Statistische Methoden der Datenanalyse

Kapitel 2: Spezielle Wahrscheinlichkeitsverteilungen

Professor Markus Schumacher
Freiburg / Sommersemester 2009

Basiert auf Vorlesungen und Folien von Glen Cowan und Abbildungen von L Feld, S, Zech, H. Kalinowski.

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 $\dfrac{N!}{n!(N-n)!}$

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

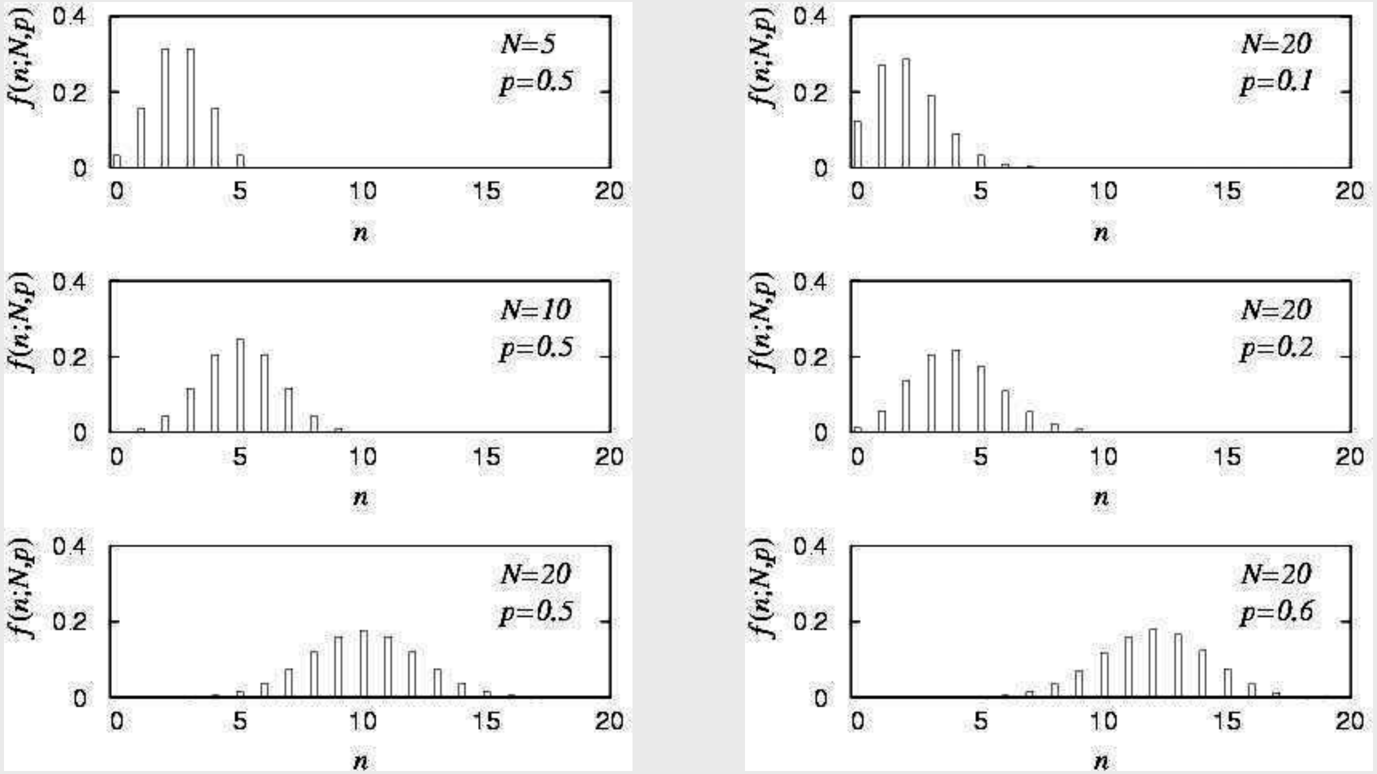
For the expectation value and variance we find:

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

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

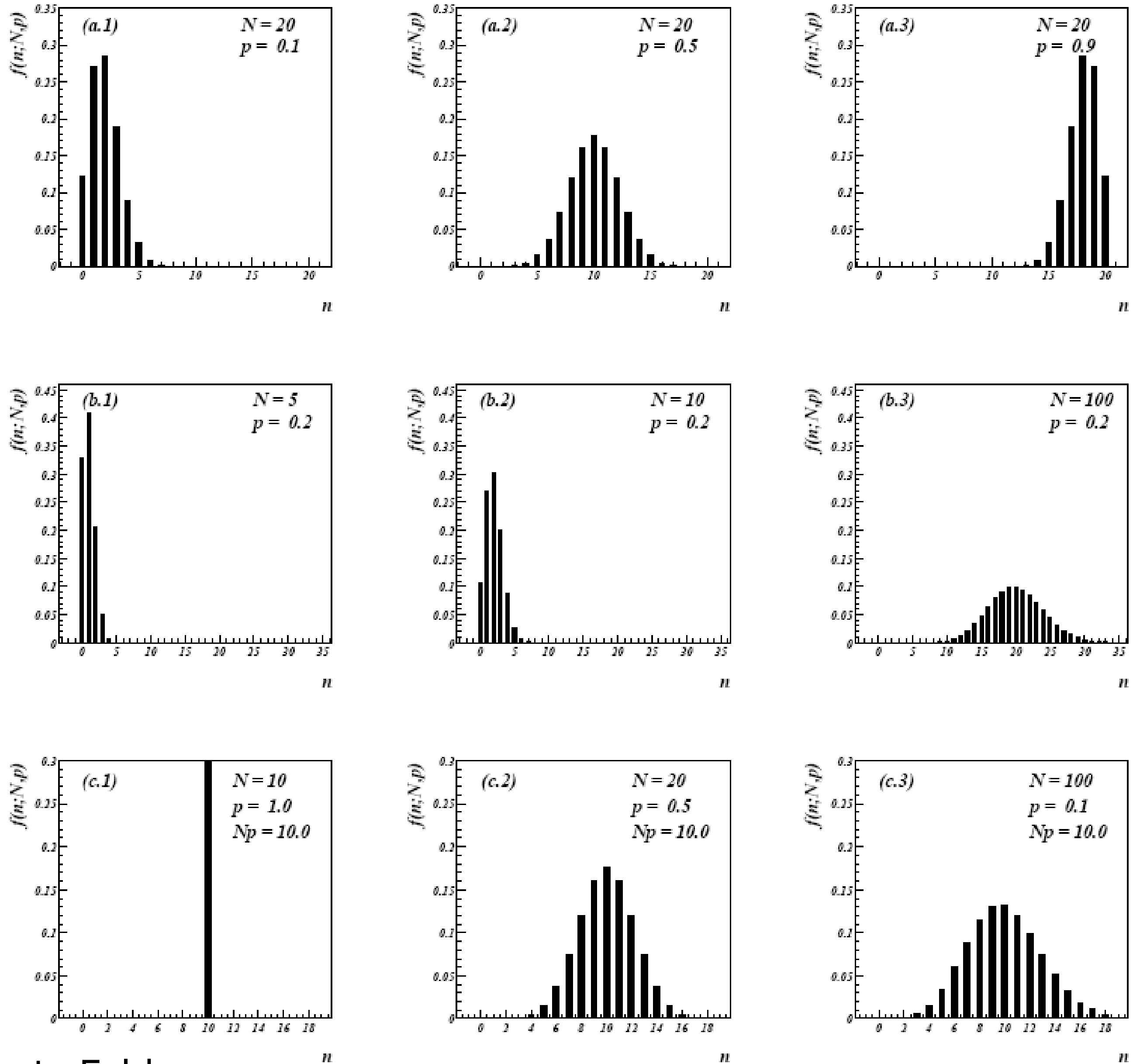
Binomial distribution (3)

Binomial distribution for several values of the parameters:



Example: observe *N* decays of W , the number *n* of which are $W \rightarrow \mu\nu$ is a binomial r.v., *p* = branching ratio.

Binomial distributions for different parameters p and N



Lutz Feld

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

$$E[n_i] = Np_i, \quad V[n_i] = Np_i(1 - p_i)$$ 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

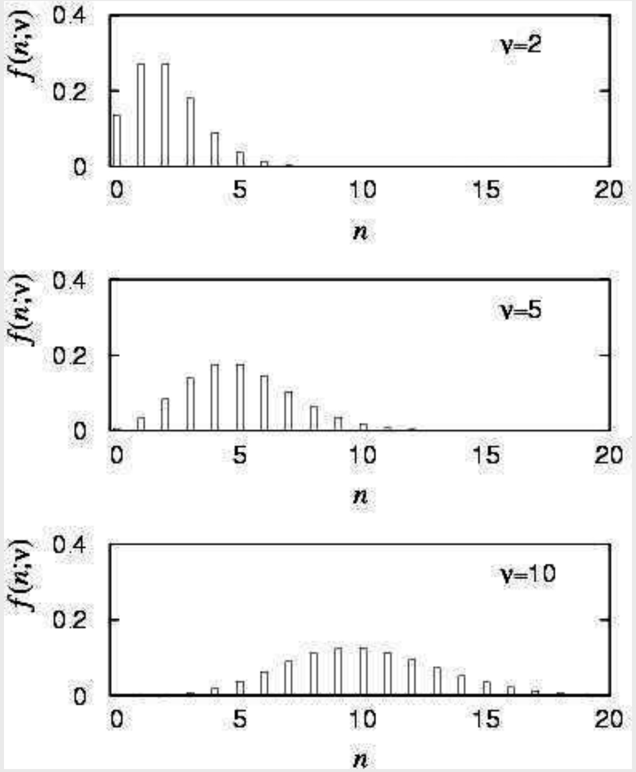
Consider binomial *n* in the limit

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

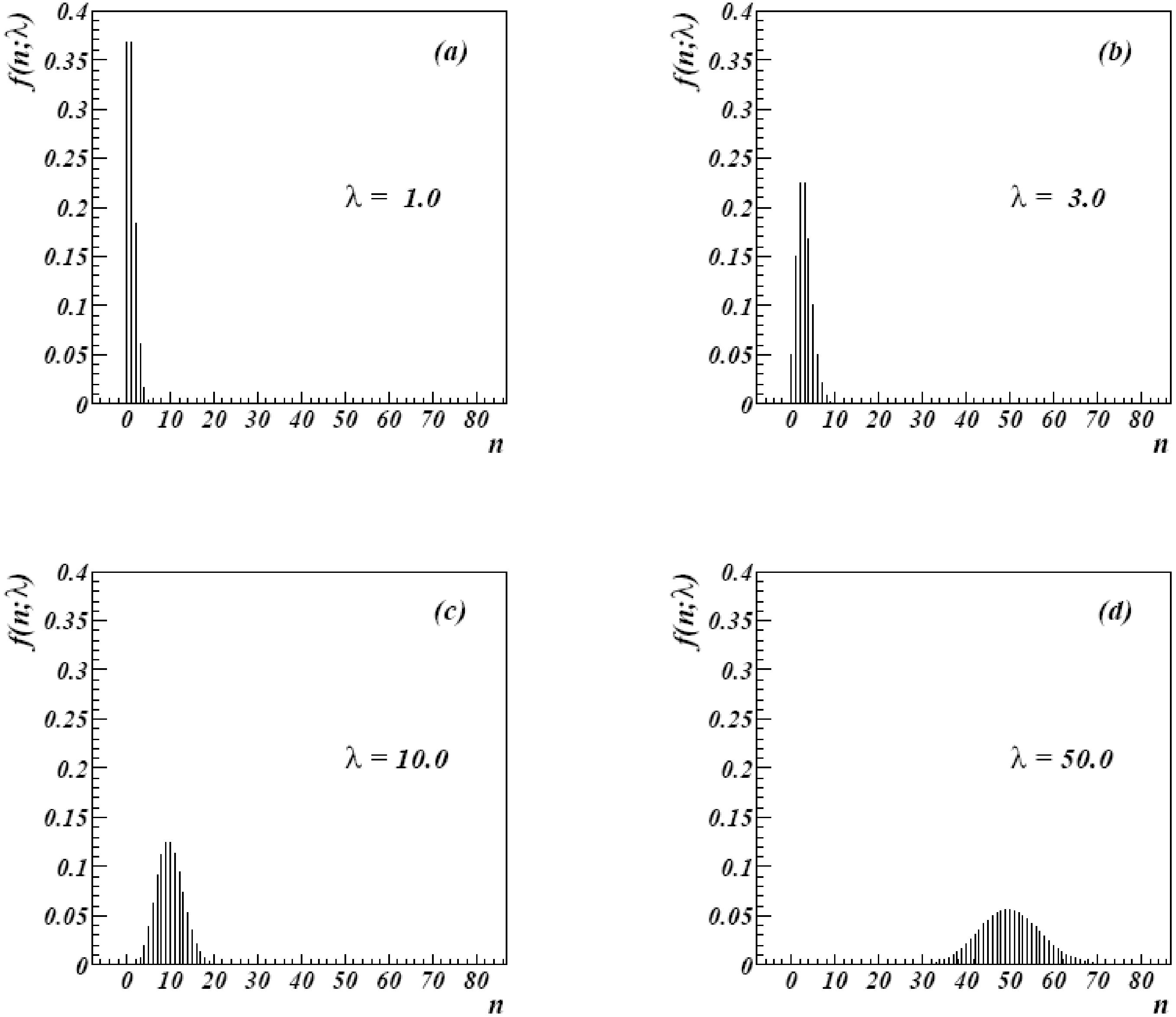
→ *n* follows the Poisson distribution:

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

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



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



f(n;λ)
(a)
λ = 1.0
(b)
λ = 3.0
(c)
λ = 10.0
(d)
λ = 50.0
0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4
0 10 20 30 40 50 60 70 80
n

Lutz Feld

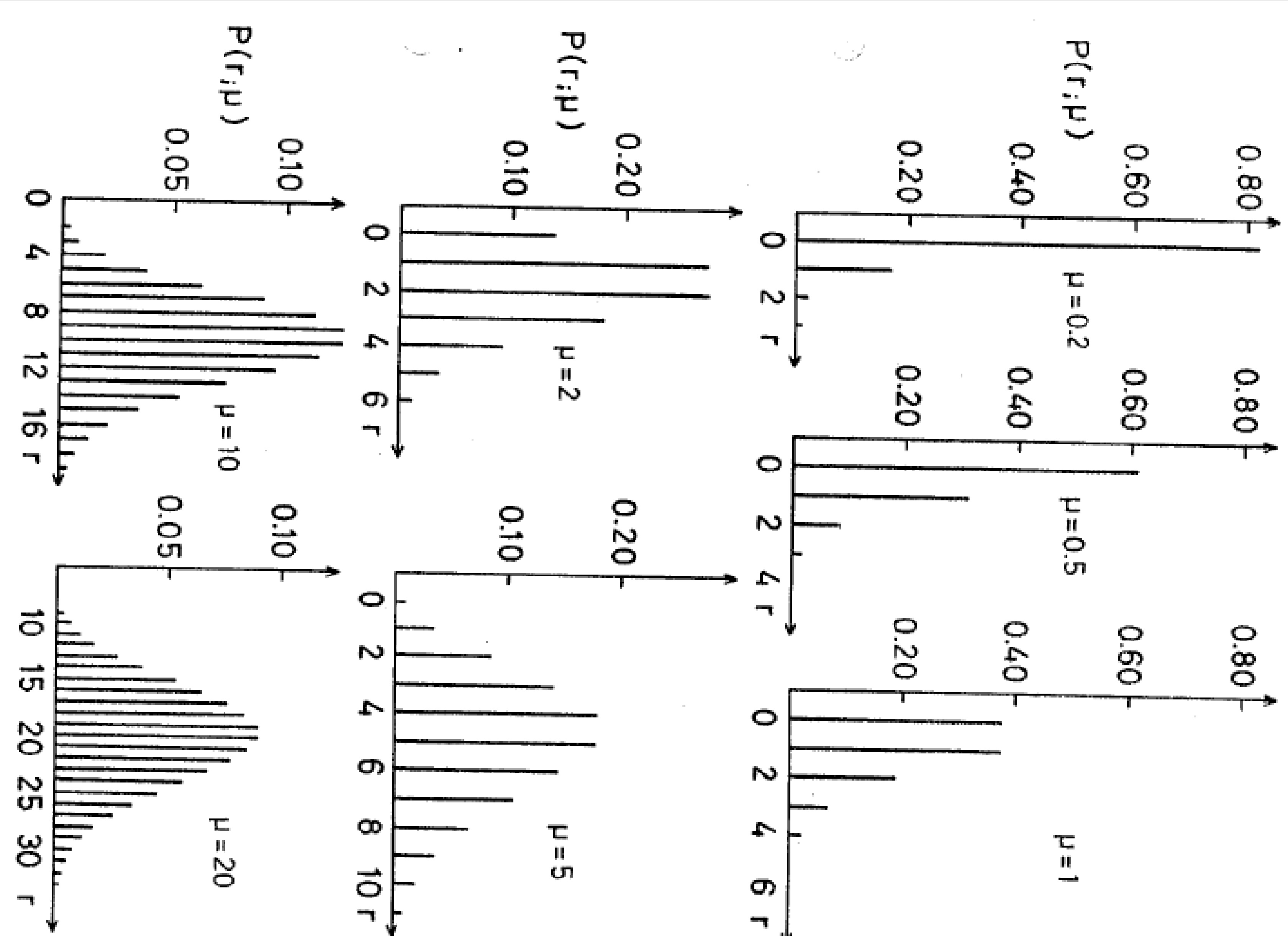


Fig. 4.3. The Poisson distribution for different mean values μ.

Frodesen et al.

Quantiles of Possion Distribution

Quantile λ_P der Poissonverteilung

$k \backslash P$	0.01	0.05	0.10	0.90	0.95	0.99
	$P = \sum_{n=0}^{k} \frac{\lambda_P^n}{n!} \exp(-\lambda_P)$					
0	4.605	2.996	2.303	0.105	0.051	0.010
1	6.638	4.744	3.890	0.532	0.355	0.149
2	8.406	6.296	5.322	1.102	0.818	0.436
3	10.045	7.754	6.681	1.745	1.366	0.823
4	11.605	9.154	7.994	2.433	1.970	1.279
5	13.108	10.513	9.275	3.152	2.613	1.785
6	14.571	11.842	10.532	3.895	3.285	2.330
7	16.000	13.148	11.771	4.656	3.981	2.906
8	17.403	14.435	12.995	5.432	4.695	3.507
9	18.783	15.705	14.206	6.221	5.425	4.130
10	20.145	16.962	15.407	7.021	6.169	4.771
11	21.490	18.208	16.598	7.829	6.924	5.428
12	22.821	19.443	17.782	8.646	7.690	6.099
13	24.139	20.669	18.958	9.470	8.464	6.782
14	25.446	21.886	20.128	10.300	9.246	7.477
15	26.743	23.097	21.292	11.135	10.036	8.181
16	28.030	24.301	22.452	11.976	10.832	8.895
17	29.310	25.499	23.606	12.822	11.634	9.616
18	30.581	26.692	24.756	13.671	12.442	10.346
19	31.845	27.879	25.903	14.525	13.255	11.082
20	33.103	29.062	27.045	15.383	14.072	11.825

Lutz Feld

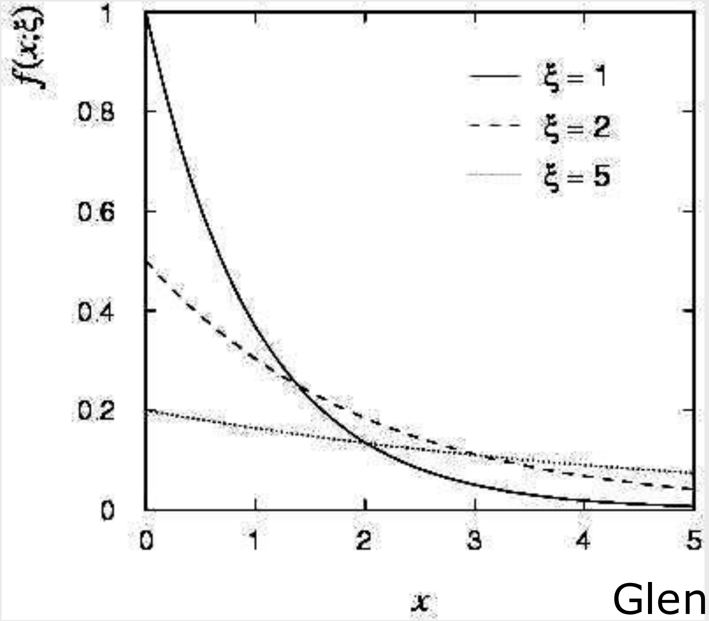
Tabelle 2.1: Quantile λ_P der Poissonverteilung. Zum Beispiel ist für $k = 0$ die Quantile $\lambda_{0.05} \approx 3$. Wenn null Ereignisse beobachtete werden, dann ist das mit (höchstens) 5% Wahrscheinlichkeit zu erwarten, wenn der Mittelwert (mindestens) $\lambda = 3$ ist.

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

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

$$E[x] = \xi$$

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



Glen Cowan

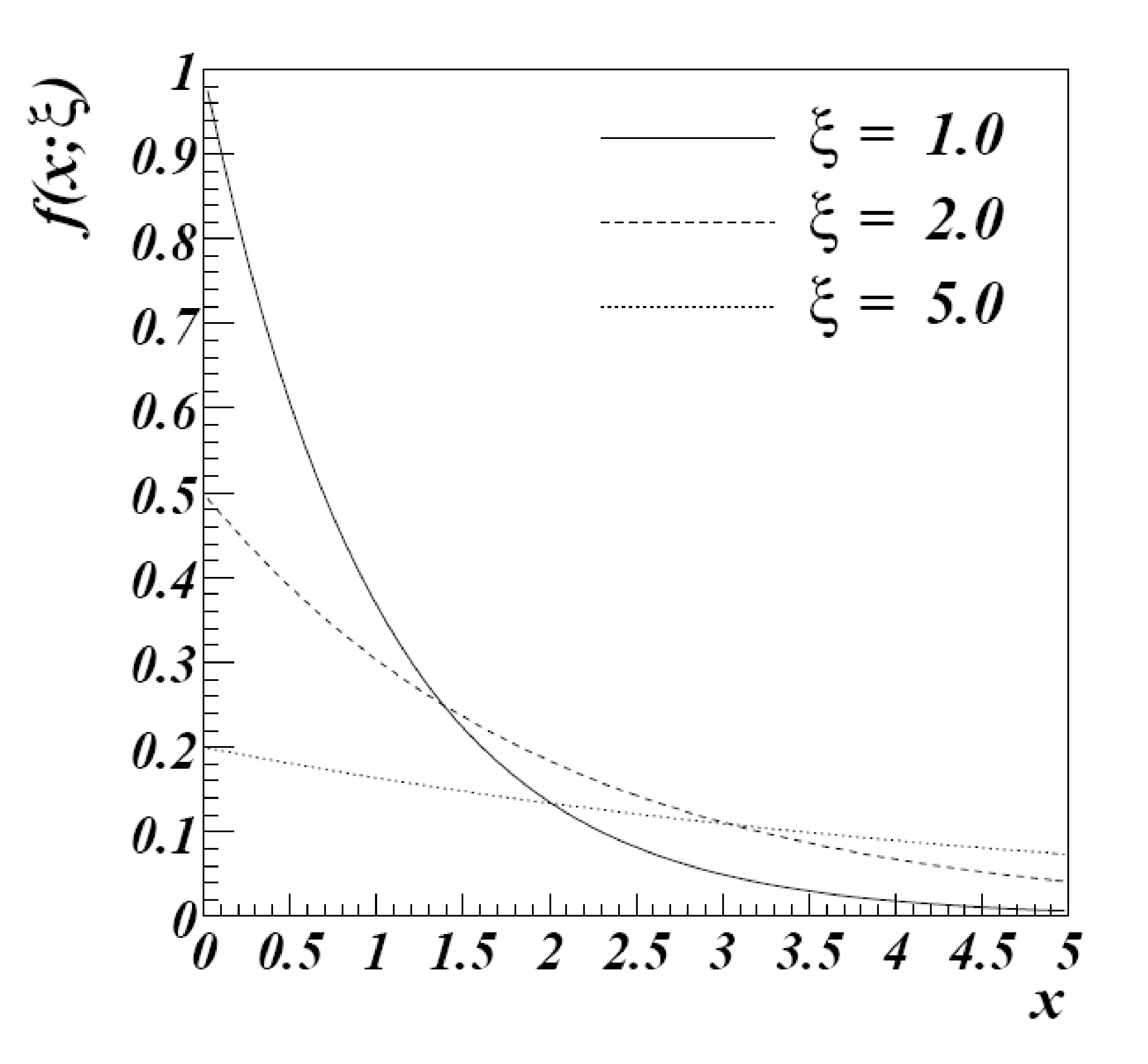
Example: proper decay time *t* of an unstable particle

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

(τ = mean lifetime)

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

Exponential Distribution



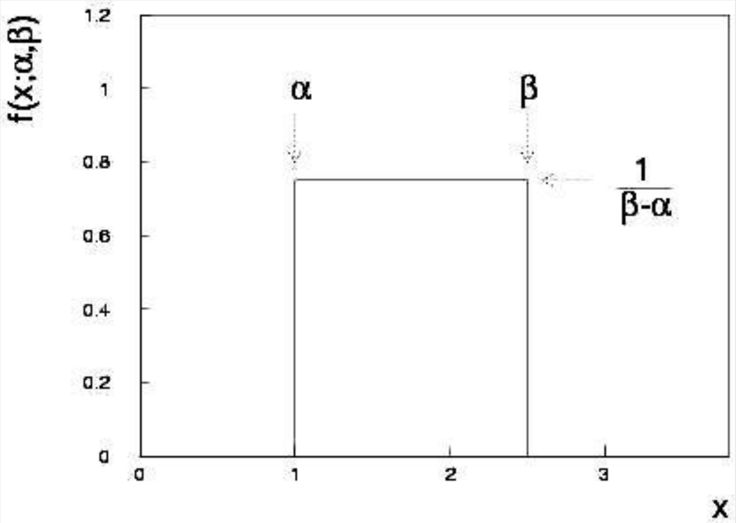
Glen Cowan

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].

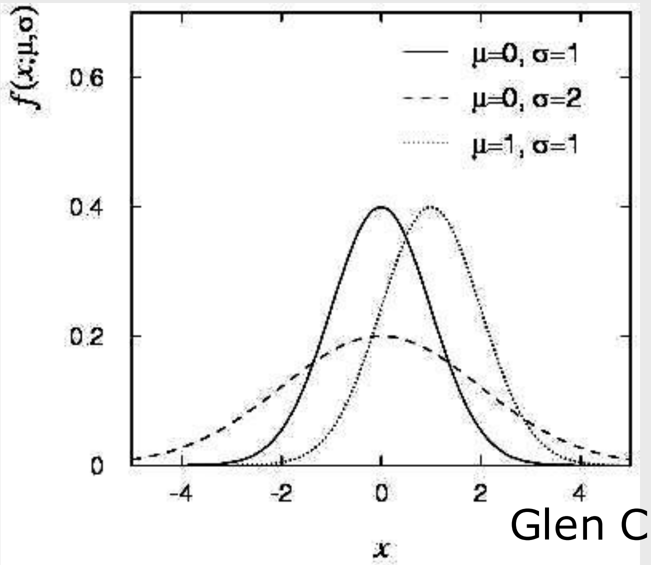
Example: for $\pi^0 \to \gamma\gamma$, E_γ is uniform in $[E_{\min}, E_{\max}]$, with

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

Glen Cowan

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

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



Glen Cowan

$$E[x] = \mu$$

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

(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’):

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

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

Normal / Gaussian Distribution

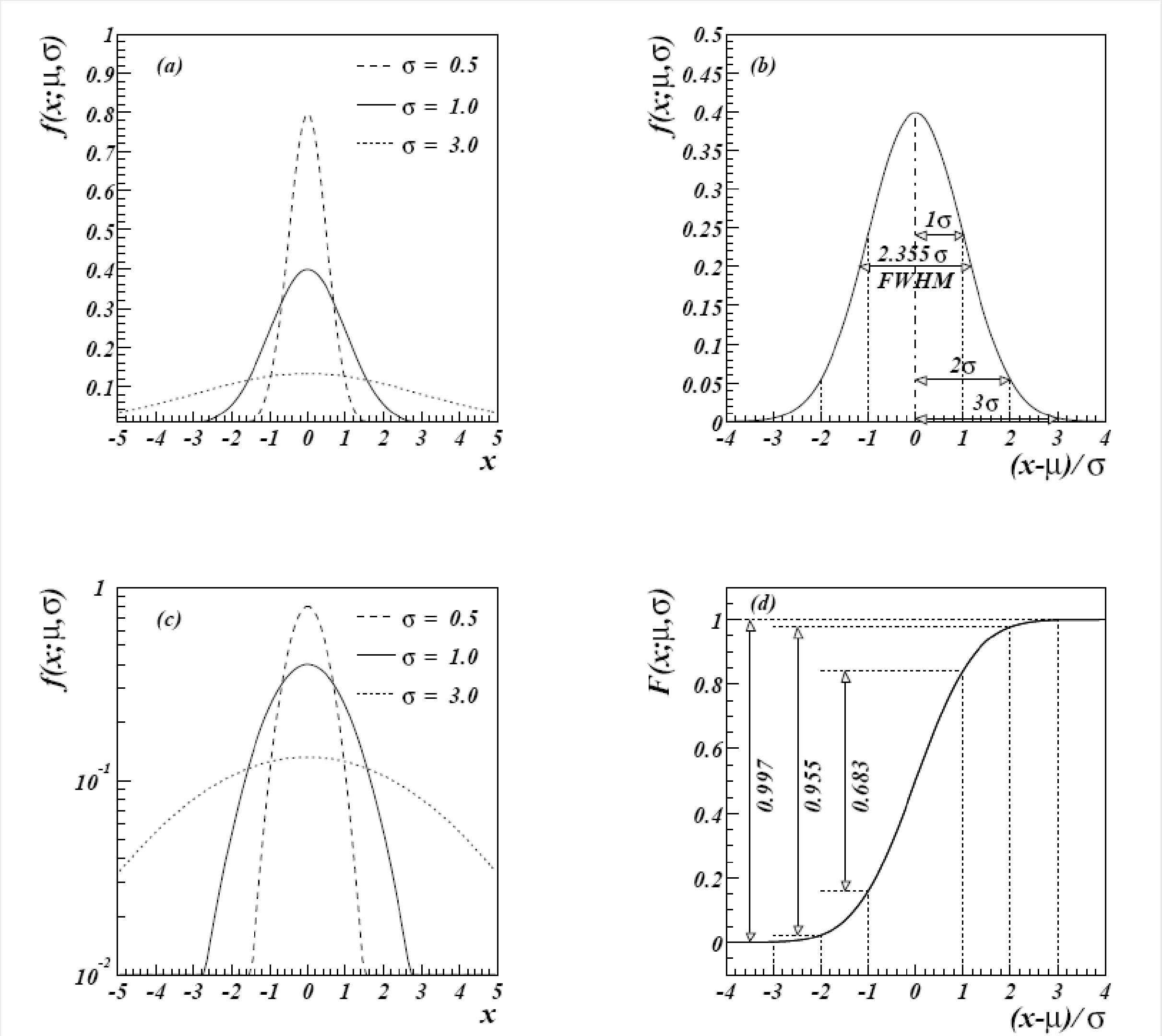
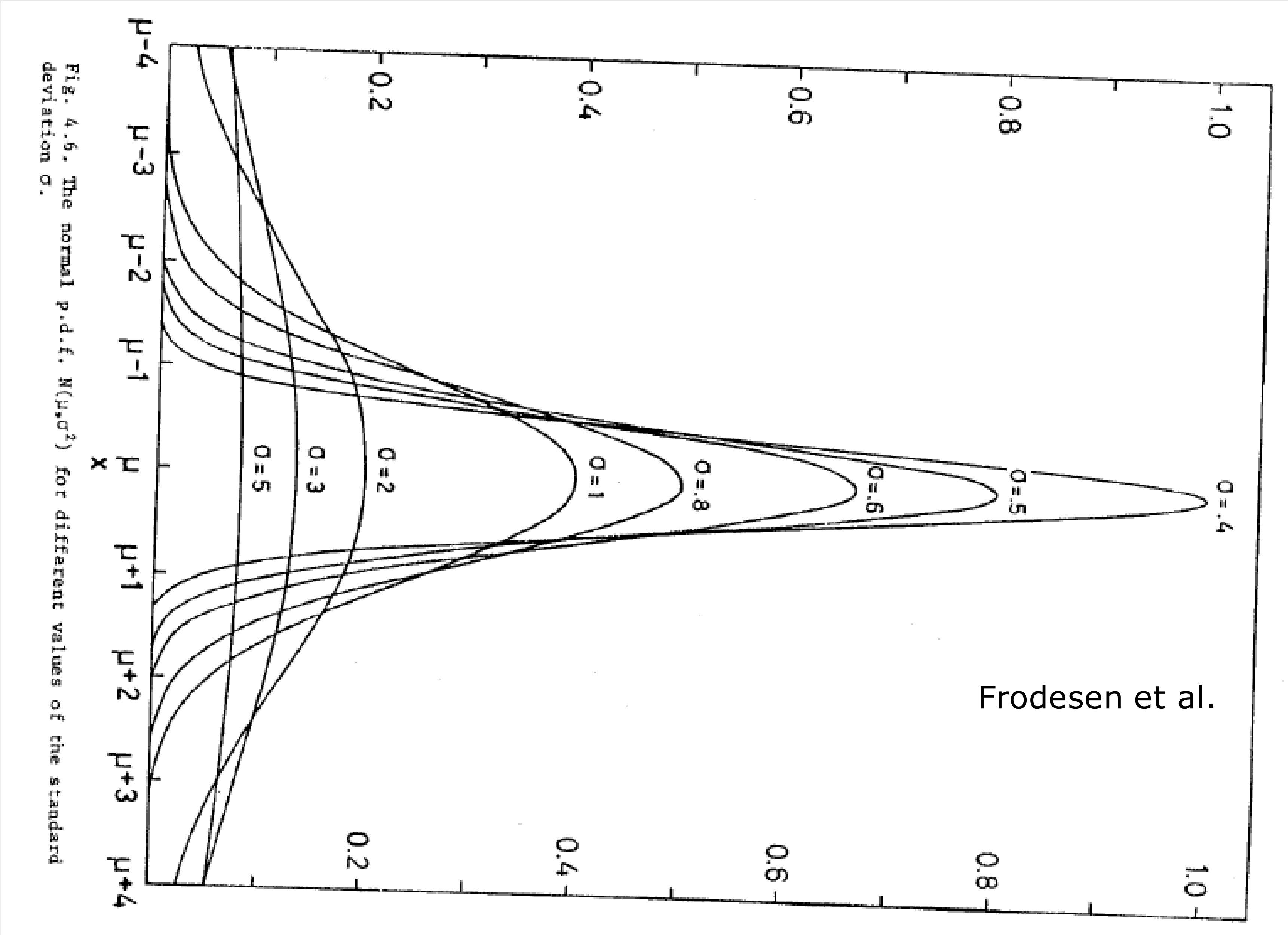
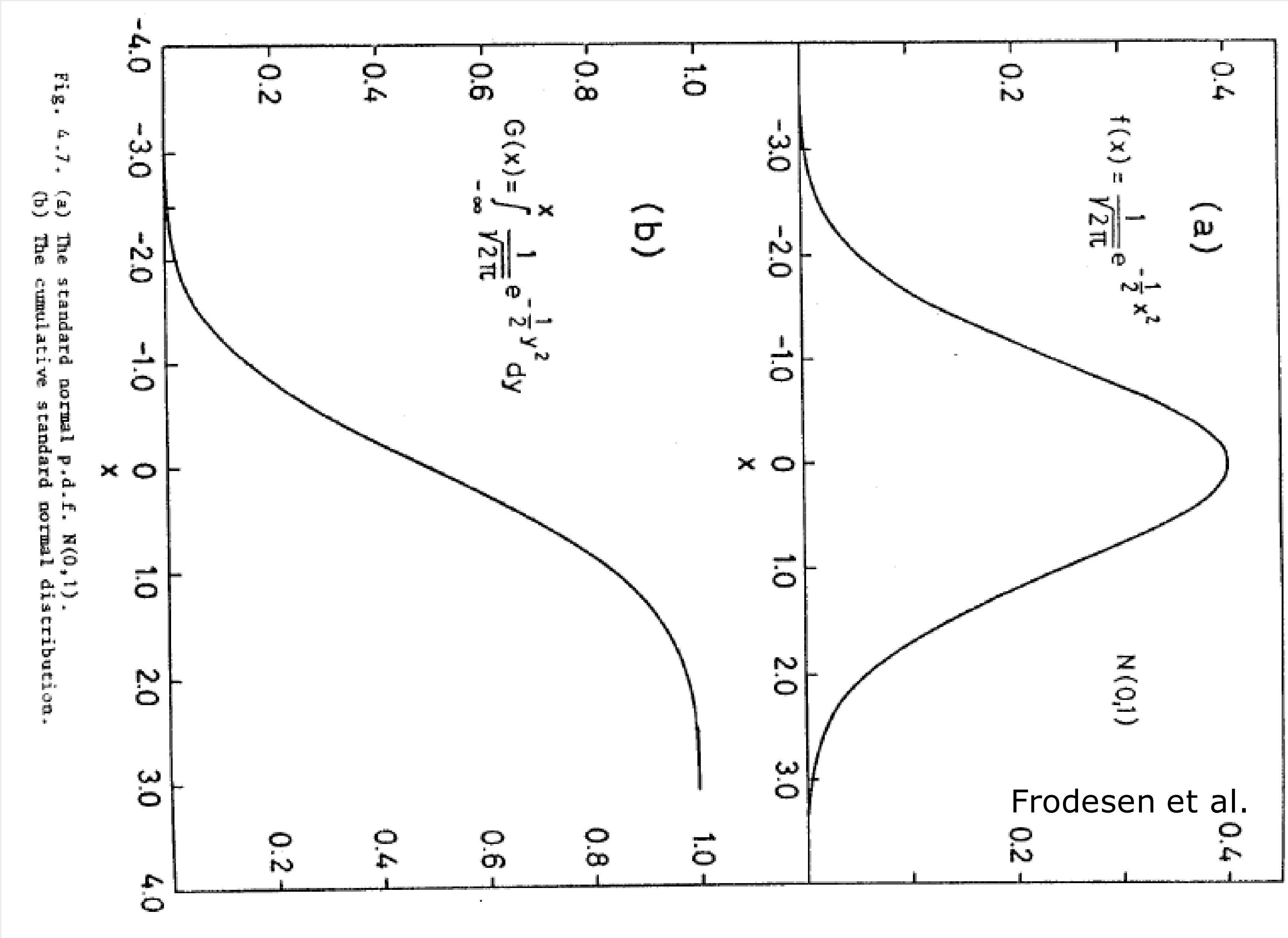


Abbildung 2.5: Gaußverteilungen für verschiedene Werte von σ und $\mu = 0$ mit linearer (a) bzw. logarithmischer (c) Ordinate. (b) zeigt die Normalverteilung, wobei für ein, zwei und drei σ Abweichung vom Mittelwert jeweils der Funktionswert gekennzeichnet ist. (d) zeigt die kumulative Wahrscheinlichkeitsverteilung. Die Werte des Integrals der Gaußverteilung innerhalb von ein, zwei und drei σ um den Mittelwert sind eingezeichnet.

Lutz Feld



0.2
0.4
0.6
0.8
1.0
μ-4
μ-3
μ-2
μ-1
μ
μ+1
μ+2
μ+3
μ+4
x
σ=.4
σ=.5
σ=.6
σ=.8
σ=1
σ=2
σ=3
σ=5
Fig. 4.6. The normal p.d.f. N(μ,σ²) for different values of the standard deviation σ.
Frodesen et al.

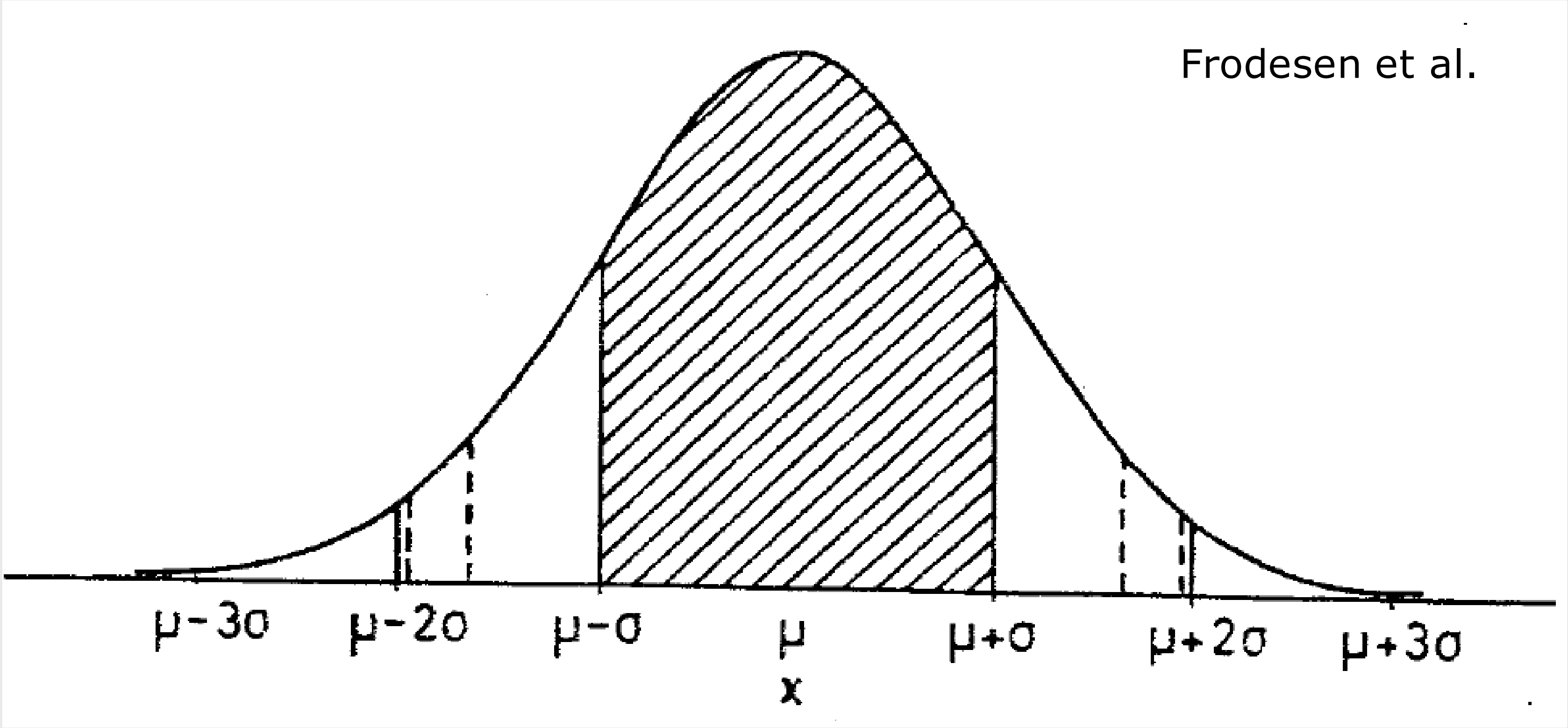

(a)
$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$
N(0,1)
(b)
$G(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$
x
Fig. 4.7. (a) The standard normal p.d.f. N(0,1).
(b) The cumulative standard normal distribution.
Frodesen et al.

$$P(\mu-\sigma \le x \le \mu+\sigma) = 2G(1) - 1 = 0.6827$$
$$P(\mu-2\sigma \le x \le \mu+2\sigma) = 2G(2) - 1 = 0.9549$$
$$P(\mu-3\sigma \le x \le \mu+3\sigma) = 2G(3) - 1 = 0.9973.$$

Frodesen et al.

$$P(\mu-1.645\sigma \le x \le \mu+1.645\sigma) = 0.90$$
$$P(\mu-1.960\sigma \le x \le \mu+1.960\sigma) = 0.95$$
$$P(\mu-2.576\sigma \le x \le \mu+2.576\sigma) = 0.99$$
$$P(\mu-3.290\sigma \le x \le \mu+3.290\sigma) = 0.999$$

Probability Contents



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.

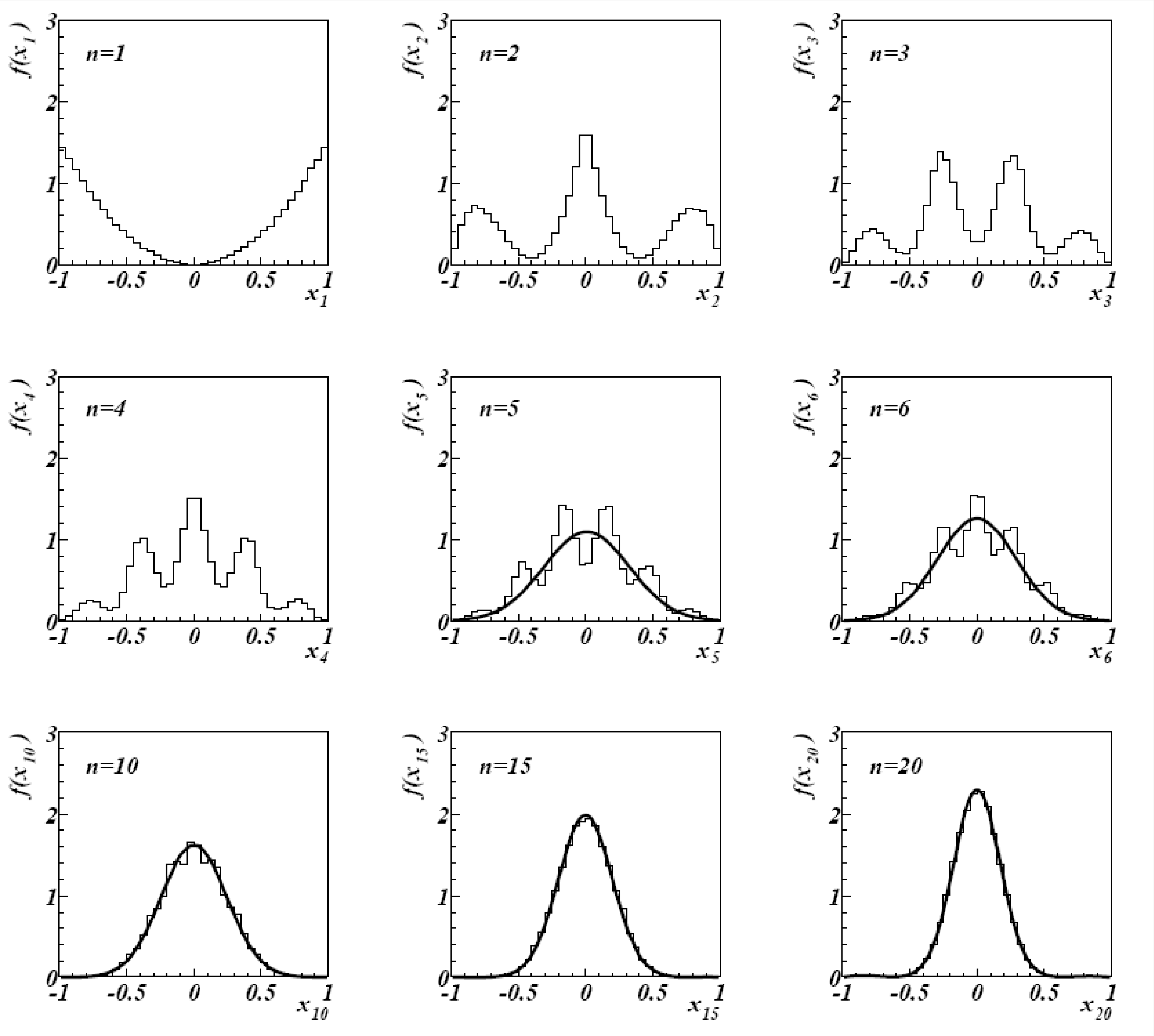
Beware of measurement errors with non-Gaussian tails.

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.)

The CLT at Work



Lutz Feld

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,

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

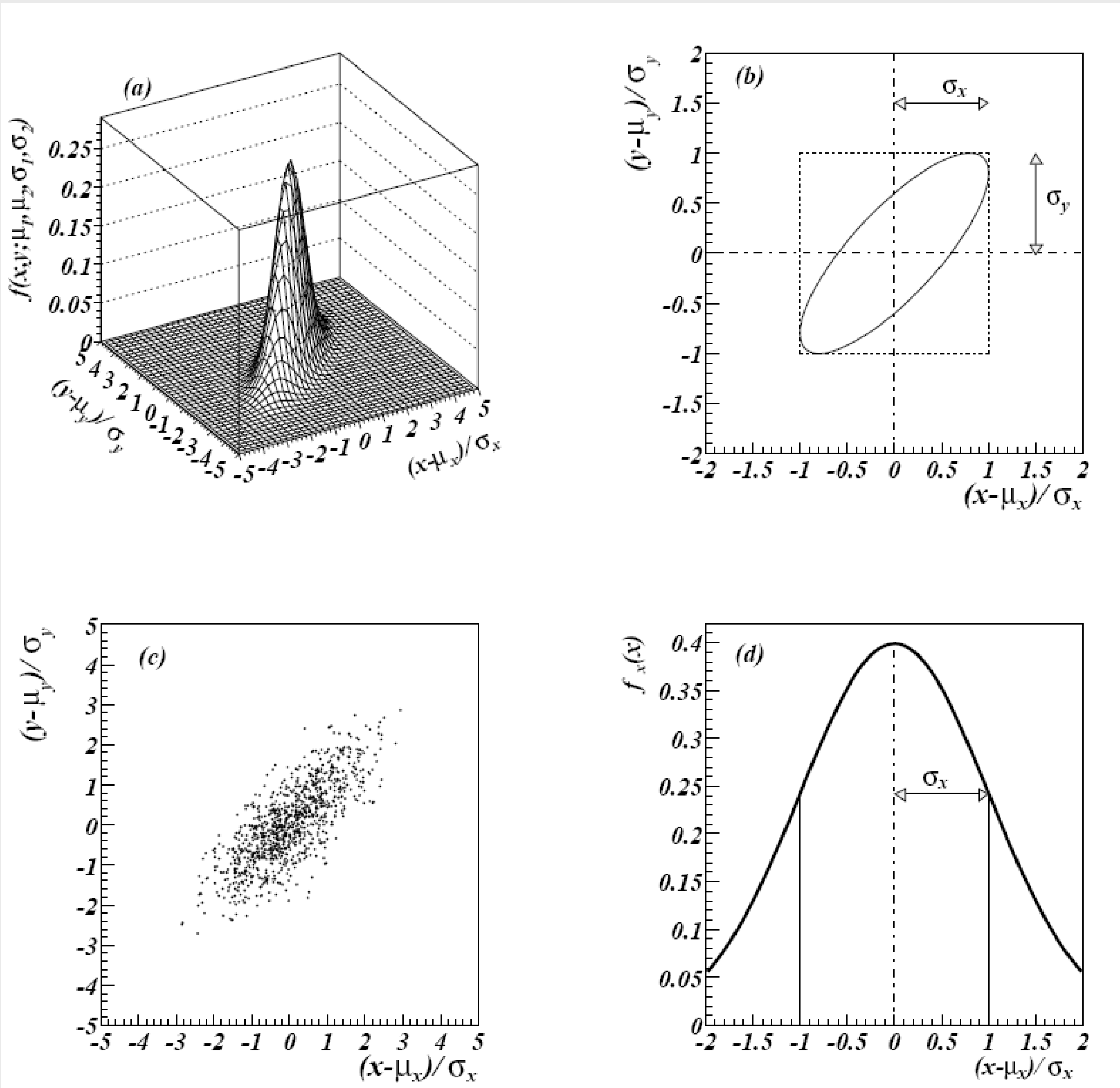
For $n = 2$ this is

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

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

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

Two Dimensional Gaussian



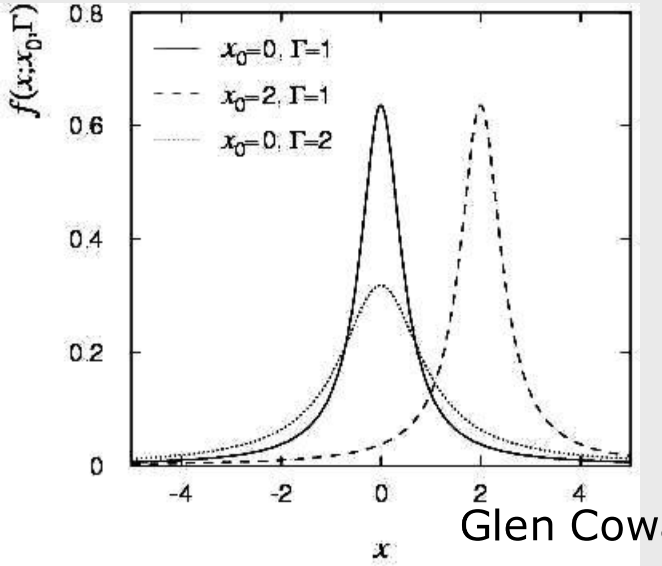
Lutz Feld

Abbildung 2.6: Zweidimensionale Gaußverteilung (a) und eine mögliche Verteilung von Meßpunkten (c) für einen Korrelationskoeffizienten von $\rho = 0.8$. (b) zeigt die 1σ Ellipse, (d) die Randverteilung in x.

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.)



Glen Cowan

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

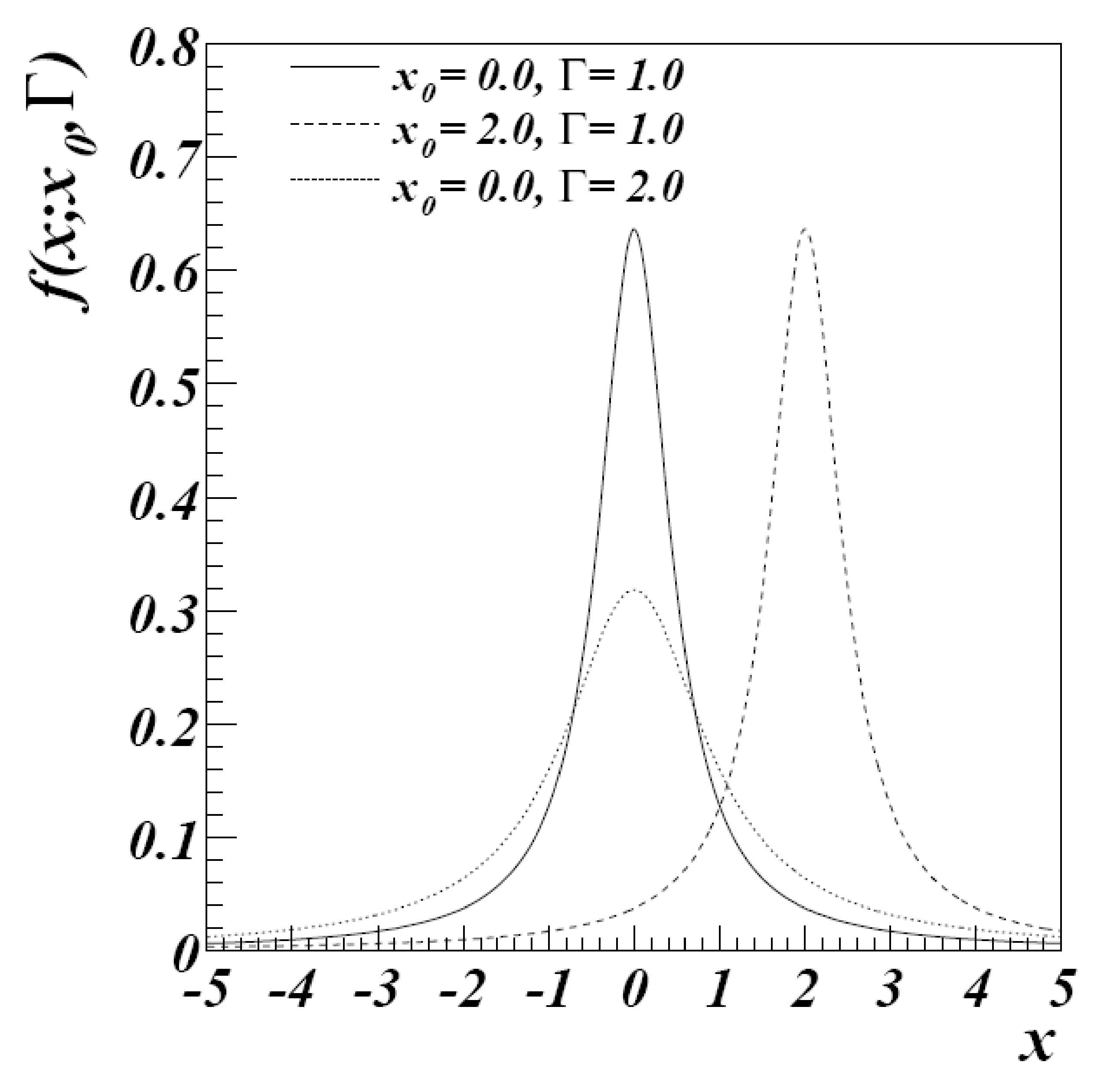
x_0 = mode (most probable value)

Γ = full width at half maximum

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

Γ = decay rate (inverse of mean lifetime)

Cauchy / Breit Wigner Distribution



Lutz Feld

Distribution	Fraction of distribution in tail				
|x|	≥ 1	≥ 2	≥ 3	≥ 4	≥ 6
Standard normal	.3173	.0455	.0027	.00006	–
Double exponential	.3679	.1353	.0498	.0183	.0025
Cauchy	.5000	.2952	.2048	.1560	.1051

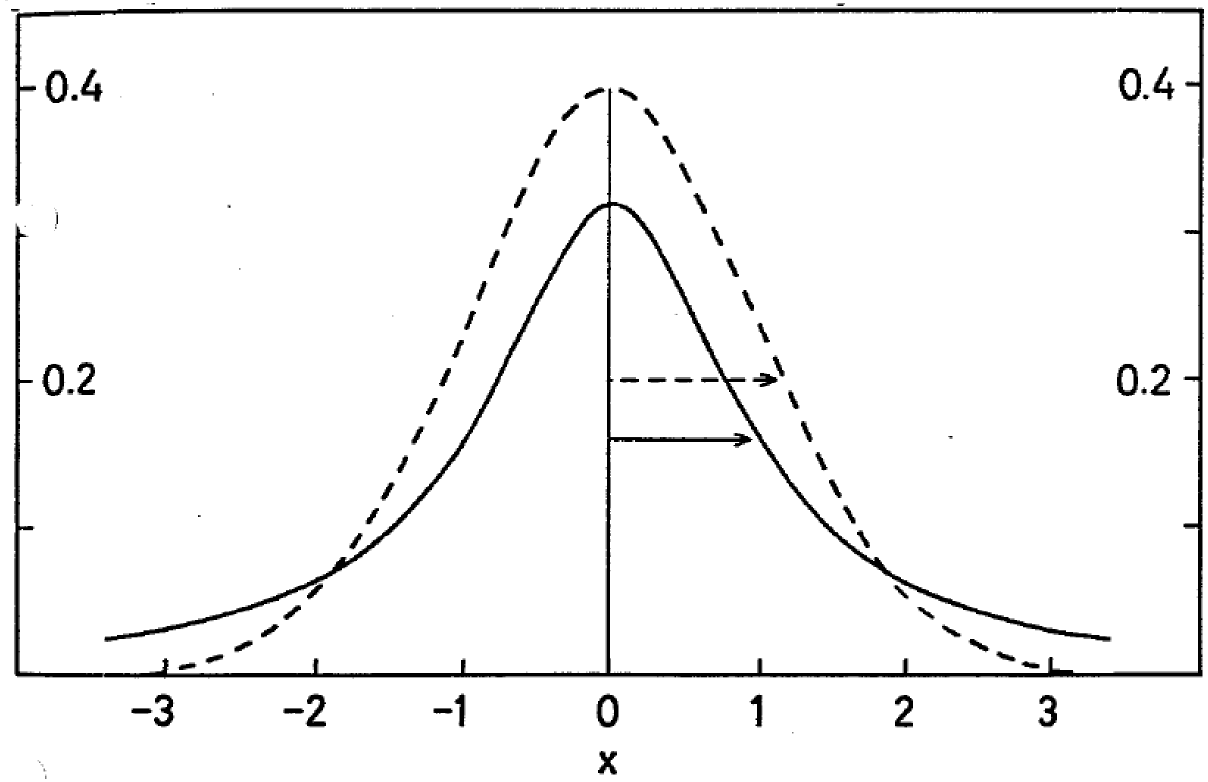
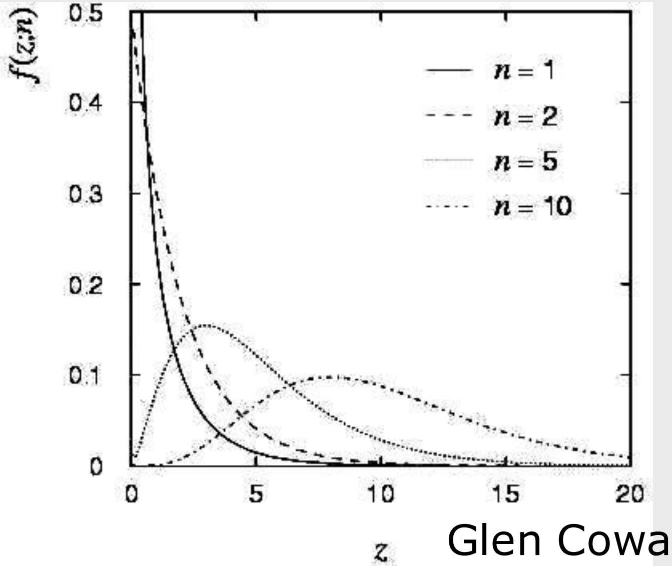


Fig. 4.9. The Cauchy or Breit-Wigner distribution (solid curve) and the standard normal distribution (dashed curve). The half-widths at half-maximum are indicated by arrows of length 1 and $\sqrt{2\ln 2} \simeq 1.18$, respectively.

Frodesen et al.

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

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



Glen Cowan

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

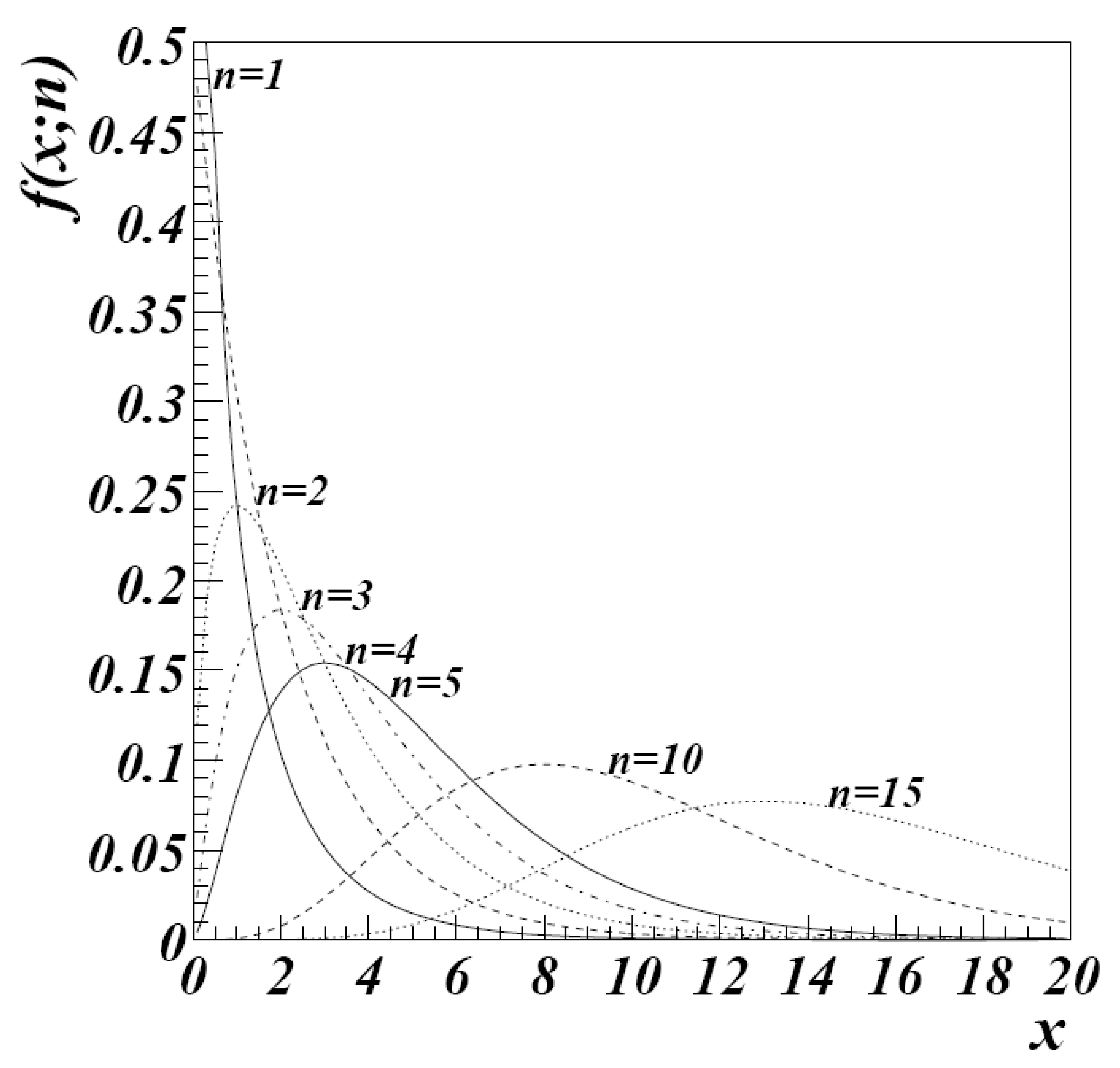
$$E[z] = n\,, \qquad V[z] = 2n\,.$$

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.


f(x;n)
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
n=1
n=2
n=3
n=4
n=5
n=10
n=15
0
2
4
6
8
10
12
14
16
18
20
x

Lutz Feld

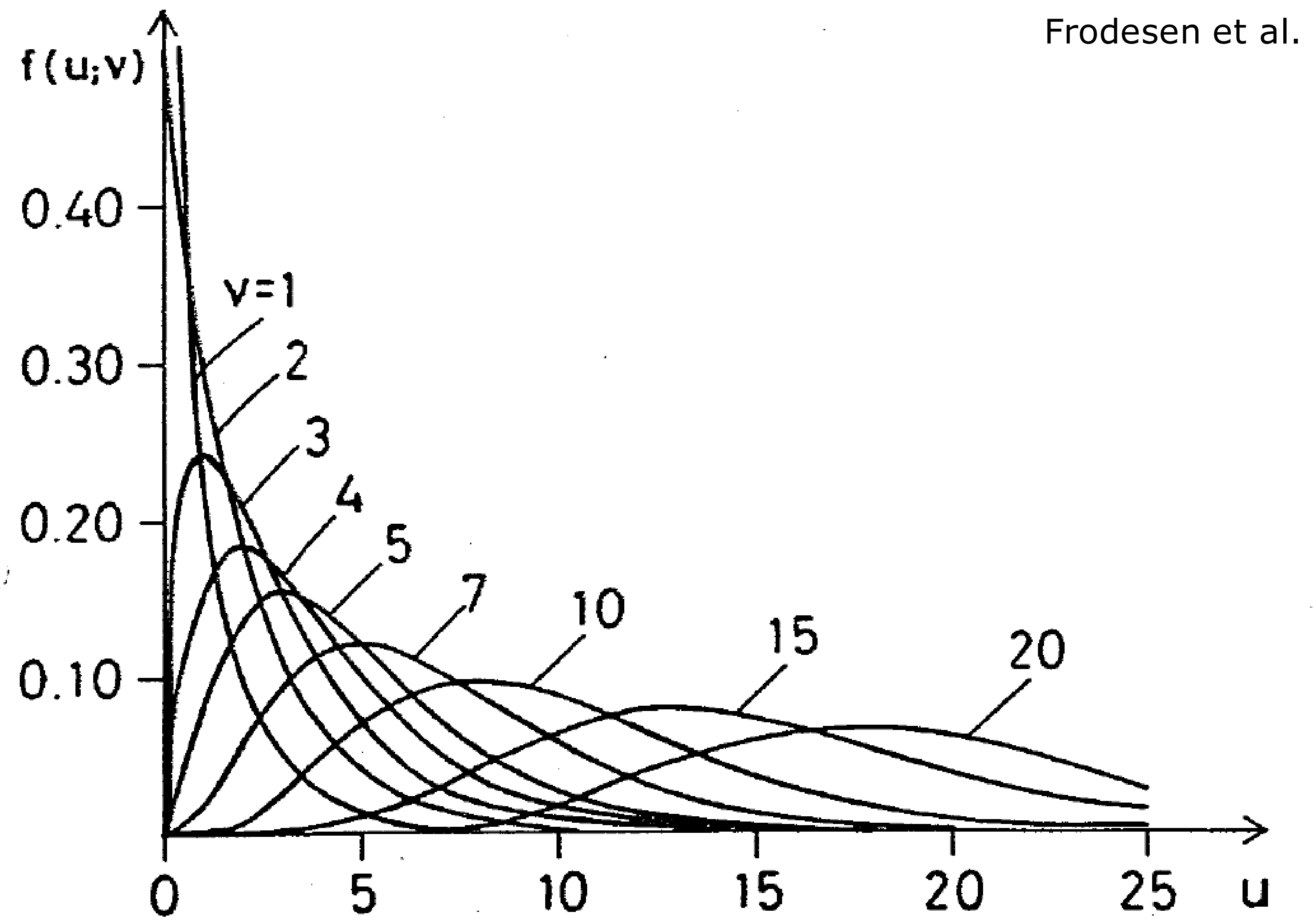
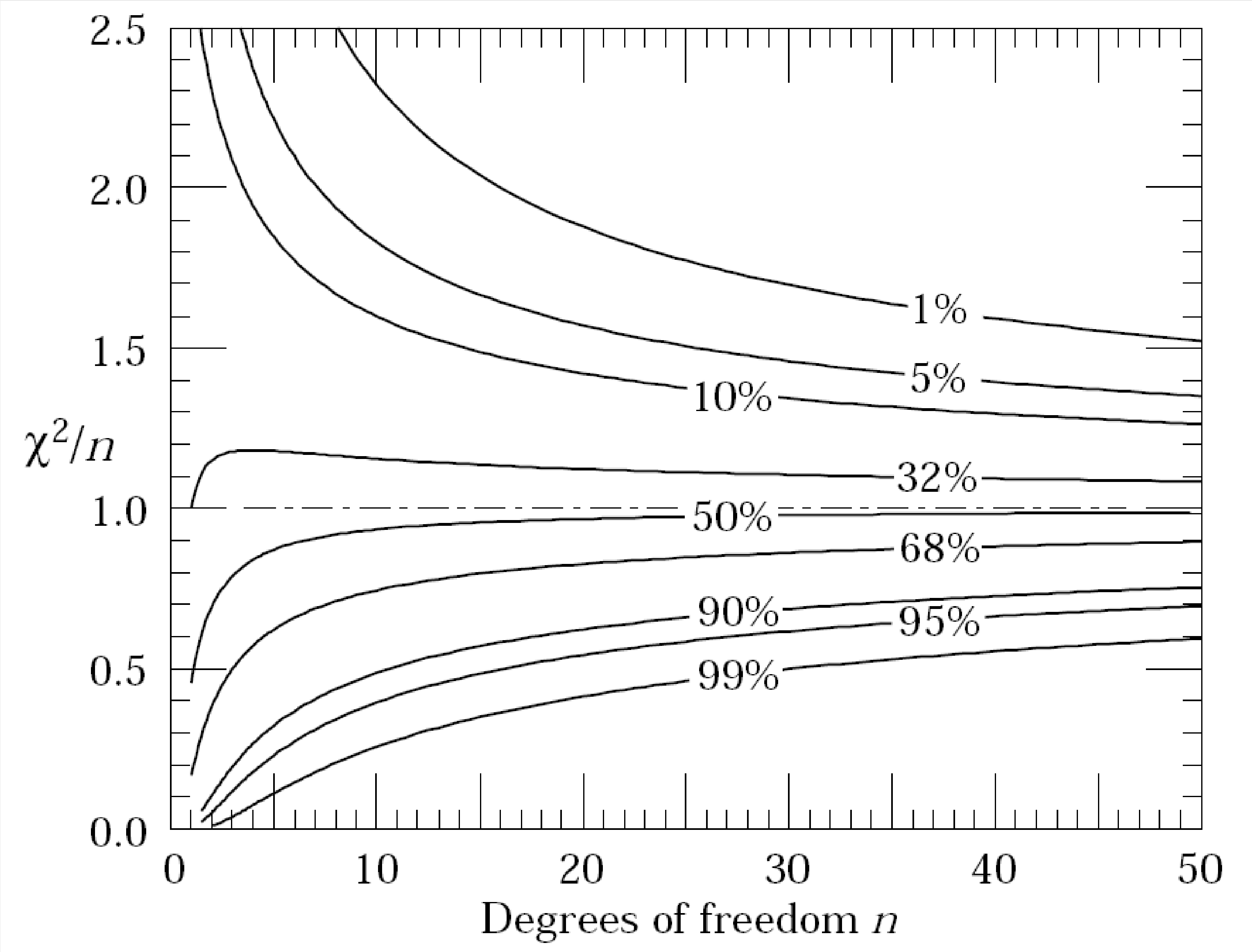


Fig. 5.1. The chi-square distribution for different degrees of freedom ν.

Quantiles of Chi Square Distribution



Lutz Feld

Quantiles of chi square distribution

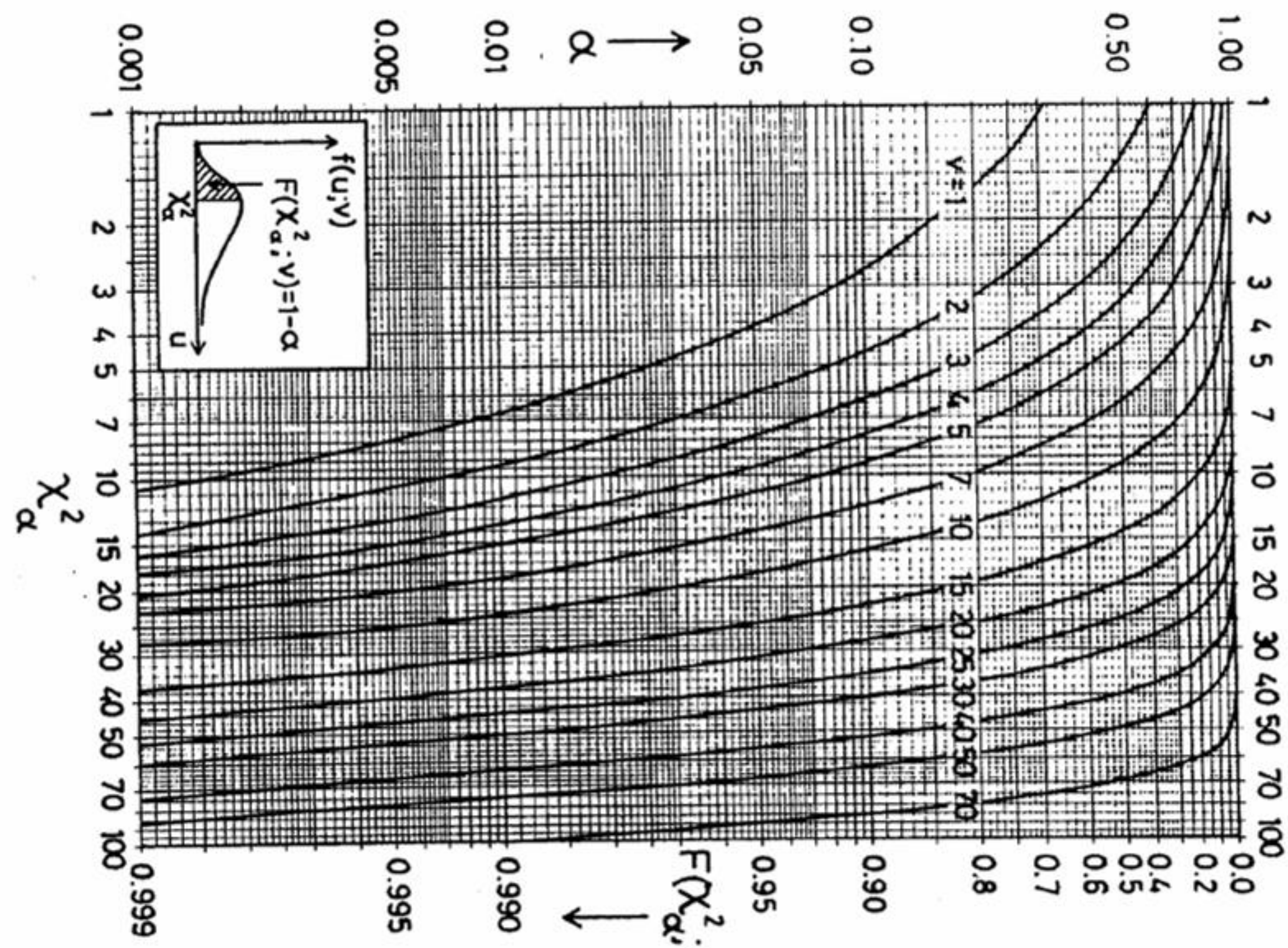


Fig. 5.2. Probability contents of the chi-square distribution.

Frodesen et al.

Gamma distribution

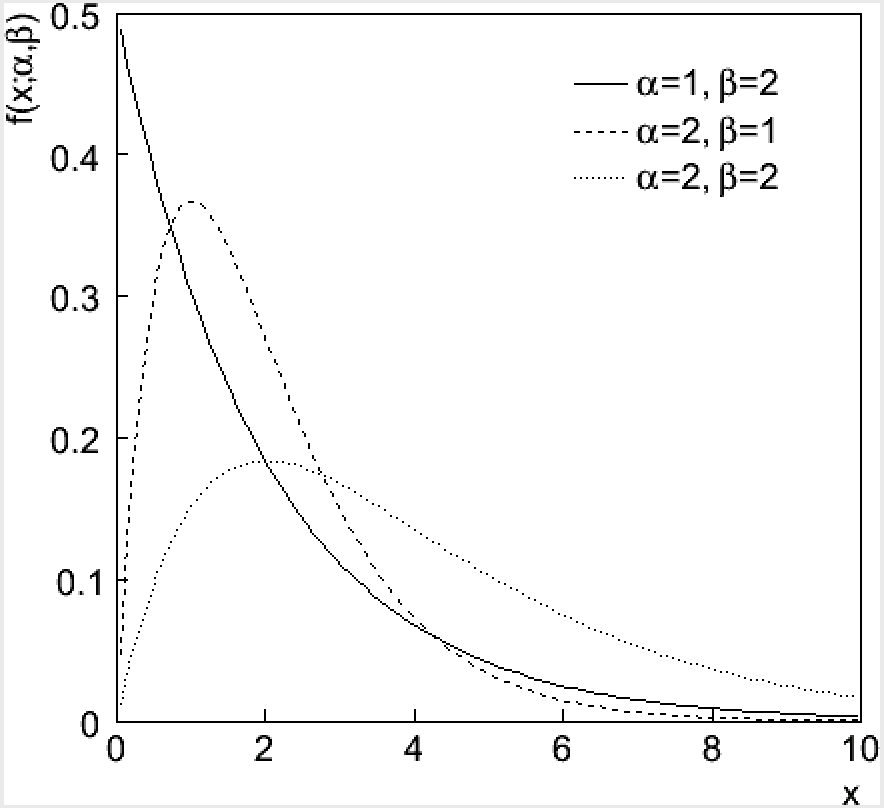
$$f(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}$$

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

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

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

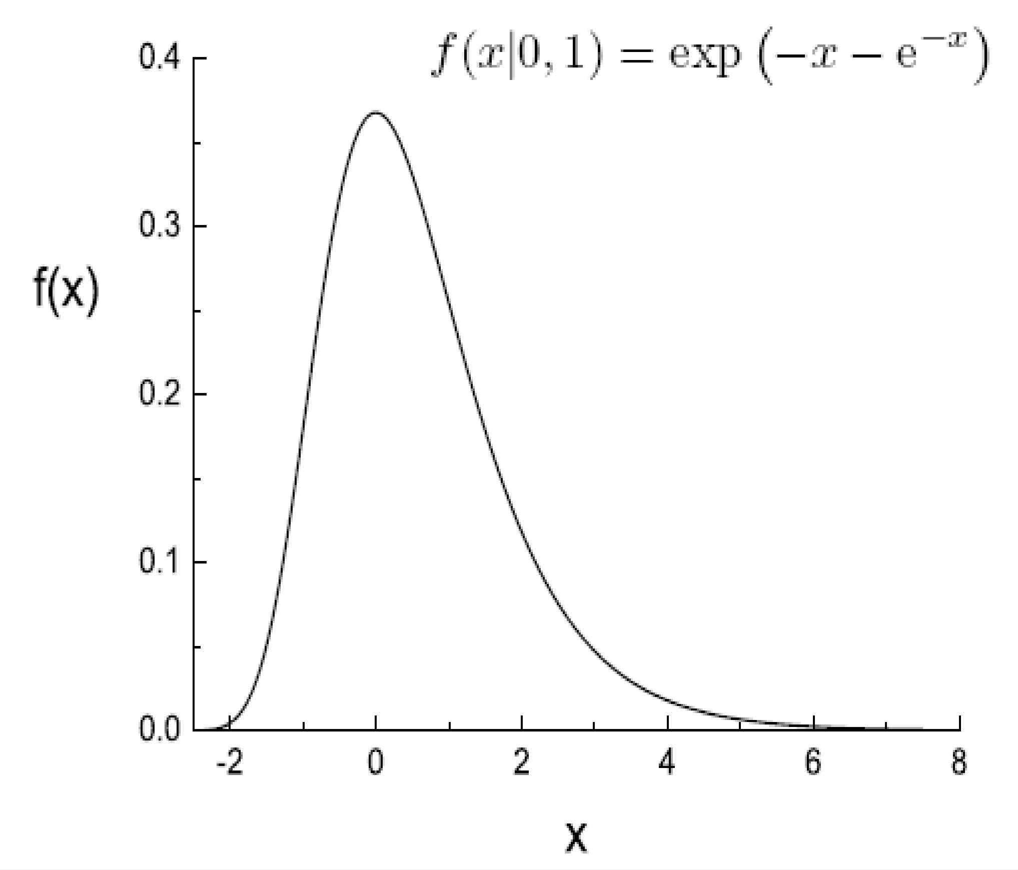


Log-Weibull-Distribution

$$f(x|x_0, s) = \frac{1}{s} \exp\left(\pm \frac{x - x_0}{s} - e^{\pm \frac{x - x_0}{s}} \right)$$

$$\mu = x_0 \mp Cs \qquad C = 0.5772$$

$$\sigma^2 = s^2 \frac{\pi^2}{6}$$



Distribution of Minimum (+) and Maximum (-) from a Sample of n RVs (n large) following a Gaussian, Poisson, Exponential and related distributions.

Landau distribution

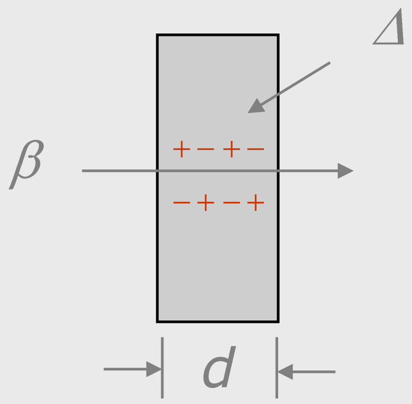
For a charged particle with $\beta = v/c$ traversing a layer of matter of thickness d, the energy loss Δ follows the Landau pdf:

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

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

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

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

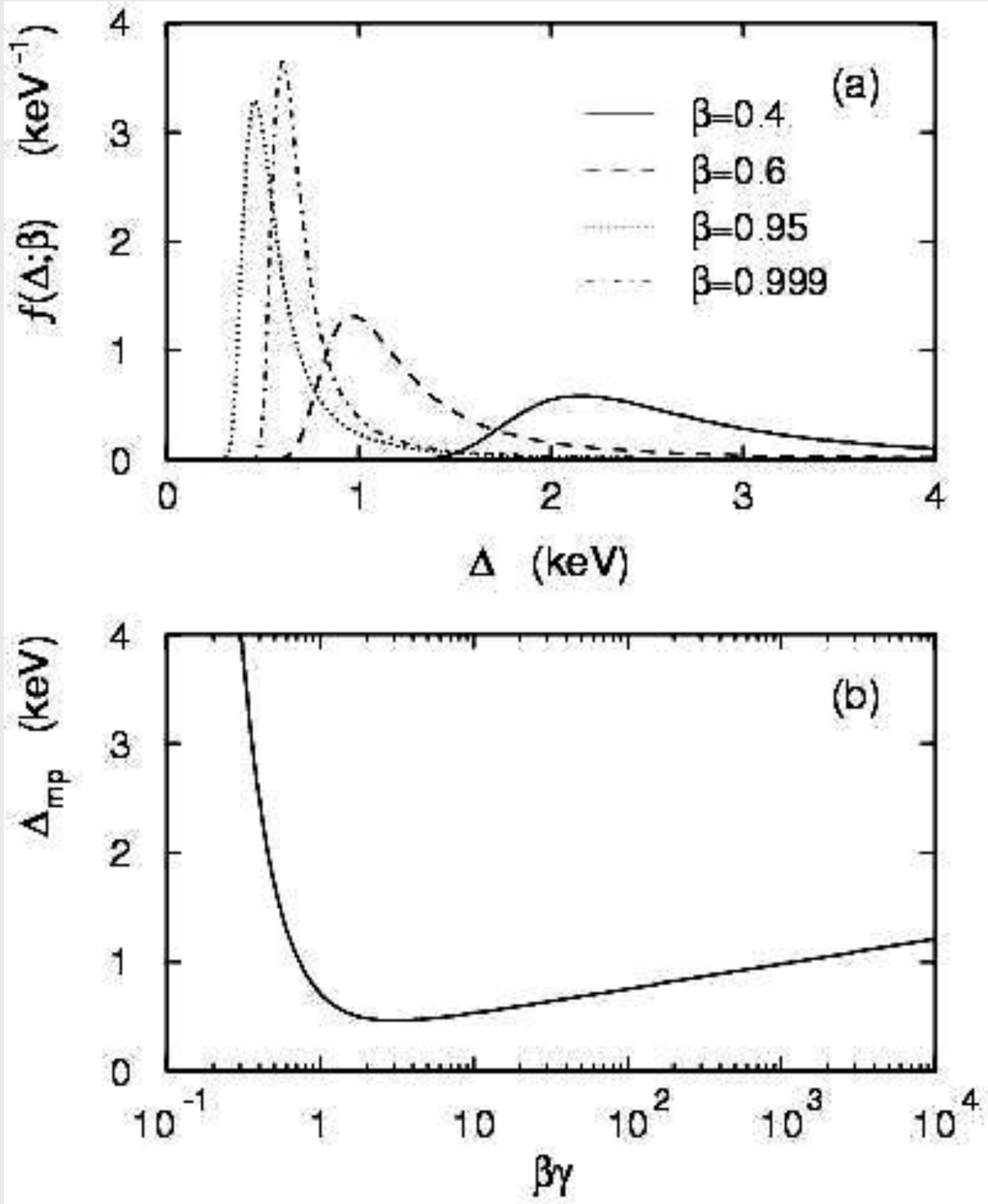


L. Landau, J. Phys. USSR **8** (1944) 201; see also
W. Allison and J. Cobb, Ann. Rev. Nucl. Part. Sci. **30** (1980) 253.

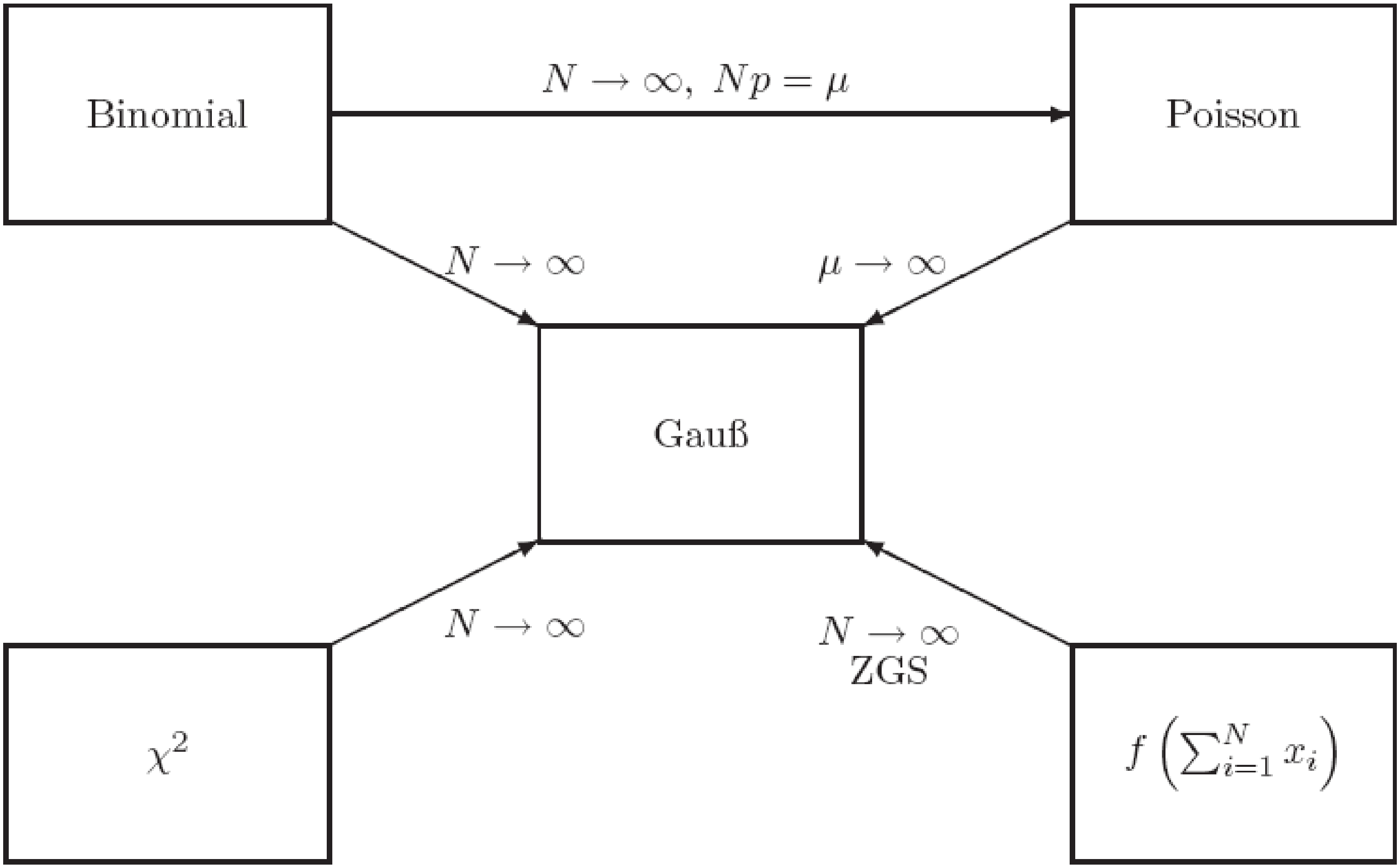
Landau distribution (2)

Long 'Landau tail'
→ all moments ∞

Mode (most probable value) sensitive to β,
→ particle i.d.



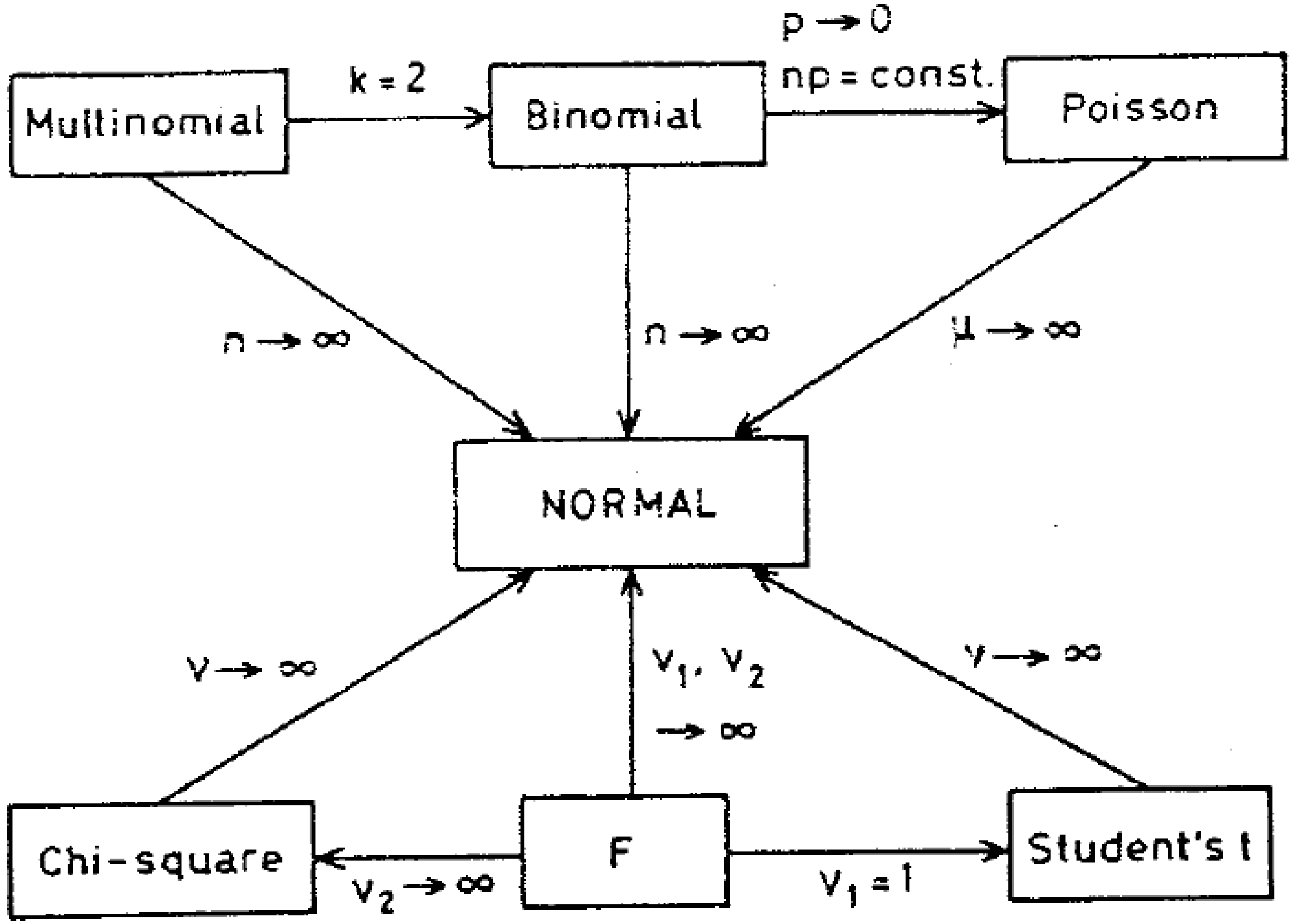
Connection between PDFs



Lutz Feld

Survey of PDFS

Name	Definitionsbereich	Wahrscheinlichkeits(dichte)	$\langle x\rangle = E[x]$	$\sigma_x = \sqrt{V[x]}$	charakter. Funktion
Binomial	$\{0,1,2,3,...,N\}$	$f(x;p,N) = \frac{N!}{x!(N-x)!}p^x(1-p)^{N-x}$	Np	$\sqrt{Np(1-p)}$	$(p(\exp(ik)-1)+1)^N$
Poisson	$\{0,1,2,3,...\}$	$f(x;\lambda) = \frac{\lambda^x}{x!}\exp(-\lambda)$	λ	$\sqrt{\lambda}$	$\exp(\lambda(\exp(ik)-1))$
Gleichvert.	$[a,b]$	$f(x;a,b) = \frac{1}{(b-a)}$	$\frac{a+b}{2}$	$\frac{b-a}{\sqrt{12}}$	$\frac{\exp(ibk)-\exp(iak)}{(b-a)ik}$
Gauß	$[-\infty,\infty]$	$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	μ	σ	$\exp\left(i\mu k - \frac{1}{2}\sigma^2k^2\right)$
Chi-Quadrat	$[0,\infty]$	$f(x;n) = \frac{1}{2^{n/2}\Gamma(n/2)}x^{\frac{n}{2}-1}\exp\left(-\frac{x}{2}\right)$	n	$2n$	$(1-2ik)^{-n/2}$
Exponentiell	$[0,\infty]$	$f(x;\xi) = \frac{1}{\xi}\exp(-x/\xi)$	ξ	ξ	$\frac{1}{1-ik\xi}$
Cauchy	$[-\infty,\infty]$	$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$	(0)	∞	$\exp(-\lvert k\rvert)$
Landau	$[0,\infty]$	$f(x;\beta) = ...$	∞	∞	



Frodesen et al.

Fig. 5.5. Relations between probability distributions.

More distributions

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

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

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

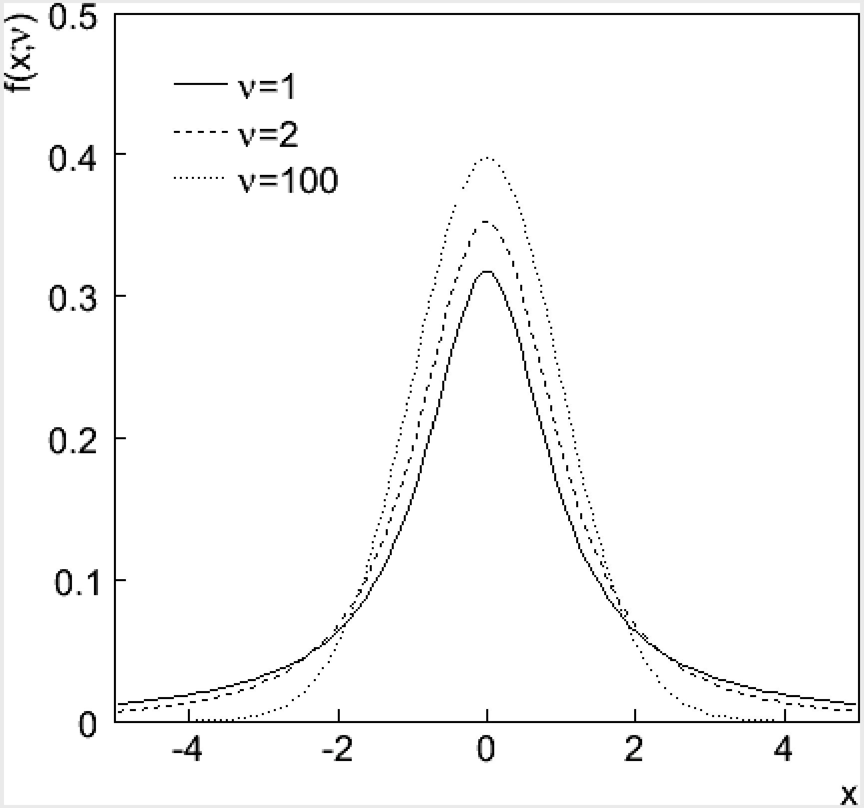
$$E[x] = 0 \quad (\nu > 1)$$

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

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

$\nu = 1$ gives Cauchy,

$\nu \to \infty$ gives Gaussian.



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

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

$t = \mathrm{x} / (z/n)^{1/2}$ 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.

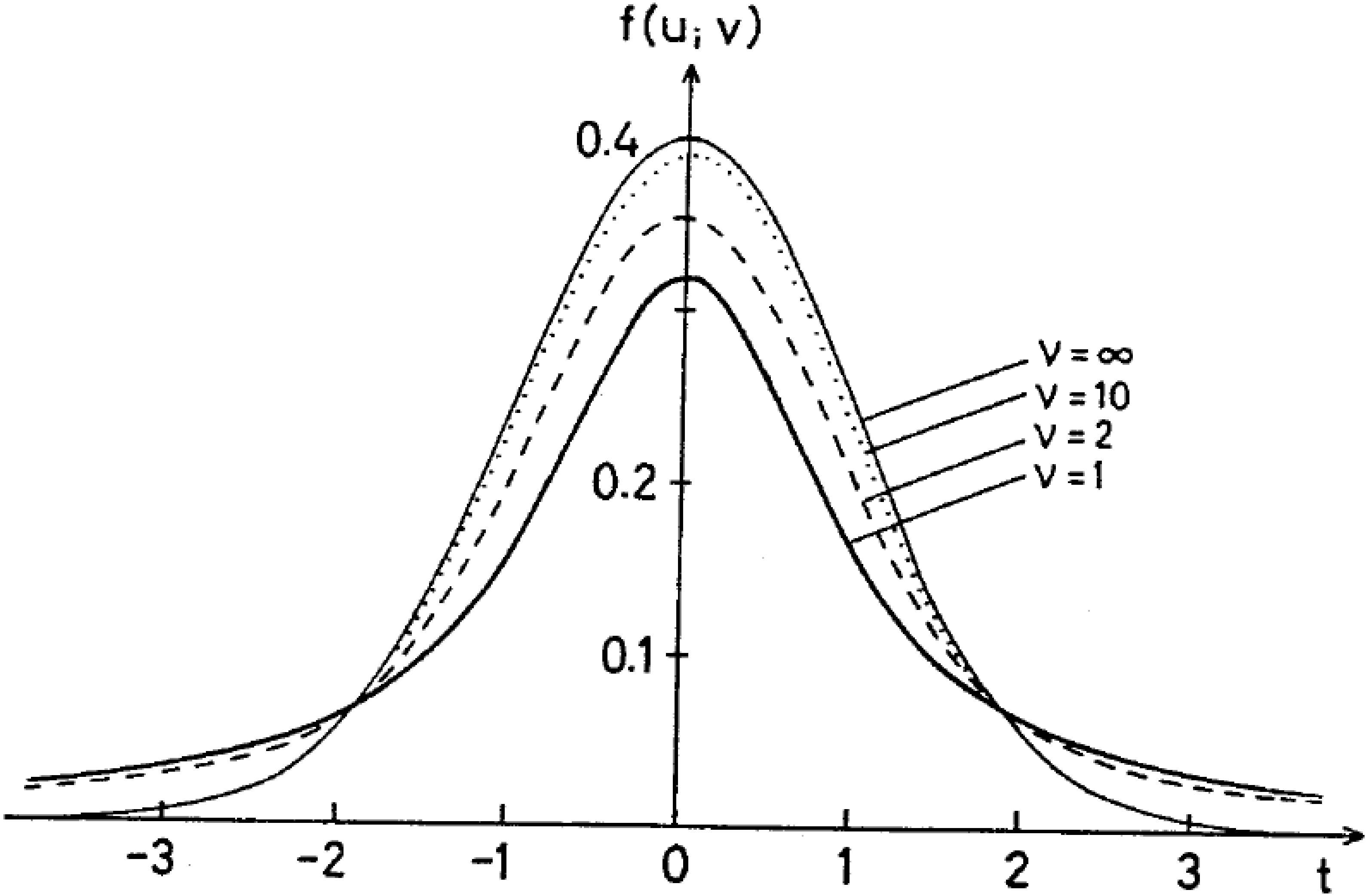


Fig. 5.3. The Student's t-distribution for different degrees of freedom ν. The special cases ν=1 and ν=∞ correspond to the Cauchy (Breit-Wigner) and the standard normal distributions, respectively.

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.