

# Statistische Methoden der Datenanalyse

## Wintersemester 2011/2012

### Albert-Ludwigs-Universität Freiburg



Prof. Markus Schumacher

Physikalisches Institut Westbau 2 OG Raum 008

Telefonnummer 07621 203 7612

E-Mail: [Markus.Schumacher@physik.uni-freiburg.de](mailto:Markus.Schumacher@physik.uni-freiburg.de)

Vorlesung basiert in weiten Teilen auf der von Glen Cowan.

[http://terascale.physik.uni-freiburg.de/lehre/ws\\_1112/statmethoden\\_ws1112](http://terascale.physik.uni-freiburg.de/lehre/ws_1112/statmethoden_ws1112)

## **Kapitel 2**

# **Katalog von ausgewählten Wahrscheinlichkeitsdichtefunktionen (Fortsetzung)**

# Multinomialverteilung

Wie Binomial aber nun  $m$  verschiedene Ausgänge des anstelle von zwei,  
Mit Wahrscheinlichkeiten für die einzelnen Ausgänge:

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

Für  $N$  Versuche suchen wir die Wahrscheinlichkeit das folgende Ergebnis zu erhalten:

$$\begin{aligned} n_1 & \text{ von Möglichkeit 1,} \\ n_2 & \text{ von Möglichkeit 2,} \\ & \dots \\ n_m & \text{ von Möglichkeit } m. \end{aligned}$$

Die Wktdichtefunktion ist die Multinomialverteilung für  $\vec{n} = (n_1, \dots, n_m)$

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

# Multinomialverteilung II

Nun betrachte Möglichkeit  $i$  as ‘Erfolg’, alle anderen als “Misserfolg”.

→ alle individuellen  $n_i$  binomialverteilt mit Parametern  $N, p_i$

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

Für die Kovarianzen ( $i$  ungleich  $j$ ) ergibt sich:

$$\begin{aligned} \text{cov}[n_i, n_j] &= E[(n_i - E[n_i])(n_j - E[n_j])] \\ &= E[n_i n_j] - E[n_i]E[n_j] \\ &= N(N - 1)p_i p_j - (Np_i)(Np_j) \\ &= -Np_i p_j. \end{aligned}$$

Negative Korrelation:  
wenn in einer Klasse mehr  
Ereignisse müssen  
irgendwo welche fehlen

$$\rho_{n_i n_j} = \frac{\text{cov}[n_i, n_j]}{\sigma_{n_i} \sigma_{n_j}} = -\sqrt{\frac{p_i p_j}{(1 - p_i)(1 - p_j)}}$$

# Multinomialverteilung (III): Beispiele

$\vec{n} = (n_1, \dots, n_m)$  repräsentiert ein Histogramm

mit  $m$  bins,  $N$  (bekannt) Gesamteinträgen

Zerfall eines instabilen Teilchens A: (1)  $A \rightarrow BC$ , (2)  $A \rightarrow DE$ , (3)  $A \rightarrow FG$

Mit bekannten Zerfallswahrscheinlichkeiten für alle drei Modi.

$(n_{BC}, n_{DE}, n_{FG})$  folgte Multinomialverteilung.

Textanalyse: Auftreten der einzelnen Buchstaben des Alphabets.

$(n_A, \dots, n_Z)$  folgt Multinomialverteilung.

# Poissonverteilung: Herleitung 1

Betrachte Grenzfall der Binomialverteilung mit

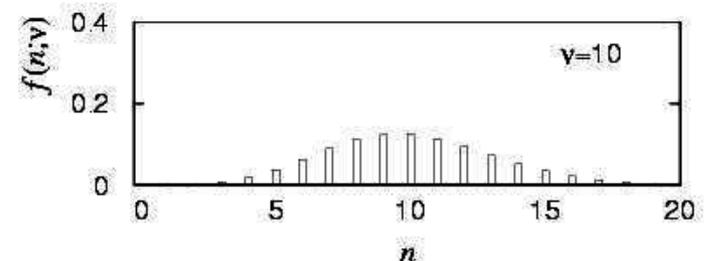
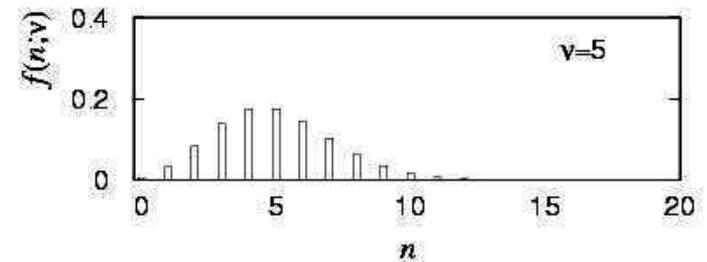
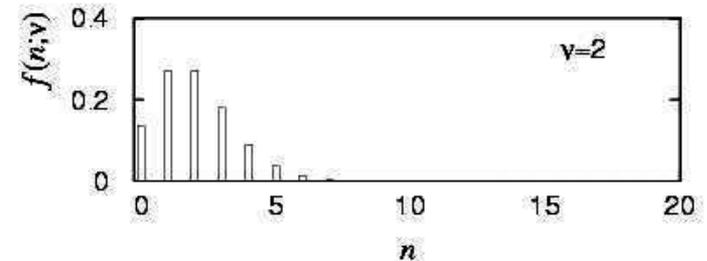
$$N \rightarrow \infty, \quad p \rightarrow 0, \quad E[n] = Np \rightarrow \nu \text{ . konstant}$$

→  $n$  folgt der Poissonverteilung:

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

$$E[n] = \sum_{n=0}^{\infty} n \frac{\lambda^n}{n!} \exp(-\lambda) = \lambda,$$

$$V[n] = \sum_{n=0}^{\infty} (n - \lambda)^2 \frac{\lambda^n}{n!} \exp(-\lambda) = \lambda.$$



# Poissonverteilung (i)

Betrachte Grenzfall der Binomialverteilung mit

$$N \rightarrow \infty, \quad p \rightarrow 0, \quad E[n] = Np \rightarrow \nu . \quad \text{konstant}$$

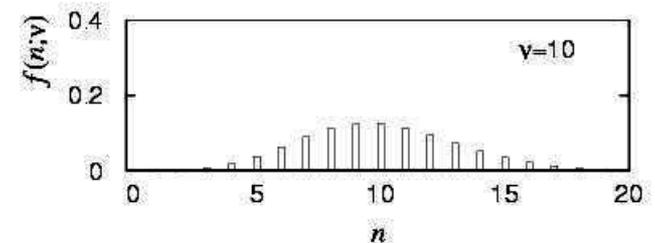
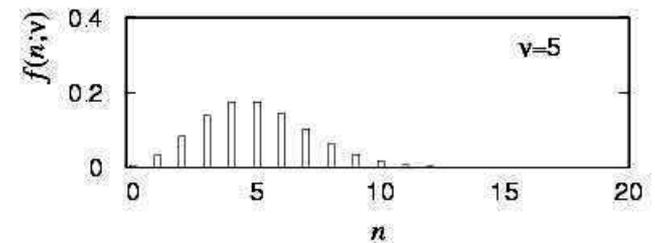
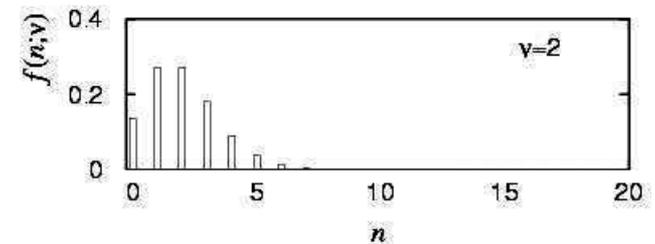
→  $n$  folgt der Poissonverteilung:

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

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

Example: number of scattering events  $n$  with cross section  $\sigma$  found for a fixed integrated luminosity, with

$$\nu = \sigma \int L dt .$$



# Poissonverteilung (i)

Quantile $\lambda_P$ der Poissonverteilung						
$k \backslash P$	$P = \sum_{n=0}^k \frac{\lambda_P^n}{n!} \exp(-\lambda_P)$					
	0.01	0.05	0.10	0.90	0.95	0.99
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

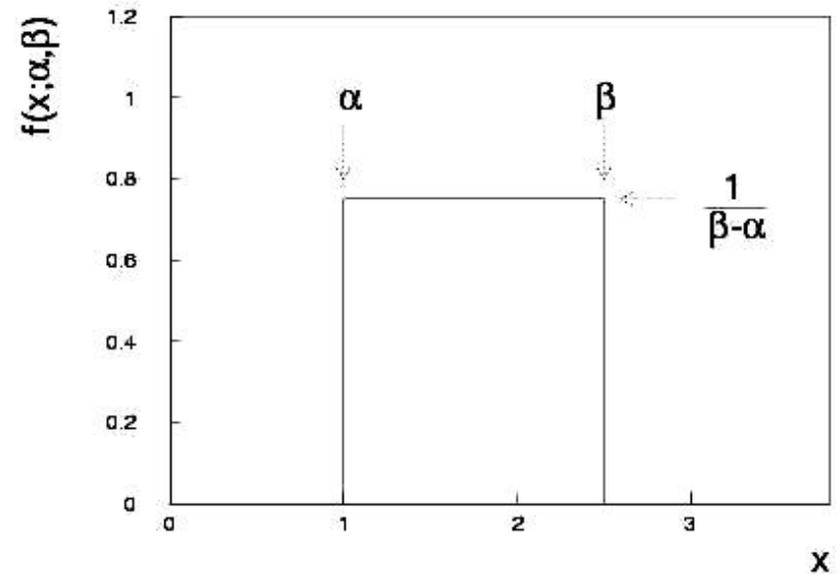
# Gleichverteilung

Betrachte kontinuierliche ZV  $x$  mit  $-\infty < x < \infty$ . Gleichverteilung lautet:

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

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

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



Bemerkung: für beliebige ZV  $x$  mit Kumulativverteilung  $F(x)$ , gilt, dass die ZV  $y = F(x)$  gleichverteilt in  $[0, 1]$  ist.

Example: for  $p^0 \rightarrow gg$ ,  $E_g$  is uniform in  $[E_{\min}, E_{\max}]$ , with

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

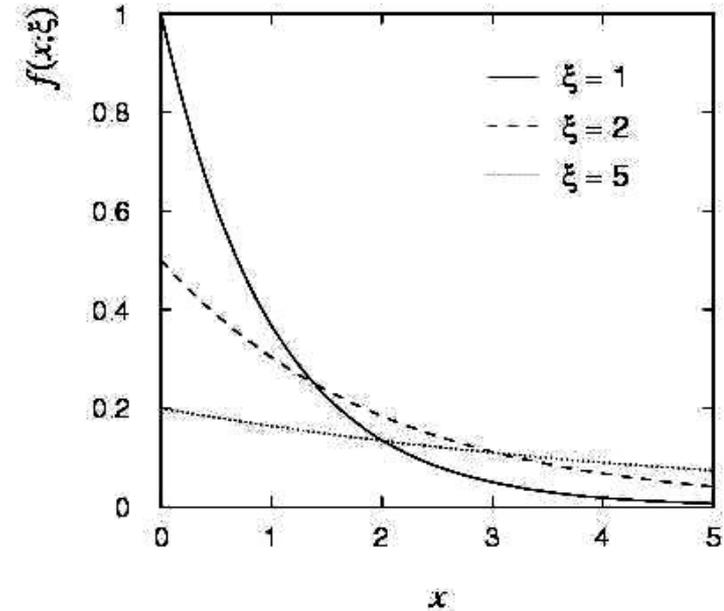
# Exponentialverteilung

Die Exponential-WDF für kontinuierliche ZV  $x$  ist definiert als:

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

$$E[x] = \xi$$

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



Beispiel: Zerfallszeit eines instabilen Teilchens

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

Kein Gedächtnis (einzigartig für Exp-WDF)  $f(t - t_0 | t \geq t_0) = f(t)$

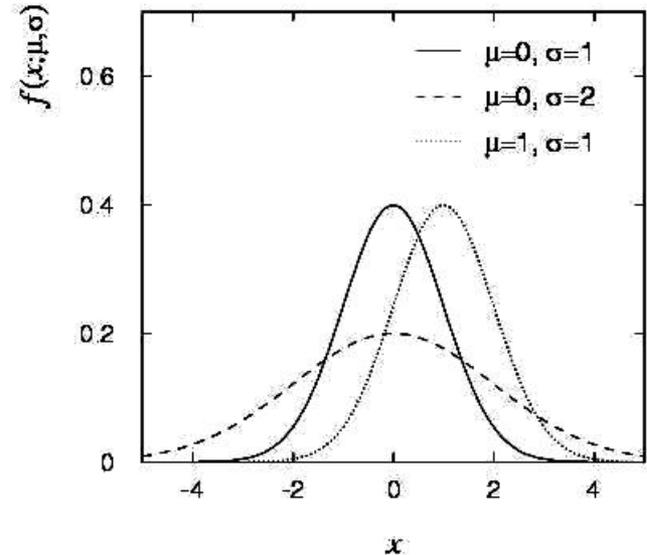
# Gauss- oder Normalverteilung

Dies Gauss- oder Normal-WDF für eine kontinuierliche ZV  $x$  ist definiert als:

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

$$E[x] = \mu$$

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



Spezialfall:  $\mu = 0, \sigma^2 = 1$  ('Standard-Gauss'):

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

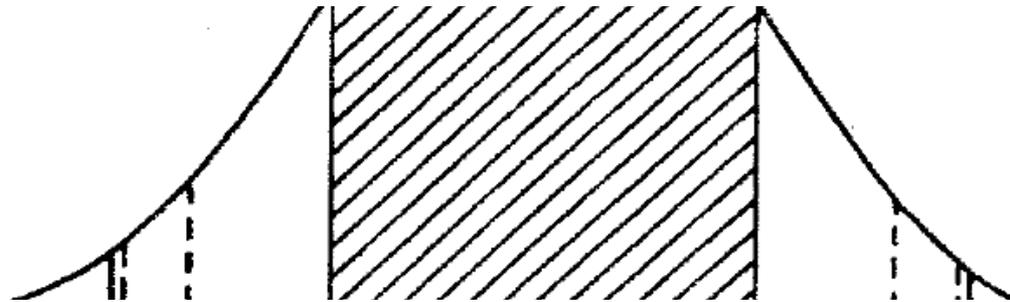
Wenn  $y$  aus Gauss-WDF mit  $\mu, \sigma^2$ , dann folgt  $x = (y - \mu) / \sigma$  der  $\varphi(x)$ .

# Gauss- oder Normalverteilung

$$P(\mu - \sigma \leq x \leq \mu + \sigma) = 2G(1) - 1 = 0.6827$$

$$P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) = 2G(2) - 1 = 0.9549$$

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



$$P(\mu - 1.645\sigma \leq x \leq \mu + 1.645\sigma) = 0.90$$

$$P(\mu - 1.960\sigma \leq x \leq \mu + 1.960\sigma) = 0.95$$

$$P(\mu - 2.576\sigma \leq x \leq \mu + 2.576\sigma) = 0.99$$

$$P(\mu - 3.290\sigma \leq x \leq \mu + 3.290\sigma) = 0.999$$

# Gauss- oder Normalverteilung

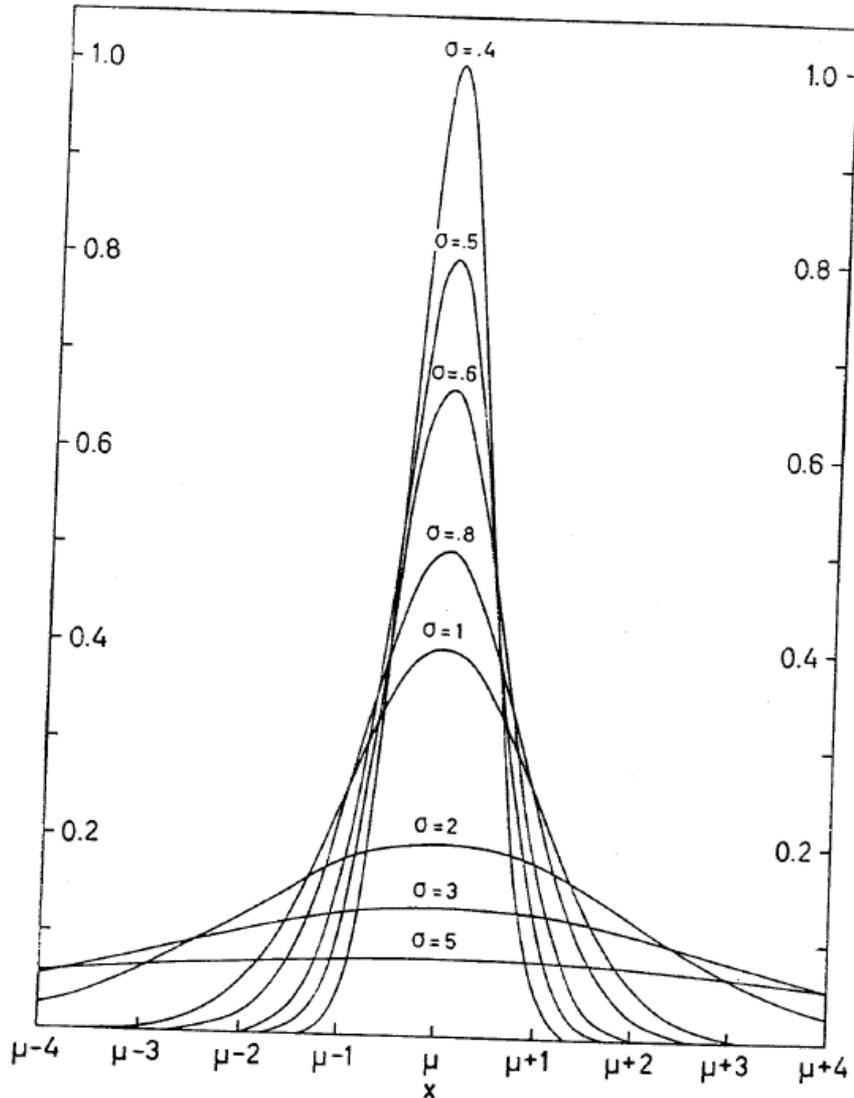


Fig. 4.6. The normal p.d.f.  $N(\mu, \sigma^2)$  for different values of the standard deviation  $\sigma$ .

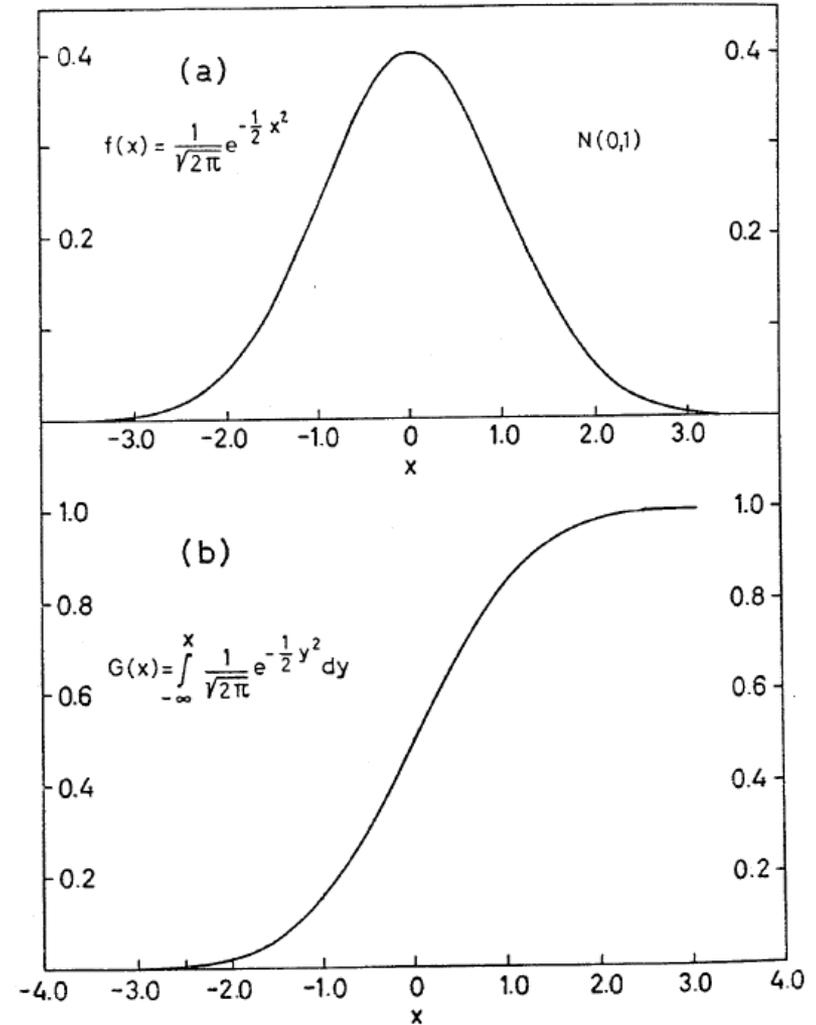
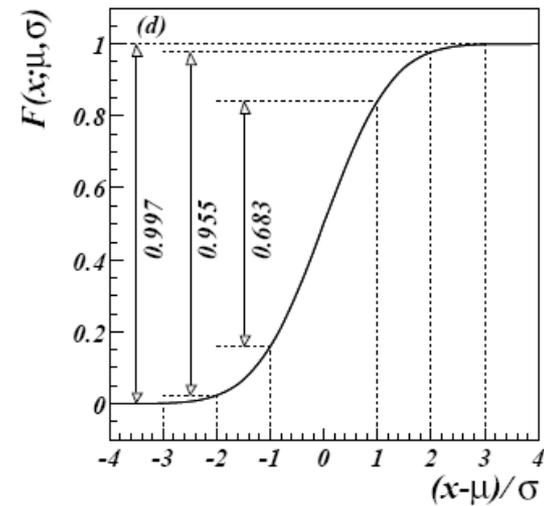
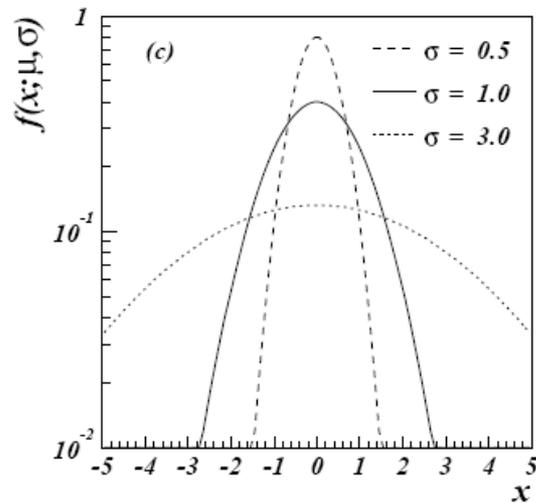
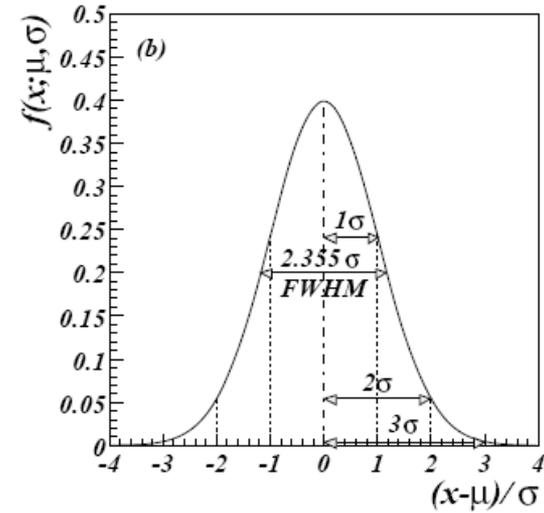
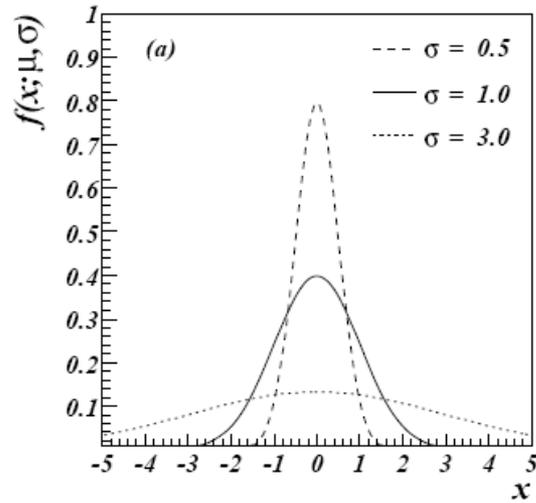


Fig. 4.7. (a) The standard normal p.d.f.  $N(0,1)$ .  
(b) The cumulative standard normal distribution.

# Gauss- oder Normalverteilung



# Gauss-WDF und Zentraler Grenzwertsatz

Die Gaussverteilung ist von so großer Bedeutung weil jede ZV, die die Summe aus einer großen Anzahl kleiner Zahlenbeiträge ist gemäß ihr verteilt ist. Dies folgt aus dem “Zentralen Grenzwertsatz”:

Gegeben  $n$  unabhängige ZV  $x_i$  mit endlicher Varianzen  $\sigma_i^2$ , aber ansonsten beliebigen WDFs. Betrachte die ZV  $y$  als Summe

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

Im Grenzfall  $n \rightarrow \infty$  gilt, dass  $y$  einer Gaussverteilung folgt mit

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

Messfehler sind oft die Summe aus vielen kleinen Beiträgen. Daher können häufig gemessene Werte als gaussverteilt angenommen werden.

Der ZGS kann unter Verwendung der charakteristischen Funktionen bewiesen werden (siehe z.B. Cowan Kapitel 10).

# Zentraler Grenzwertsatz (ZGS) (II)

Für endliche  $n$ , gilt der ZGS in “guter” Näherung wenn die Fluktuationen der Summe der ZV nicht durch einen (oder wenige) Beiträge dominiert wird.



Vorsicht vor Messungen mit nicht-Gausschen Fehlern

Gutes Beispiel: Geschwindigkeitskomponente  $v_x$  von Luftmolekülen

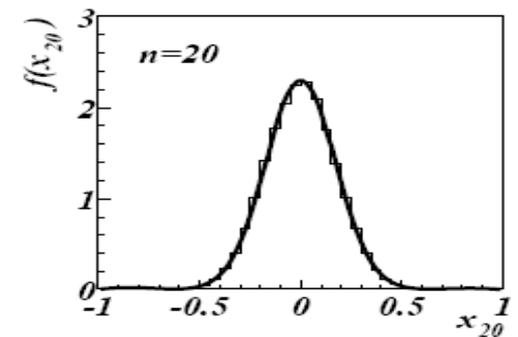
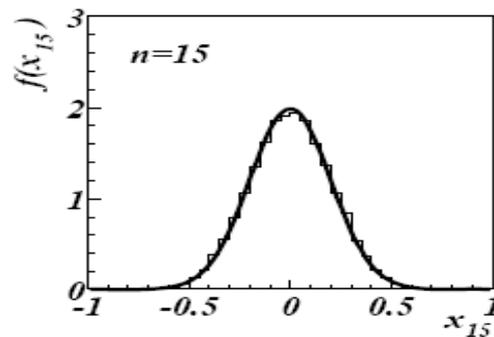
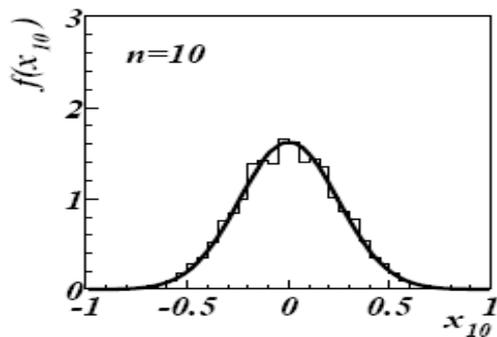
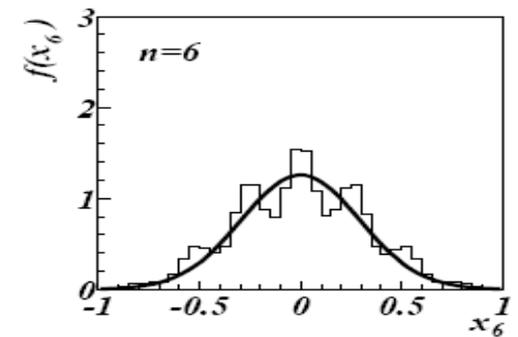
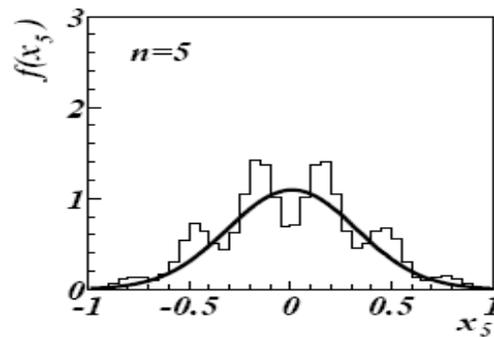
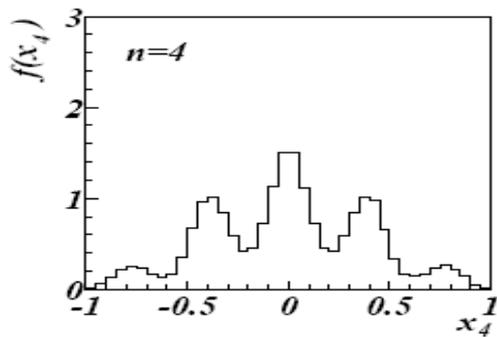
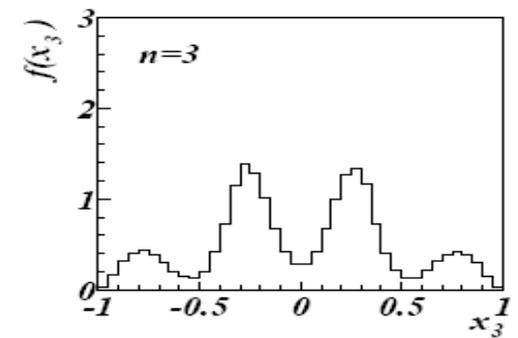
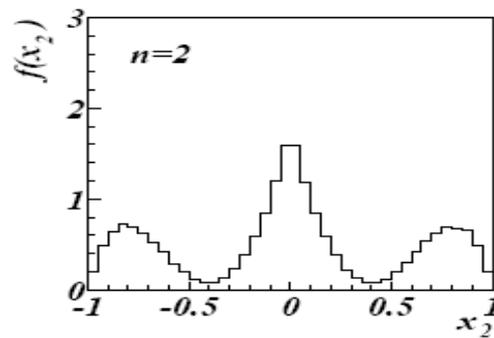
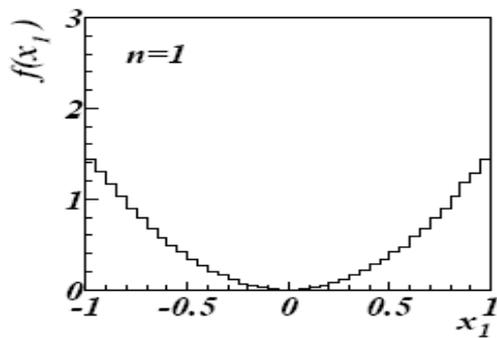
“OK”-Beispiel: totale Ablenkung durch Coulomb-Vielfachstreuung  
(Seltende Ablenkungen unter gr. Winkeln ergeben nicht-Gaussche Ausläufer)

Schlechtes Beispiel: Energieverlust von geladenem Teilchen in dünner Gasschicht. (Seltende Kollisionen tragen Großteil des Energieverlustes → Landau-WDF).

Gutes Beispiel: Größe des Menschen. Viele Faktoren beeinflussen Größe.

Schlechtes Beispiel: Gewicht des Menschen. Dominiert durch Essverhalten.

# Central Limit Theorem (2)



# Mehrdimensionale Gaussverteilung

Mehrdimensionale Gauss-WDF für Vektor von ZV  $\vec{x} = (x_1, \dots, x_n)$  :

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

$\vec{x}$ ,  $\vec{\mu}$  sind Spaltenvektoren  $\vec{x}^T$ ,  $\vec{\mu}^T$  sind transponierte (Zeilen-)Vektoren

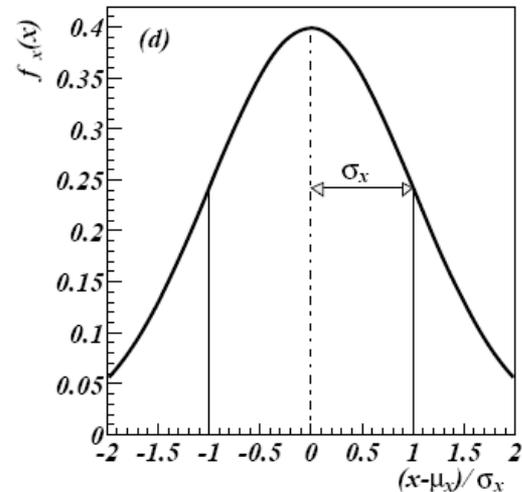
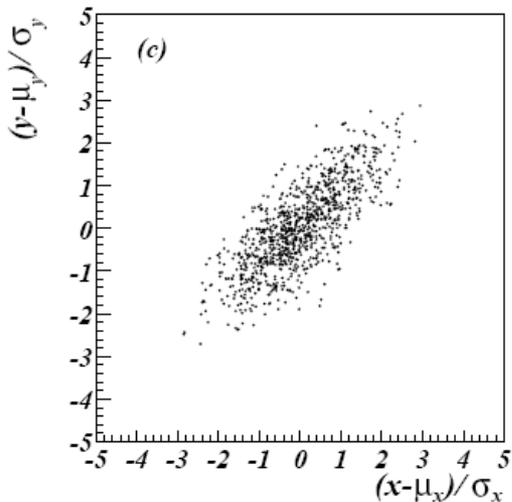
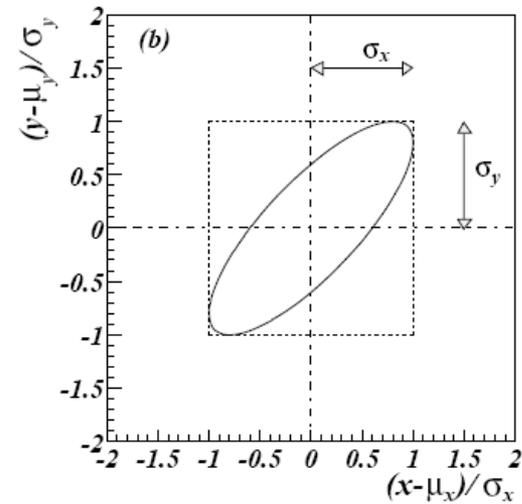
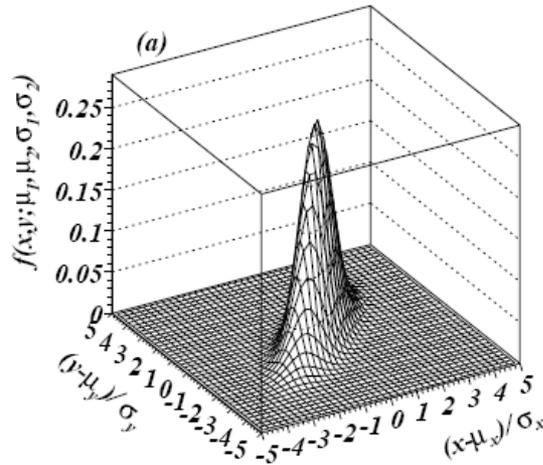
$$E[x_i] = \mu_i, \quad \text{COV}[x_i, x_j] = V_{ij} .$$

Für  $n = 2$  ist die WDF gegeben durch:

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

wobei  $\rho = \text{cov}[x_1, x_2]/(\sigma_1\sigma_2)$  der Korrelationskoeffizient ist.

# Mehrdimensionale Gaussverteilung



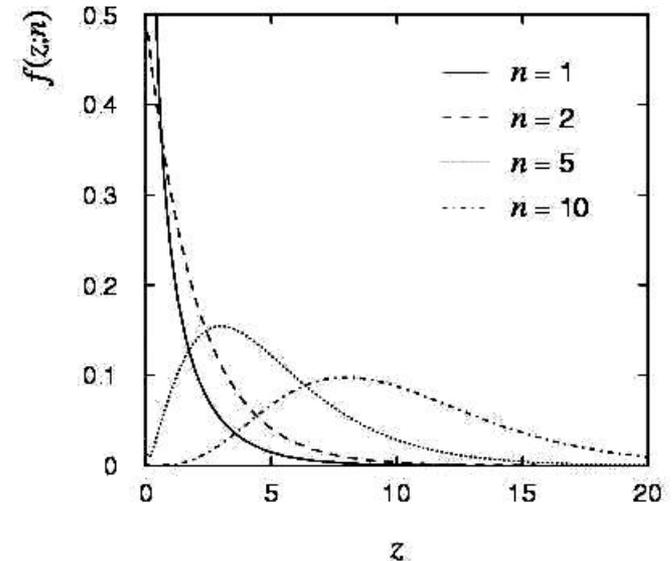
# Chi-Quadrat-( $\chi^2$ )-Verteilung

Die Chi-Quadrat-WDF für kontinuierliche ZV  $z$  ( $z \geq 0$ ) ist definiert als:

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

$n = 1, 2, \dots =$  Anzahl der Freiheitsgrade "FG"

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



Für unabhängige Gauss-ZV  $x_i$ ,  $i = 1, \dots, n$ , mit Mittelwerten  $\mu_i$ , Varianzen  $\sigma_i^2$ ,

$$z = \sum_{i=1}^n \frac{(x_i - \mu_i)^2}{\sigma_i^2} \quad \text{folgt einer } \chi^2 \text{ WDF mit } n \text{ FG.}$$

Beispiel: Test der Güte der Anpassung im besonderen im Zusammenhang mit der Methode der kleinsten Quadrate.

# Quantiles of chi square distribution

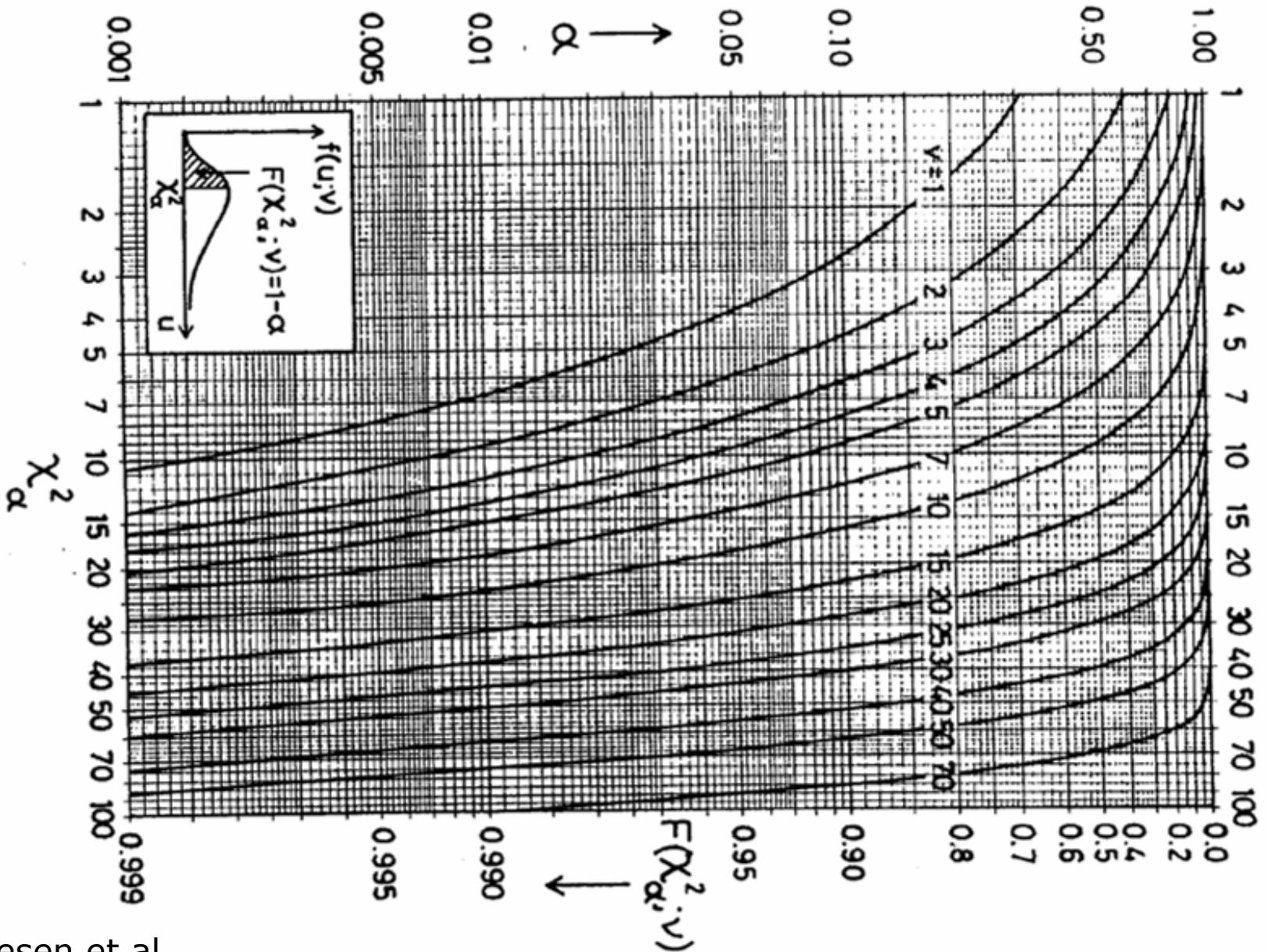
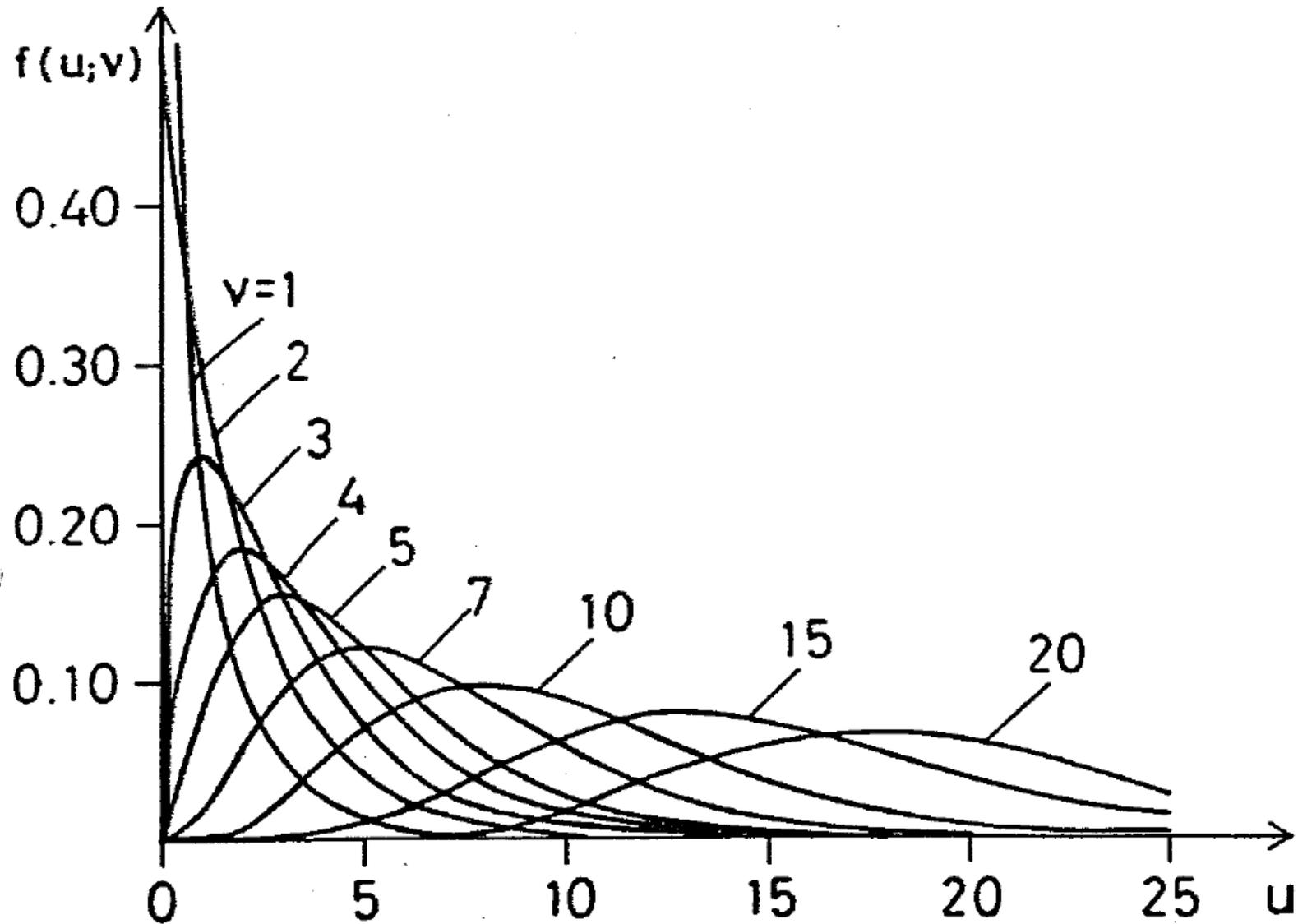


Fig. 5.2. Probability contents of the chi-square distribution.

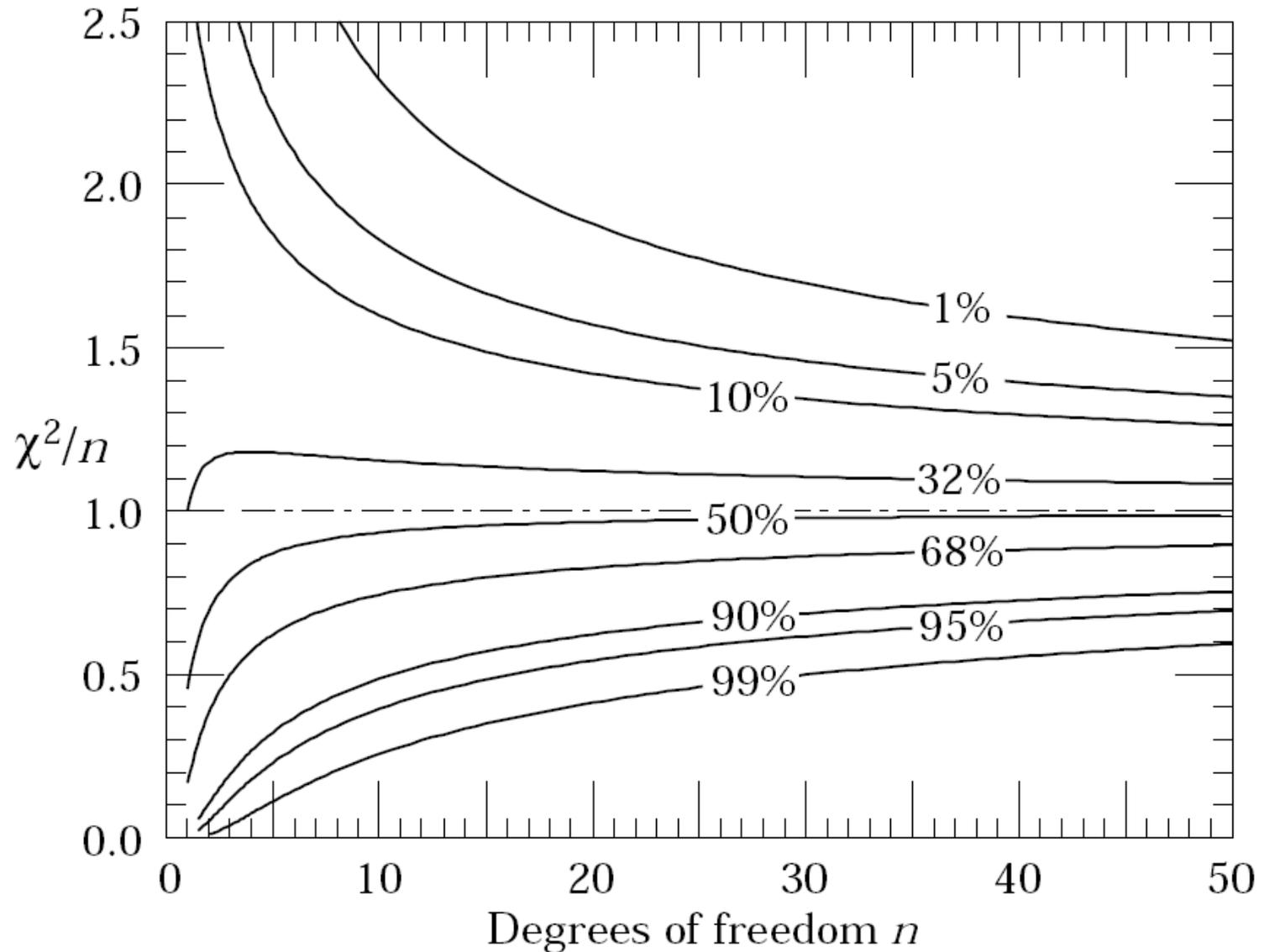
Frodesen et al.

# Quantiles of chi square distribution



F1 Fig. 5.1. The chi-square distribution for different degrees of freedom  $v$ .

# Quantiles of chi square distribution



Frodesen et al.

# Cauchy- und Breit-Wigner-Verteilung

Die Breit-Wigner-WDF für eine kontinuierliche ZV  $x$  ist definiert als

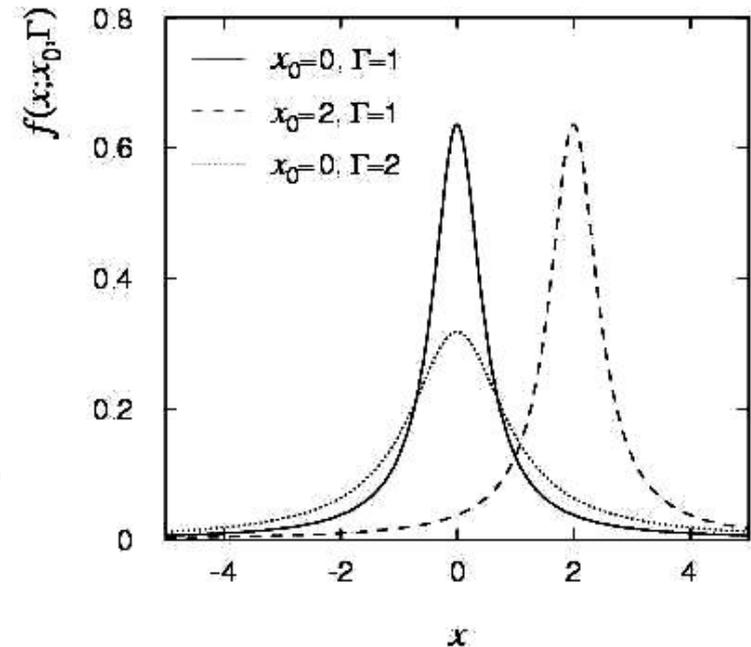
$$f(x; \Gamma, x_0) = \frac{1}{\pi} \frac{\Gamma/2}{\Gamma^2/4 + (x - x_0)^2}$$

(Spezialfall: Cauchy-Verteilung  
für  $\Gamma = 2$  und  $x_0 = 0$ )

$E[x]$  nicht wohldefiniert,  $V[x] \rightarrow \infty$ .

$x_0$  = Modalwerte (wahrscheinlichster Wert)

$\Gamma$  = volle Breite auf halber Maximalhöhe



Beispiel: Massenverteilung eines instabilen Teilchens z.B.  $\rho$ ,  $K^*$ ,  $f^0$ , ...

$\Gamma$  = Zerfallsbreite (Reziproke der mittleren Lebensdauer)

# Cauchy / Breit Wigner Distribution

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

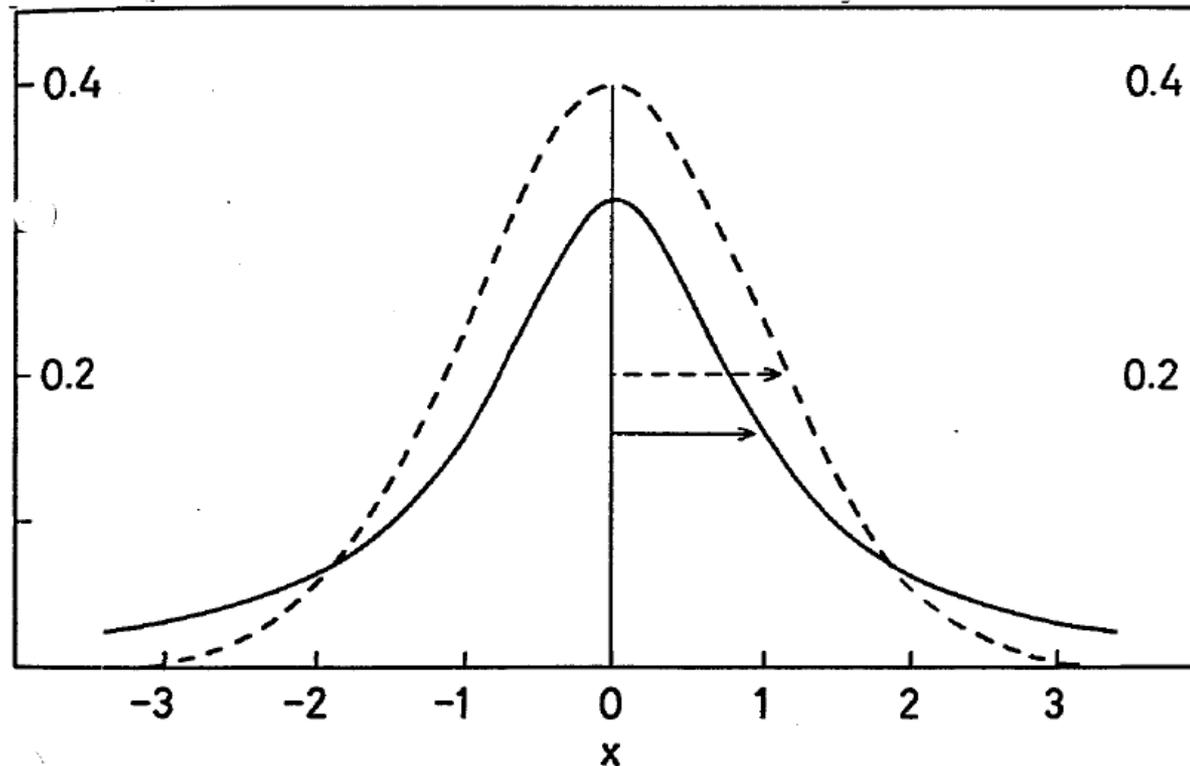


Fig. 4.9. The Cauchy or Breit-Wigner distribution (solid curve) and the standard normal distribution (dashed curve). The half-widths at half-maximum are indicated by arrows of length 1 and  $\sqrt{2\ln 2} \approx 1.18$ , respectively.

Frodesen et al.

# Landau distribution

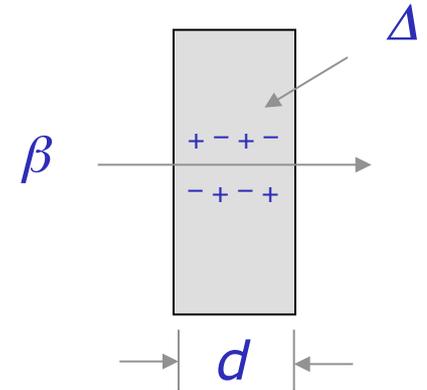
Der Energieverlust  $\Delta$  eines geladenen Teilchens mit Geschwindigkeit  $\beta = v/c$ , welches eine Materialschicht der Dicke  $d$  durchquert, folgt der Landau-Verteilung.

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

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

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

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



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 (II)

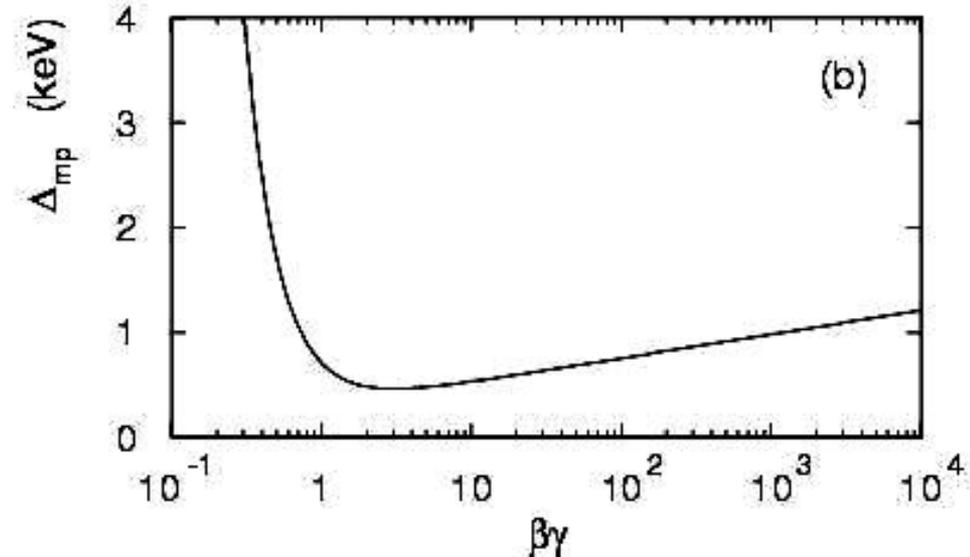
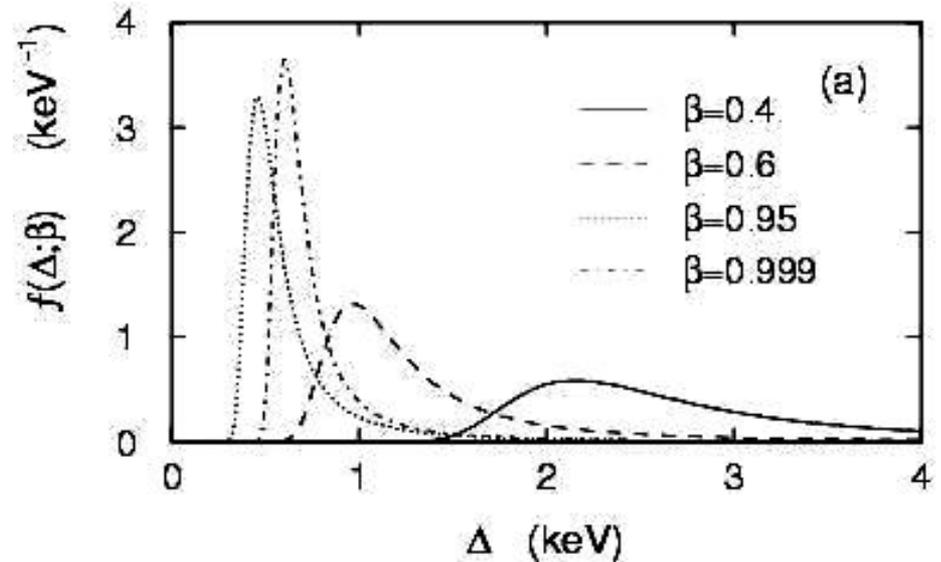
Lange “Landau-Ausläufer”

→ alle Momente  $\infty$

Modalwert sensitiv

auf Geschwindigkeit  $\beta$ ,

→ Teilchenidentifikation



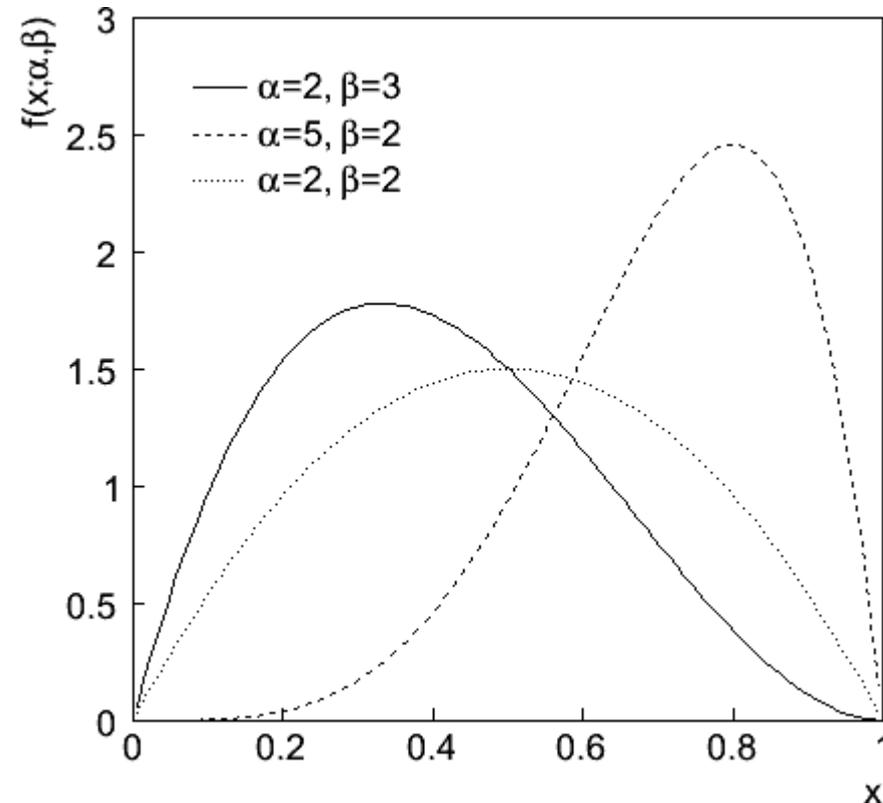
# Beta-Verteilung

Die Beta-WDF für eine kontinuierliche ZV  $x$  im Intervall  $[0,1]$  ist definiert als

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

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

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



Wird oft verwendet um WDF zu approximieren, die eine kontinuierliche ZV in einem gewissen Bereich beschrieben  $[a,b]$  beschreiben.

# Gamma-Verteilung

Die Gamma-WDF für eine kontinuierliche ZV  $x$  im Intervall  $[0, \infty]$  ist definiert als

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

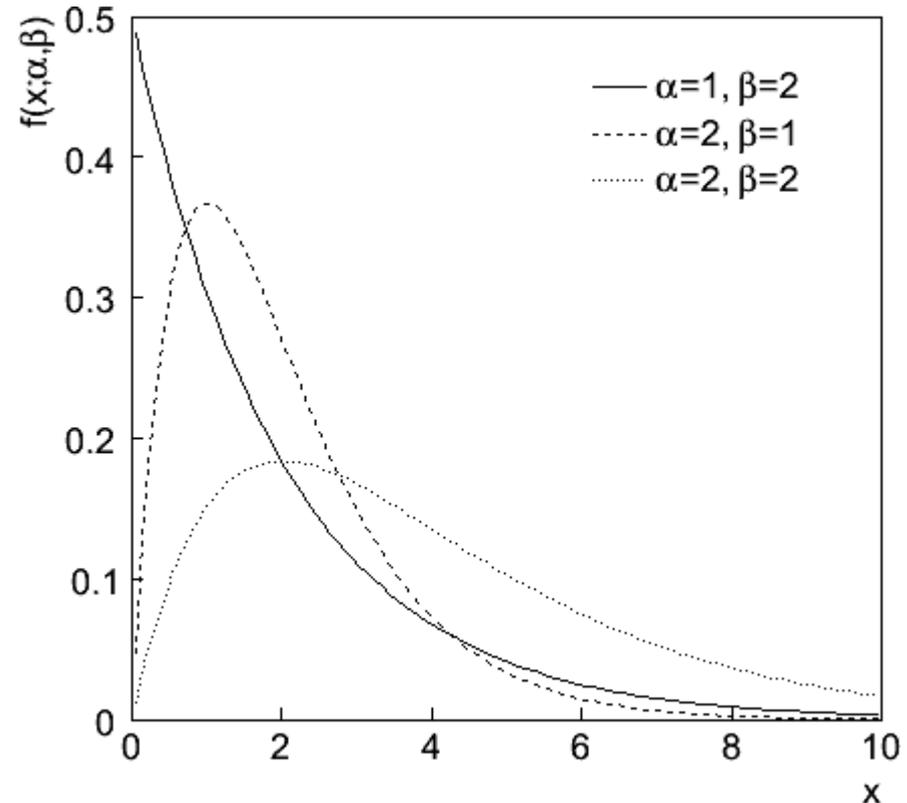
$$E[x] = \alpha\beta$$

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

Oft benutzt um eine WDF im Bereich  $[0, \infty]$  zu parametrisieren.

Beschreibt auch z.B.

- Summe von  $n$  ZV aus Exponentialverteilung
- Zeit bis zum r.v.s or time until  $n$ ten Ereignis in einem Poisson-Prozess



# Studentsche $t$ -Verteilung

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

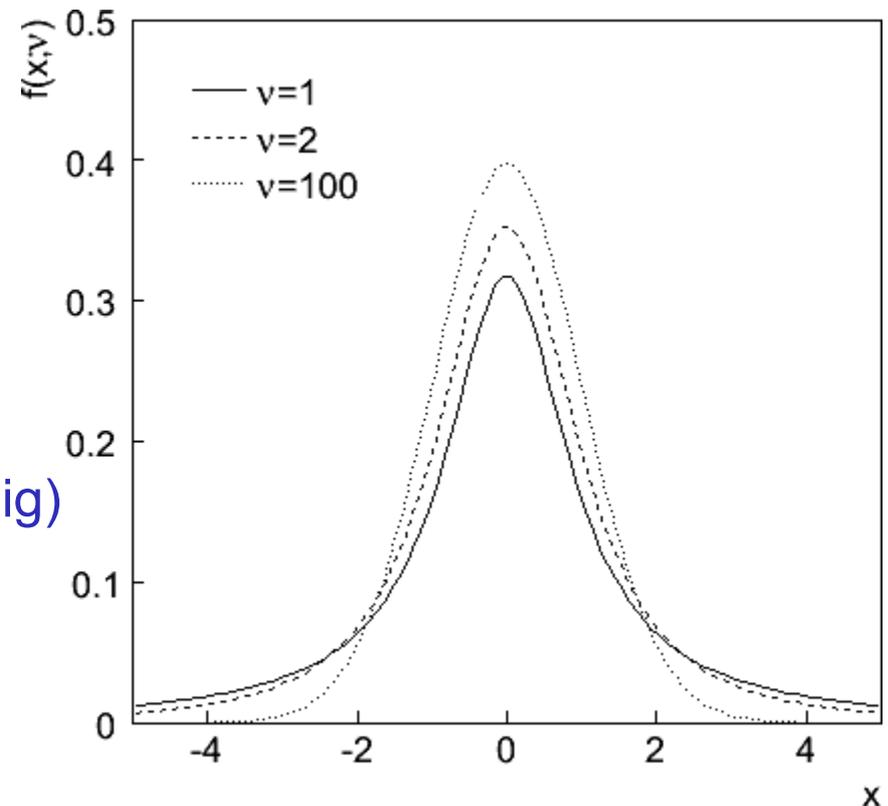
$$E[x] = 0 \quad (\nu > 1)$$

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

$\nu$  = Anzahl der Freiheitsgrade  
(nicht notwendigerweise ganzzahlig)

$\nu = 1$  ergibt Cauchy-WDF

$\nu \rightarrow \infty$  ergibt Gauss-WDF



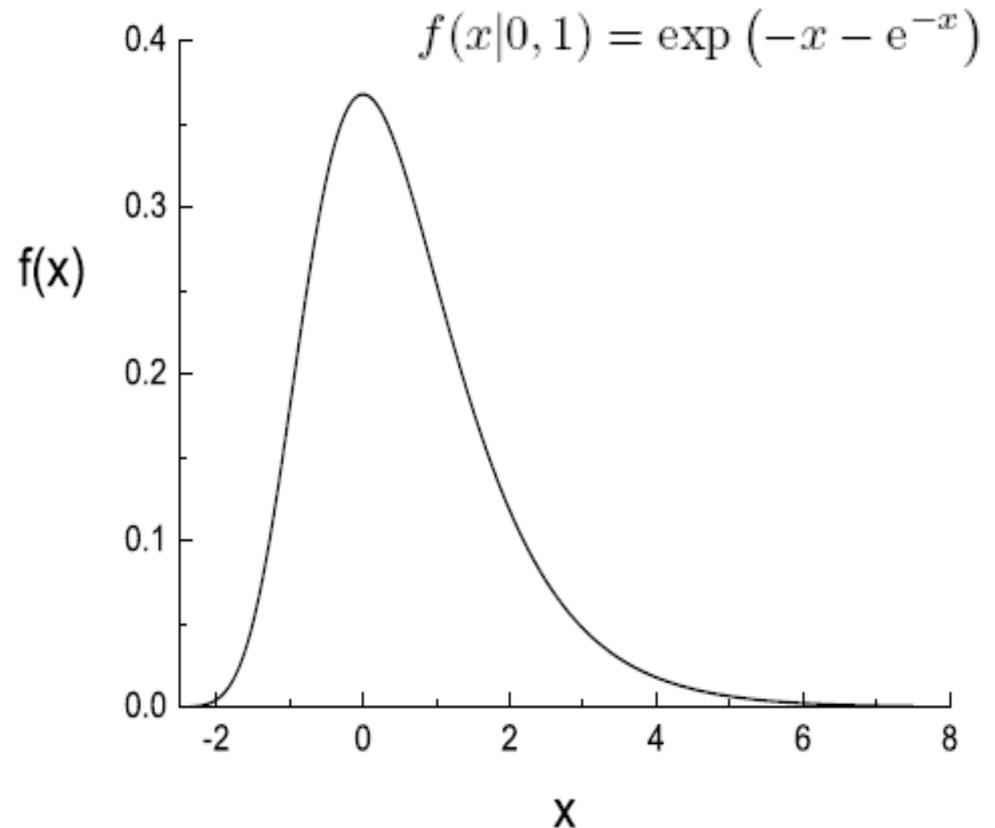
# Log-Weibull-Distribution

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

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

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

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



# Studentsche $t$ -Verteilung

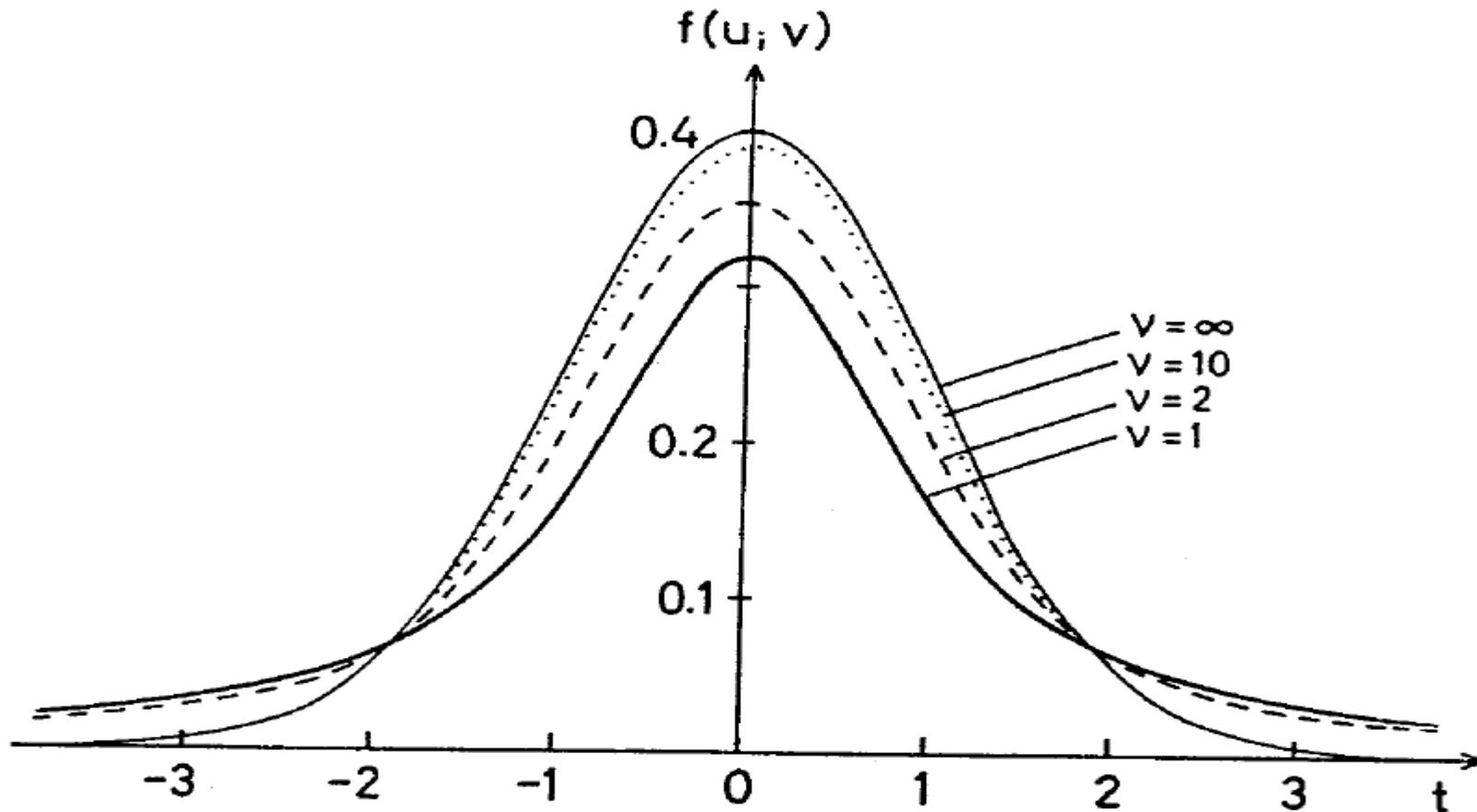


Fig. 5.3. The Student's  $t$ -distribution for different degrees of freedom  $v$ . The special cases  $v=1$  and  $v=\infty$  correspond to the Cauchy (Breit-Wigner) and the standard normal distributions, respectively.

# Studentsche $t$ -Verteilung (II)

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.

# Studentsche $t$ -Verteilung (II)

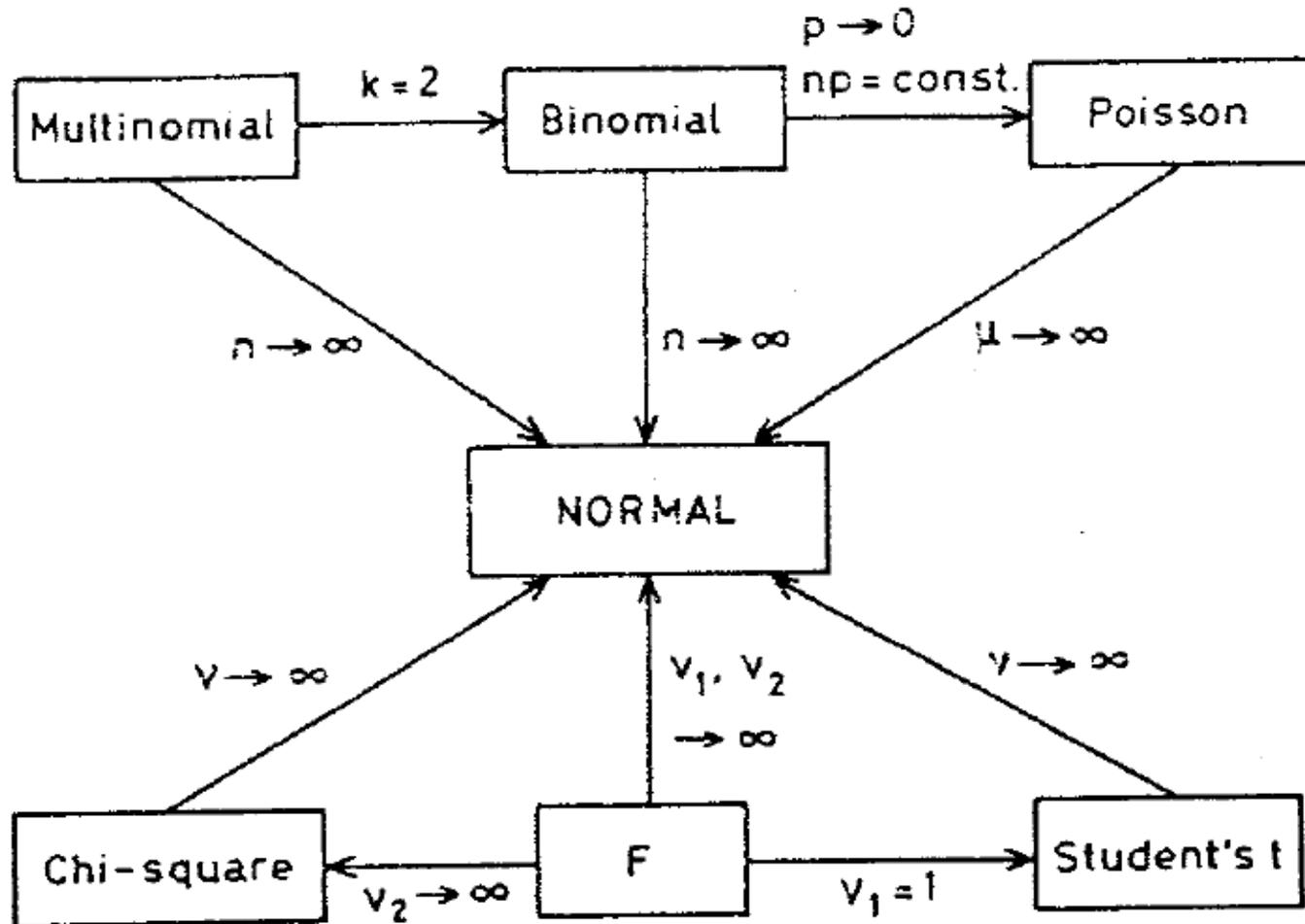
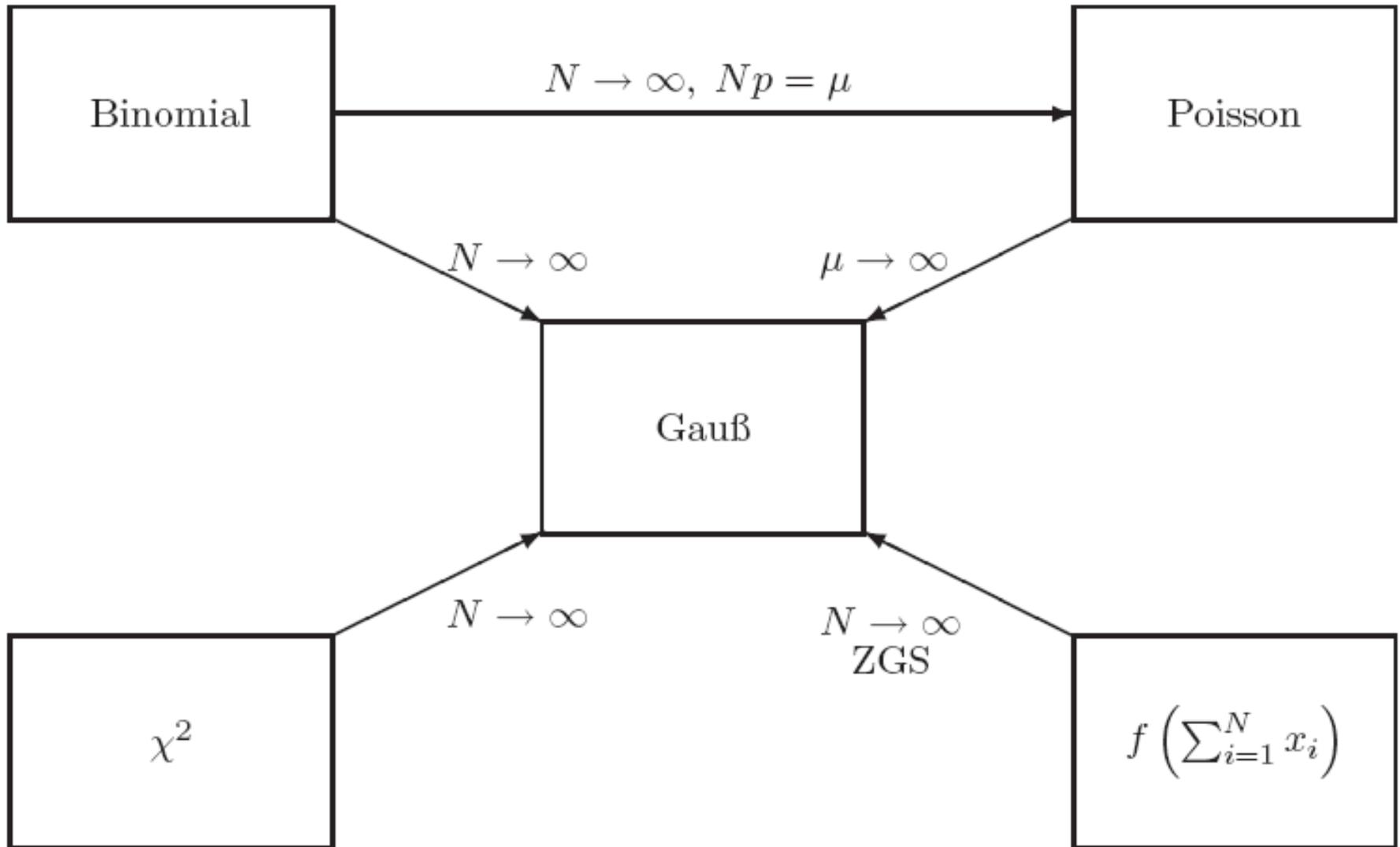


Fig. 5.5. Relations between probability distributions.

# Studentsche $t$ -Verteilung (II)

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

# Studentsche $t$ -Verteilung (II)



## **Kapitel 3**

# **Die Monte-Carlo-Methode für Erstellung von Wahrscheinlichkeitsdichtefunktionen**

# Die Monte-Carlo-(MC)-Methode

MC-Methode ist eine numerische Technik zur Bestimmung von

- Wahrscheinlichkeitsdichtefunktionen
- Transformation von Zufallsvariablen
- Bestimmung von Integralen
- Erwartungswerte
- Faltungen

mit Hilfe von Zufallszahlen

- Anwendungen:
- Generierung von Ereignissen/Messungen gemäß eines theoretischen Modells
  - Simulation des Ansprechverhaltens eines Nachweisapparates/ einer Messapratrur
  - ....

# The MC-Methode (II)

Die einzelnen Schritte sind:

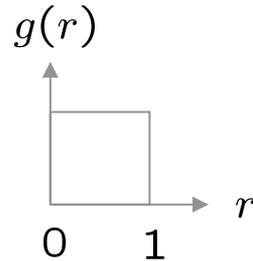
- (1) Generiere eine Sequenz von Zufallszahlen  $r_1, r_2, \dots, r_m$  gemäß Gleichverteilung im Intervall  $[0, 1]$ .
- (2) Verwende diese um eine weitere Sequenz  $x_1, x_2, \dots, x_n$  zu erzeugen, die gemäß einer vorgegebenen WDF  $f(x)$  verteilt sind. ( $x$  kann Vektor sein).
- (3) Verwende  $x_i$  Werte um Eigenschaften von  $f(x)$  zu bestimmen, z.B. Erwartungswerte oder Anteil  $x$  Werten mit  $a < x < b$  ergibt

→ MC-Berechnung = integration (zumindest formal)  $\int_a^b f(x) dx$ .

MC generierte Werte = “simulierte Daten”

→ oft verwendet um Gültigkeit statistischer Methoden

Besonders nützlich bei:- vieldimensionalen  $f(x)$   
- komplizierten Unterräumen des “ $x$ ”<sup>n</sup>



# Zufallszahlengenerator

Ziel: Erzeugung gleichmäßig verteilter Werte im Intervall  $[0, 1]$ .

Werfe Münze für z.B. 32 Bitnummer... (too tiring).

→ 'Zufallszahlengenerator'

= Computeralgorithmus zur Erzeugung von  $r_1, r_2, \dots, r_n$ .

Beispiel: Multiplikativer linear kongruenter Generator (MLCG)

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

$$n_i = \text{ganze Zahl}$$

$$a = \text{Multiplikator}$$

$$m = \text{Modulus/Rest}$$

$$n_0 = \text{"Seed"} = \text{"Saatzahl"} \text{ (Startwert)}$$

Bemerkung: mod = modulus (Reste), z.B.  $27 \bmod 5 = 2$ .

Diese Regel erzeugt eine Sequenz von Zahlen  $n_0, n_1, \dots$

# Zufallszahlengenerator (2)

Die Sequenz ist (unglücklicherweise) periodisch!

Beispiel: (siehe Brandt Kapitel 4):  $a = 3$ ,  $m = 7$ ,  $n_0 = 1$

$$n_1 = (3 \cdot 1) \bmod 7 = 3$$

$$n_2 = (3 \cdot 3) \bmod 7 = 2$$

$$n_3 = (3 \cdot 2) \bmod 7 = 6$$

$$n_4 = (3 \cdot 6) \bmod 7 = 4$$

$$n_5 = (3 \cdot 4) \bmod 7 = 5$$

$$n_6 = (3 \cdot 5) \bmod 7 = 1 \quad \leftarrow \text{Folge wiederholt sich}$$

Wähle  $a$ ,  $m$  so, dass eine lang Periode (Maximum =  $m - 1$ ) erreicht wird;  $m$  normalerweise nahe an der größten "Integer"-Zahl, die auf dem Computer repräsentiert werden kann.

Verwende nur ein Untermenge der einzelnen Periode der Sequenz.

# Zufallszahlengenerator (3)

$r_i = n_i/m$  sind  $[0, 1]$  aber sind sie “zufällig”?

Wähle  $a, m$  so, dass die  $r_i$  eine Reihe von test für Zufälligkeit erfüllen:

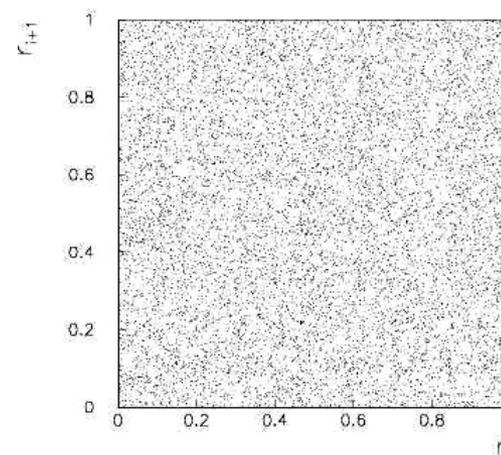
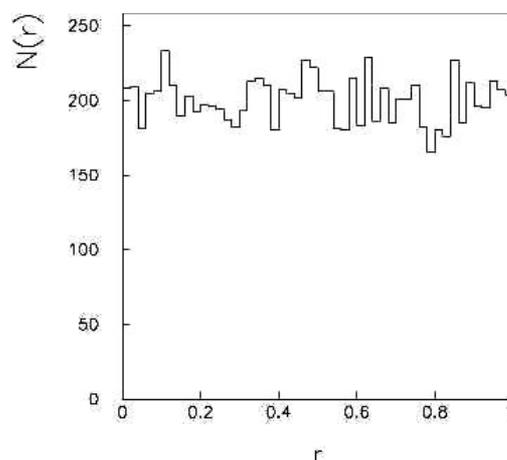
gleichförmig verteilt in  $[0, 1]$ ,

alle Werte unabhängig (keine Korrelationen zwischen Paaren),

z.B. L'Ecuyer, Commun. ACM **31** (1988) 742 schlägt

$a = 40692$

$m = 2147483399$



Es gibt weit bessere Generatoren, e.g. **TRandom3**, basierend auf “Mersenne twister” Algorithmus, Periode =  $2^{19937} - 1$  (a “Mersenne prime”).

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

# Die Transformationsmethode

Anwendung der Transformationsmethode für Zufallsvariablen

bisher:  $f(x)$                        $a(x)$                        $\rightarrow$                        $g(a)$   
WDF für  $x$                       Funktion                      WDF für  $a$   
gegeben                      gegeben                      gesucht

jetzt:  $g(r)$                        $x(r)$                        $\rightarrow$                        $f(x)$   
Gleichverteilung                      Transformation                      WDF für  $x$   
in  $r$  gegeben                      gesucht                      gegeben

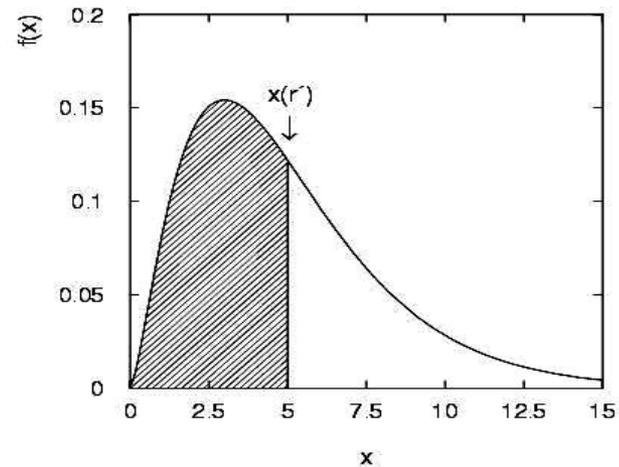
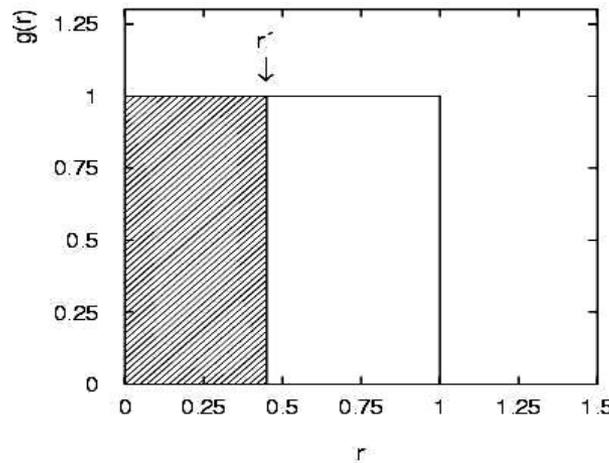
Ziel: gegeben Sequenz  $r_1, r_2, \dots, r_n$  gleichförmig in  $[0, 1]$ ,  
erzeuge  $x_1, x_2, \dots, x_n$  die  $f(x)$  folgen durch Auffinden  
einer geeigneten Transformation  $x(r)$ .

# The transformation method

Verlange: Wkt, dass  $r$  in  $][r, r+dr]$  =  $g(r)dr$   
= Wkt., dass  $x$  in  $[x(r), x(r)+dx(r)]$  =  $f(x) dx$

Oder äquivalent:  $P(r \leq r') = P(x \leq x(r'))$

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



Da  $g(r)=1$  gilt:  $F(x) = r$  d.h.  $x = F^{-1}(r)$

Benötigt: Kumulativfunktion analytisch bestimmbar und invertierbar.

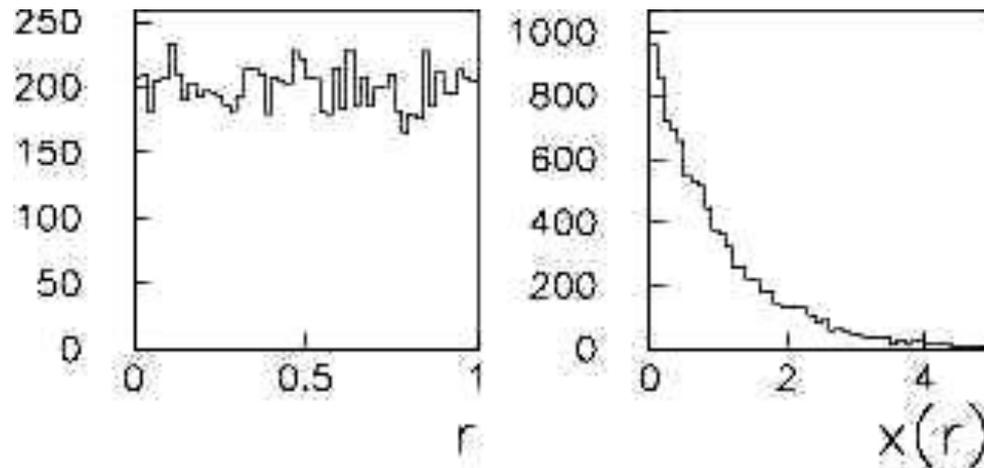
# Beispiel für die Transformationsmethode

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

Setze  $\int_0^x \frac{1}{\xi} e^{-x'/\xi} dx' = r$  und löse nach  $x(r)$  auf.

$$-e^{(-x/\xi)} + 1 = r$$

→  $x(r) = -\xi \ln(1 - r)$  ( $x(r) = -\xi \ln r$  geht auch)



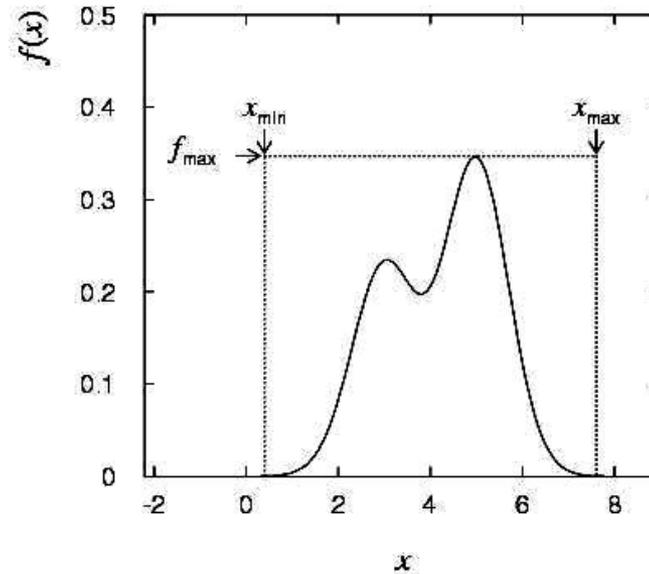
Vorteil: 100% Effizienz, d.h. aus jedem  $r_i$  wird ein  $x_i$  erzeugt.

# Weitere Beispiele für die Transformationsmethode

Wahrscheinlichkeitsdichte	Wertebereich	Algorithmus
$f(x) = \frac{1}{b-a}$	$[a, b[$	$x = (b-a) \cdot z + a$
$f(x) = 2x$	$[0, 1[$	$x = \max(z_1, z_2)$ or $x = \sqrt{z}$
$f(x) \sim x^{r-1}$	$[a, b[$	$x = [(b^r - a^r) \cdot z + a^r]^{1/r}$
$f(x) \sim \frac{1}{x}$	$[a, b[$	$a \cdot (b/a)^z$
$f(x) = \frac{1}{x^2}$	$]1, \infty]$	$x = 1/z$
$f(x) = \frac{1}{k} e^{-x/k}$	$]0, \infty]$	$x = -k \ln z$
$f(x) = x e^{-x}$	$]0, \infty]$	$x = -\ln(z_1 \cdot z_2)$
$f(x) = -\ln x$	$[0, 1[$	$x = z_1 \cdot z_2$
Gauss: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{-\frac{x^2}{2\sigma^2}}$	$[-\infty, \infty]$	$x = \sigma \sqrt{-\ln z_1^2} \cdot \cos(2\pi z_2)$
Breit-Wigner: $f(x) = \frac{\Gamma}{2\pi} \cdot \frac{1}{(x-\mu)^2 + (\Gamma/2)^2}$	$[-\infty, \infty]$	$x = [\tan \pi(z - 0.5)] \cdot \Gamma/2 + \mu$

# Die von-Neumannsche-Zurückweisungsmethode

Enclose the pdf in a box:



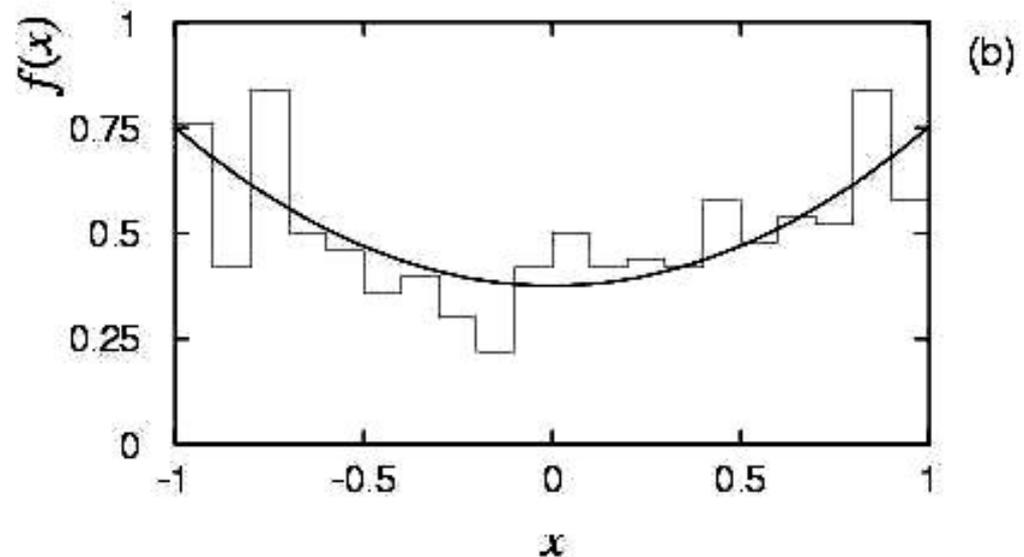
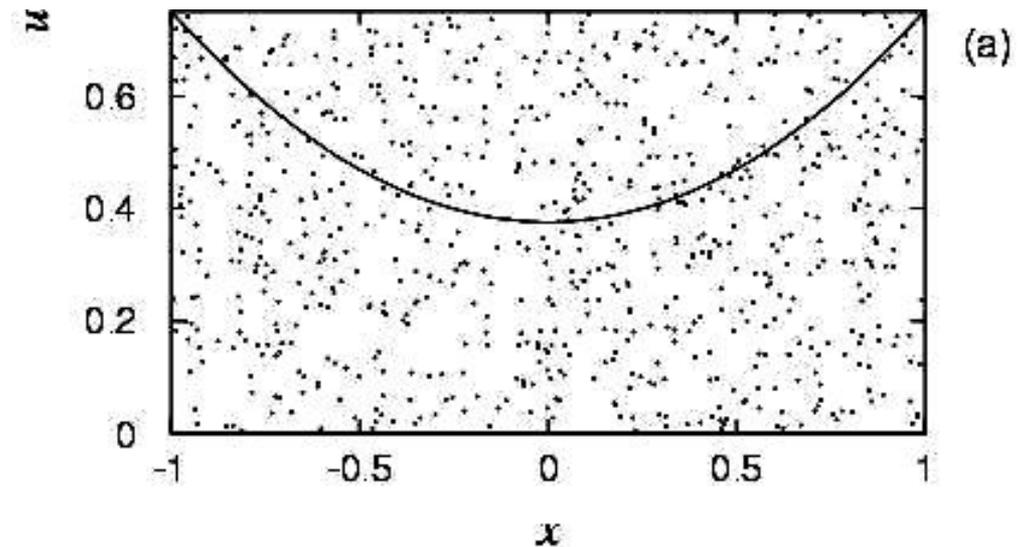
- (1) Generate a random number  $x$ , uniform in  $[x_{\min}, x_{\max}]$ , i.e.  $x = x_{\min} + r_1(x_{\max} - x_{\min})$ ,  $r_1$  is uniform in  $[0,1]$ .
- (2) Generate a 2nd independent random number  $u$  uniformly distributed between 0 and  $f_{\max}$ , i.e.  $u = r_2 f_{\max}$ .
- (3) If  $u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

# Die von-Neumannsche-Zurückweisungsmethode

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

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

Wenn Punkt unterhalb der Kurve, dann behalte ihn und Fülle  $x$ -Wert in Histogramm.

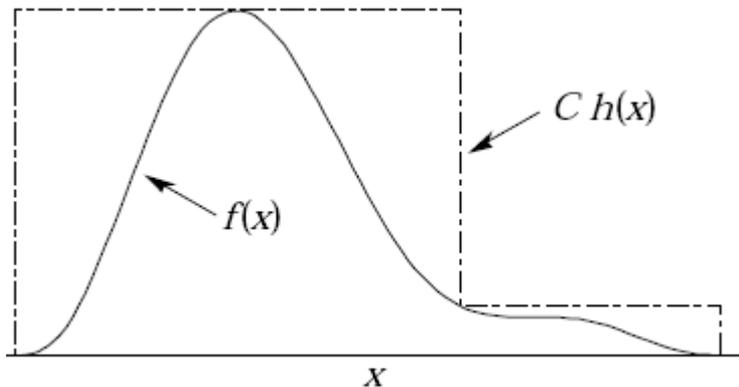


# Effizienzsteigerung für Zurückweisungsmethode

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.

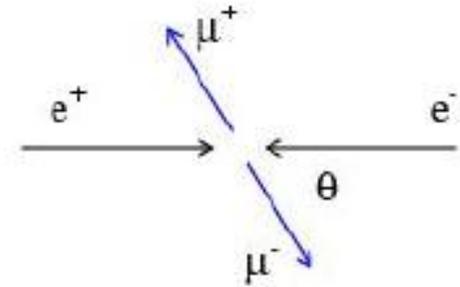


Generate points uniformly over  $C h(x)$ .

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

# Monte Carlo event generators

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



Generate  $\cos\theta$  and  $\phi$ :

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

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

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

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

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

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

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

# 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  $\Delta$ ),  
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.

Estimate ‘efficiencies’ = #events found / # events generated.

Programming package: **GEANT**

# Wrapping up lecture

We've now seen <sup>5</sup> the Monte Carlo method:  
calculations based on sequences of random numbers,  
used to simulate particle collisions, detector response.

So far, we've mainly been talking about **probability**.

But suppose now we are faced with experimental data.  
We want to infer something about the (probabilistic) processes  
that produced the data.

This is **statistics**, the main subject of the following lectures.