider a continuous r.v. x with $-\infty < x < \infty$. Uniform pdf is:

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

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

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



N.B. For any r.v. x with cumulative distribution $F(x)$, $y = F(x)$ is uniform in [0,1].

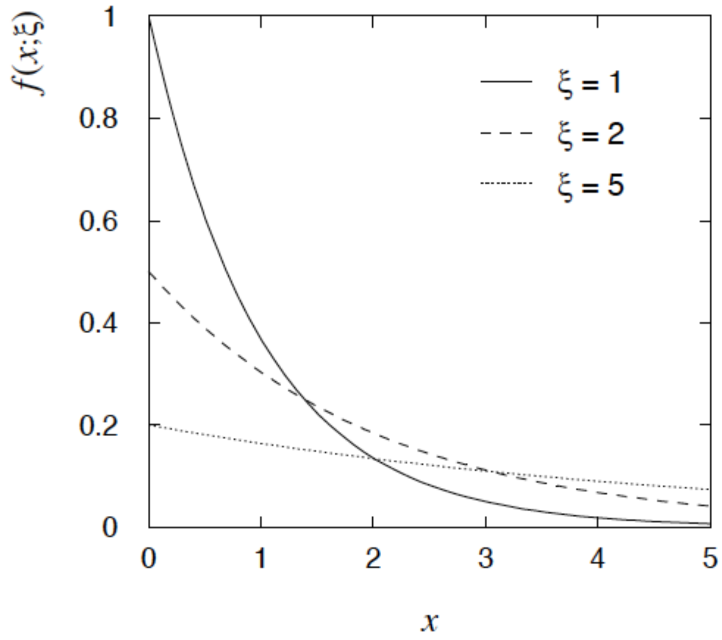
Example: for $\pi^0 \to \gamma\gamma$, E_γ is uniform in $[E_{min}, E_{max}]$, with

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

Exponential distribution

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

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



$$E[x] = \xi$$

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

Example: proper decay time *t* of an unstable particle

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

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

Gaussian distribution

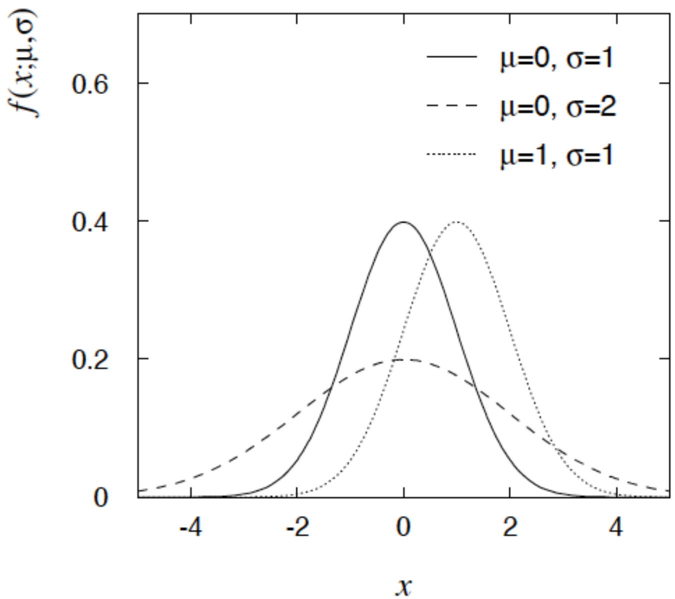
The Gaussian (normal) pdf for a continuous r.v. *x* is defined by:

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

$$E[x] = \mu$$

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

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



Special case: $\mu = 0$, $\sigma^2 = 1$ ('standard Gaussian'):

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

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

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, 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)

The CLT can be proved using characteristic functions (Fourier transforms), see, e.g., SDA Chapter 10.

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 component 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. (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,

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

For $n = 2$ this is

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

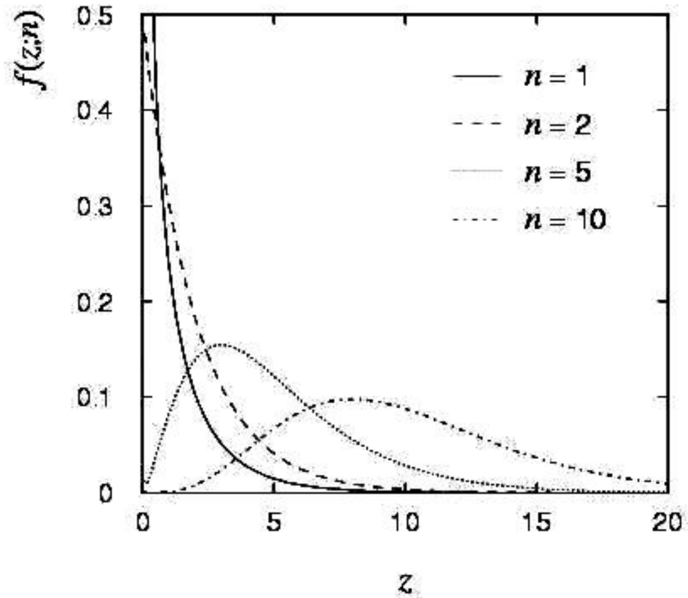
$$\times \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\}$$

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

Chi-square (χ^2) distribution

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

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



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

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

For independent Gaussian x_i, $i = 1, \ldots, n$, 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.

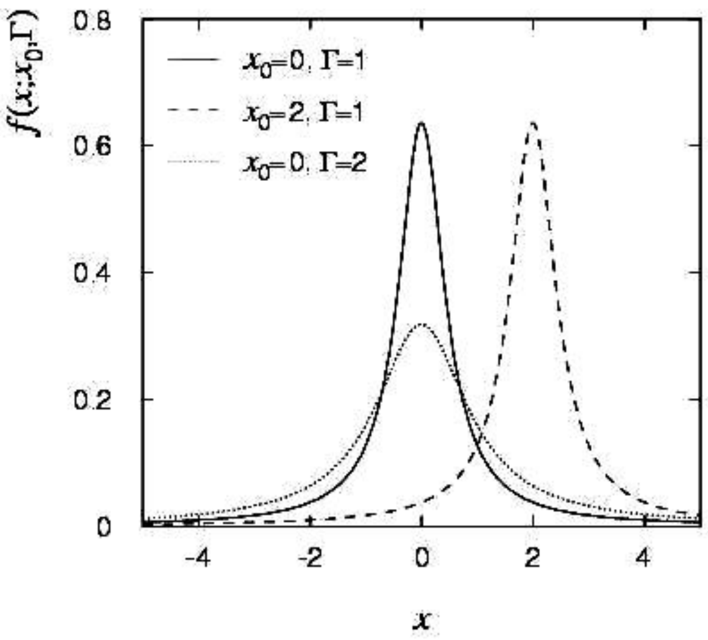
Example: goodness-of-fit test variable especially in conjunction with method of least squares.

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

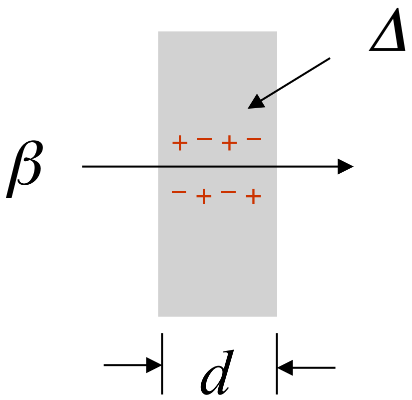
Example: mass of resonance particle, e.g. ρ, K^*, ϕ^0, ...

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.

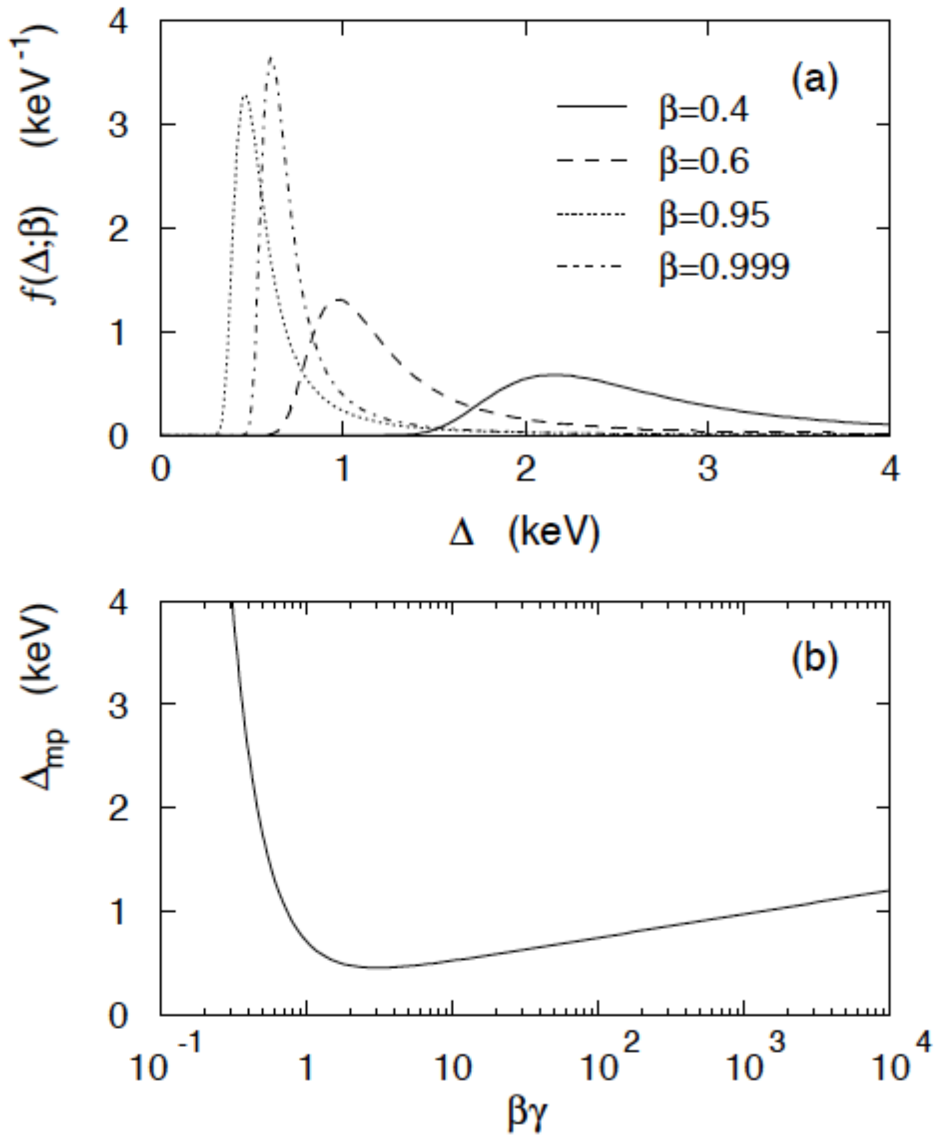
Landau distribution (2)

Long 'Landau tail'

→ all moments ∞

Mode (most probable value) sensitive to β ,

→ particle i.d.



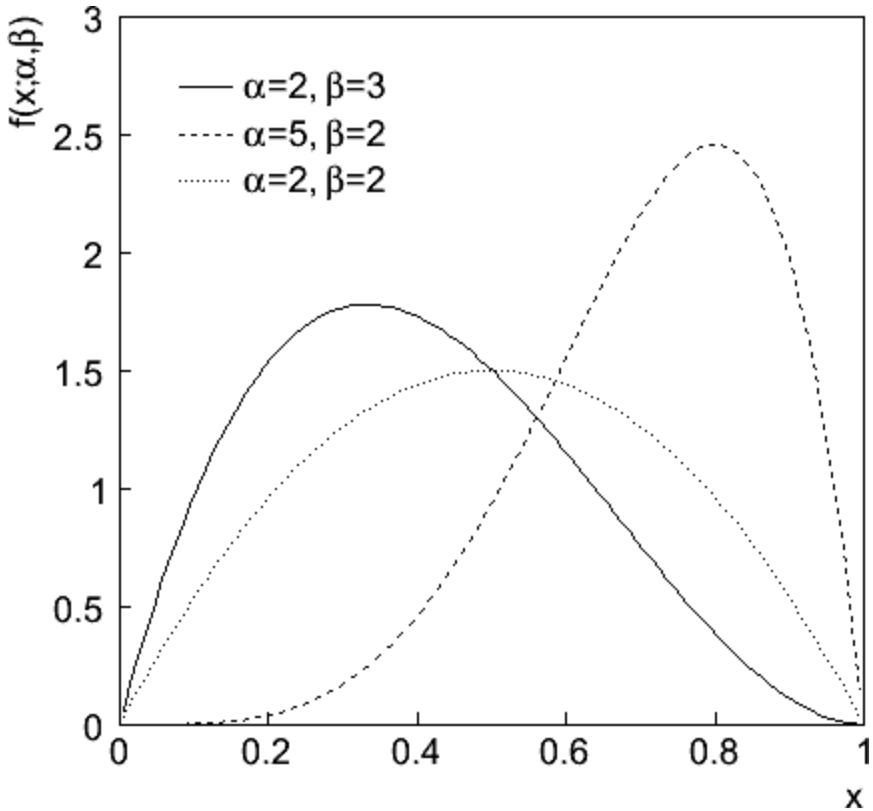
Beta distribution

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

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

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

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



Gamma distribution

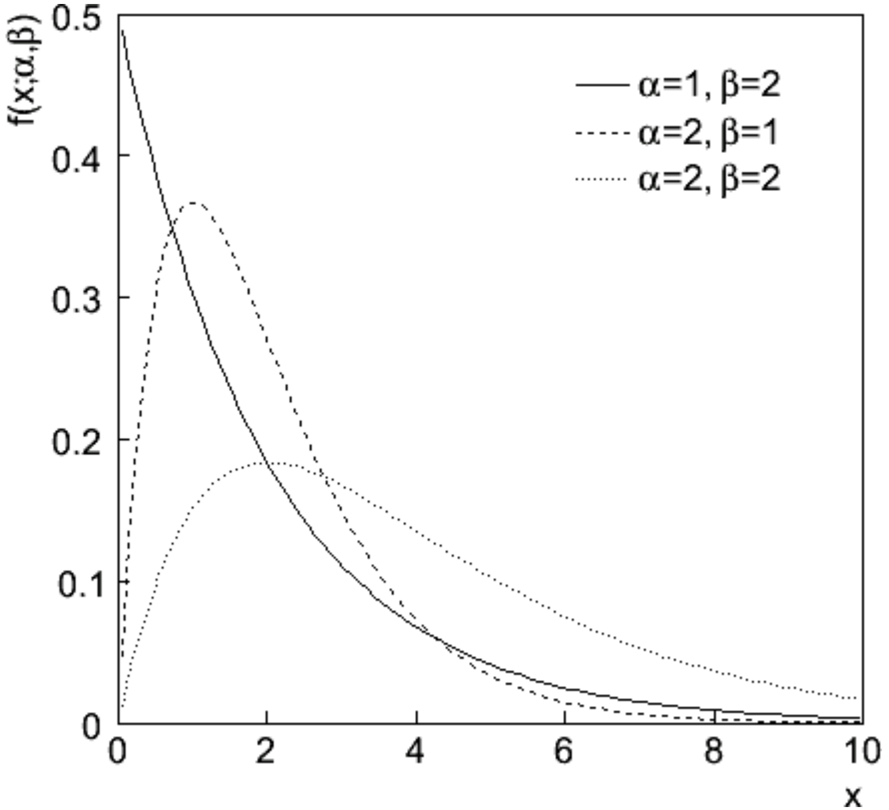
$$f(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}$$

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

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

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

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



Student's *t* distribution

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

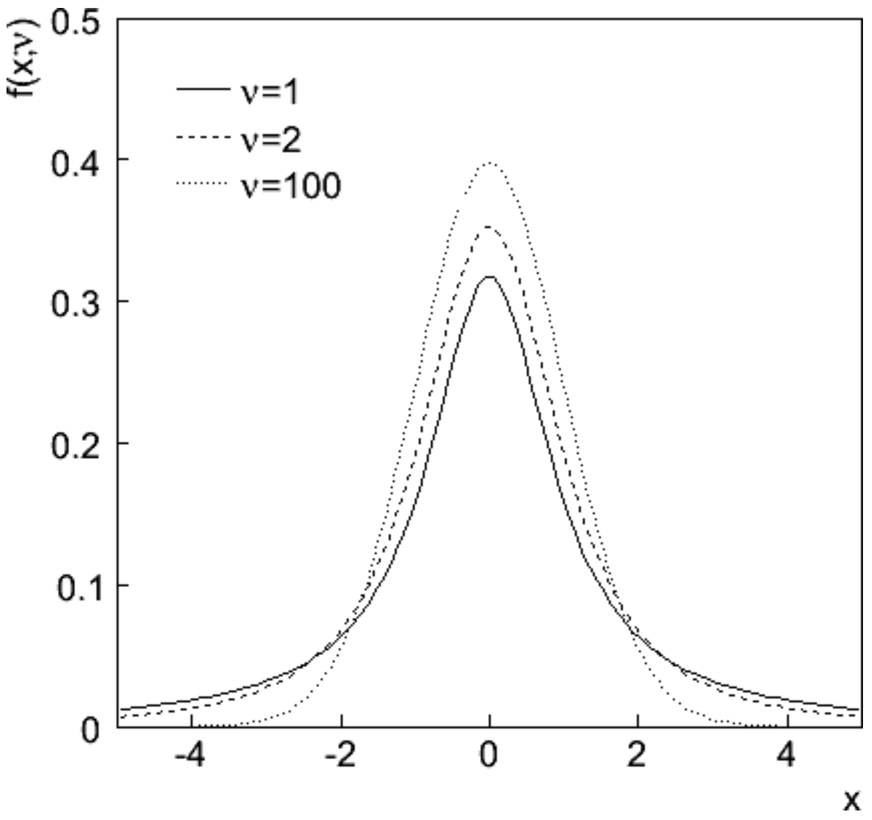
$$E[x] = 0 \quad (\nu > 1)$$

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

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

ν = 1 gives Cauchy,

$\nu \to \infty$ gives Gaussian.



Student's t distribution (2)

If $x \sim$ 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, ($\nu \rightarrow \infty$, 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 under the pseudonym "Student" for the Guinness Brewery.

Wrapping up lecture Stat 2

We’ve looked at a number of important distributions:

Binomial, Multinomial, Poisson, Uniform, Exponential
Gaussian, Chi-square, Cauchy, Landau, Beta,
Gamma, Student's t

and we’ve seen the important Central Limit Theorem:

explains why Gaussian r.v.s come up so often

For a more complete catalogue see e.g. the handbook on statistical distributions by Christian Walck from

`http://www.physto.se/~walck/`