Who's Counting?

Basic Counting Statistics

Stochastic Nuclear Observables

2 sources of stochastic observables x in nuclear science:

1) Nuclear phenomena are governed by quantal wave functions and inherent statistics

2) Detection of processes occurs with imperfect efficiency ($\varepsilon < 1$) and finite resolution distributing sharp events x_0 over a range in x.

Stochastic observables x have a range of values with frequencies determined by probability distribution P(x).
characterize by set of moments of P

$$<x^n> = \int x^n P(x)dx; \quad n\ 0, 1, 2, \ldots$$

with the normalization $<x^0> = 1$. First moment of P:

$$E(x) = <x> = \int x P(x)dx$$

second central moment = "variance" of P(x):

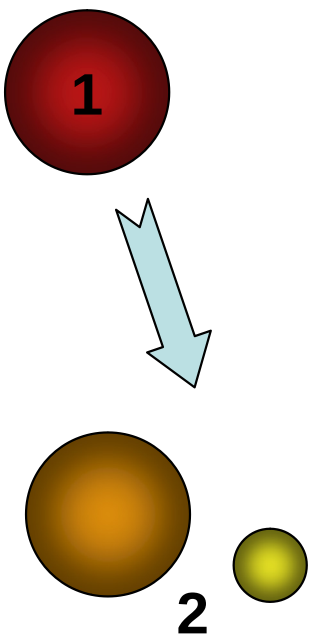
$$\sigma_x^2 = <x^2 - <x^2> >$$

Uncertainty and Statistics

Nucleus is a quantal system described by a wave function ψ(x,...;t)

(x,...;t) are the degrees of freedom of the system and time.

Probability density (e.g., for x, integrate over others)



$$\frac{dP(x,t)}{dx} = |\psi(x,t)|^2$$

Normalization

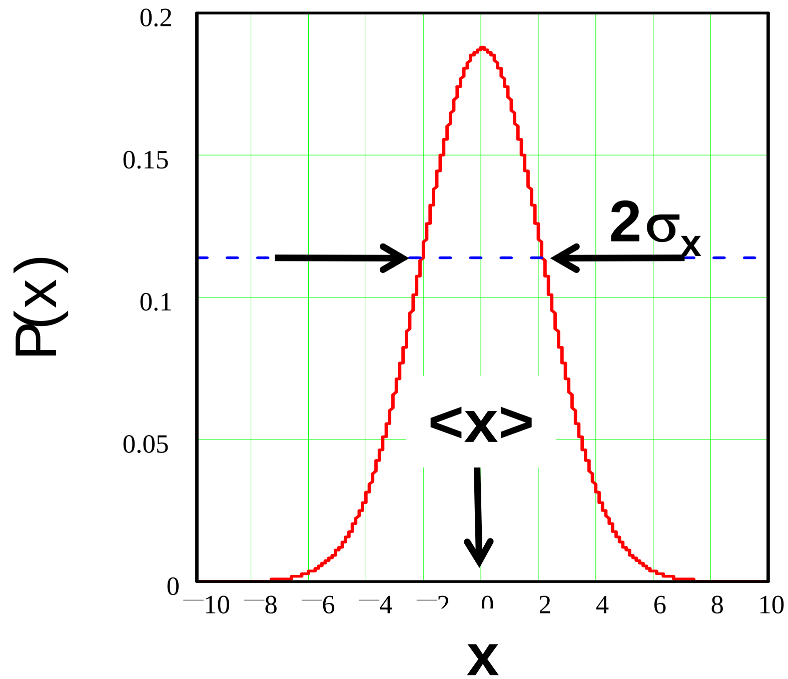
$$P(\cancel{x},t) = \int_{-\infty}^{+\infty} dx \frac{dP(x,t)}{dx} = \int_{-\infty}^{+\infty} dx |\psi(x,t)|^2 = 1$$

$$\textit{Transition between states } 1 \to 2,\ \Gamma \approx \frac{\hbar}{2\pi} |\langle M_{12} \rangle|^2 \rho(E)$$

$$\frac{dP_1(x,t)}{dx} = |\psi(x)|^2 e^{-\frac{\Gamma}{\hbar}\cdot t} \propto e^{-\lambda \cdot t} \quad \textit{state } 1 \textit{ disappears}$$

Probability rate λ for disappearance (decay of 1) can vary over many orders of magnitude → no certainty

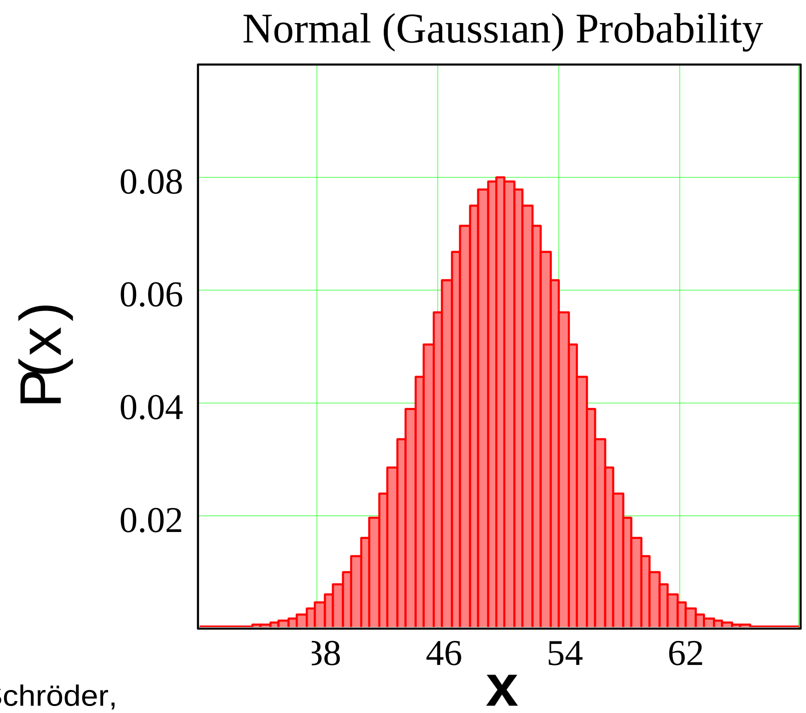
The Normal Distribution



Continuous function or discrete distribution (over bins)

$$\boxed{P(x)=\frac{1}{\sqrt{2\pi\sigma_X^2}}\cdot\exp\left\{-\frac{\left(x-\langle x\rangle\right)^2}{2\sigma_X^2}\right\}}$$

$$\Gamma_{FWHM}=2\sigma_x\cdot\sqrt{2\ln 2}=2.35\cdot\sigma_x$$



Normalized probability

$$P(x<x_1)=\frac{1}{\sqrt{2\pi\sigma_X^2}}\cdot\int_{-\infty}^{x_1}dx\exp\left\{-\frac{\left(x-\langle x\rangle\right)^2}{2\sigma_X^2}\right\}$$

Experimental Mean Counts and Variance

What can be measured: ensemble (sampling) averages (expectation values) and uncertainties

n	n-<n>	(n-<n>)^2
36076	129.6	16796.16
35753	-193.4	37403.56
35907	-39.4	1552.36
36116	169.6	28764.16
35884	-62.4	3893.76
36136	189.6	35948.16
35741	-205.4	42189.16
35640	-306.4	93880.96
36124	177.6	31541.76
36087	140.6	19768.36
35946	-1.5E-12	**3463.76**
<n>	**<n-<n>>**	**$\sigma^2{}_n$**

^{236}U (0.25mg) source, count α particles emitted during N = 10 time intervals (samples @1 min). λ =??

Average count n in a sample:

$$\langle n\rangle = \frac{1}{N}\cdot\sum_{i=1}^{N} n_i \quad (\neq \langle n\rangle_{population}\ unknown)$$

Variance of n in the individual samples

$$s^2 = \sigma^2 = \frac{1}{N-1}\cdot\sum_{i=1}^{N}(n_i - \langle n\rangle)^2$$

Variance ("error") of the sample average $\langle n\rangle \neq \langle n\rangle_{population}$

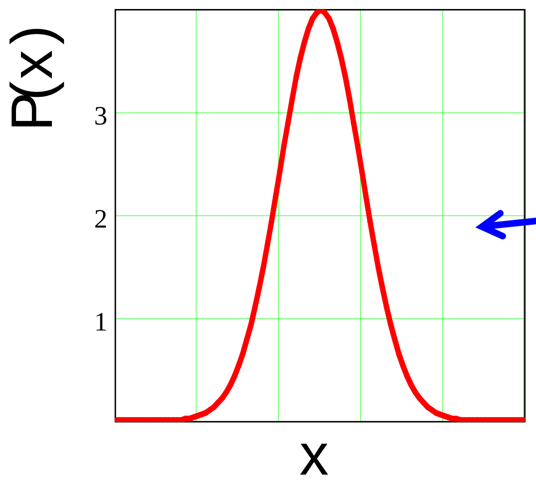
$$\sigma_n^2 = \frac{s^2}{N} = \frac{1}{N(N-1)}\cdot\sum_{i=1}^{N}(n_i - \langle n\rangle)^2$$

$$Std.deviation: \sigma_n = \sqrt{\sigma_n^2} \approx \sqrt{<n>/N} = 59$$

$$Result: \langle n\rangle \approx \langle n\rangle_{pop} = (35496 \pm 59)\,\mathrm{min}^{-1}$$

same for 1 sample
10 times as large

Sample Statistics

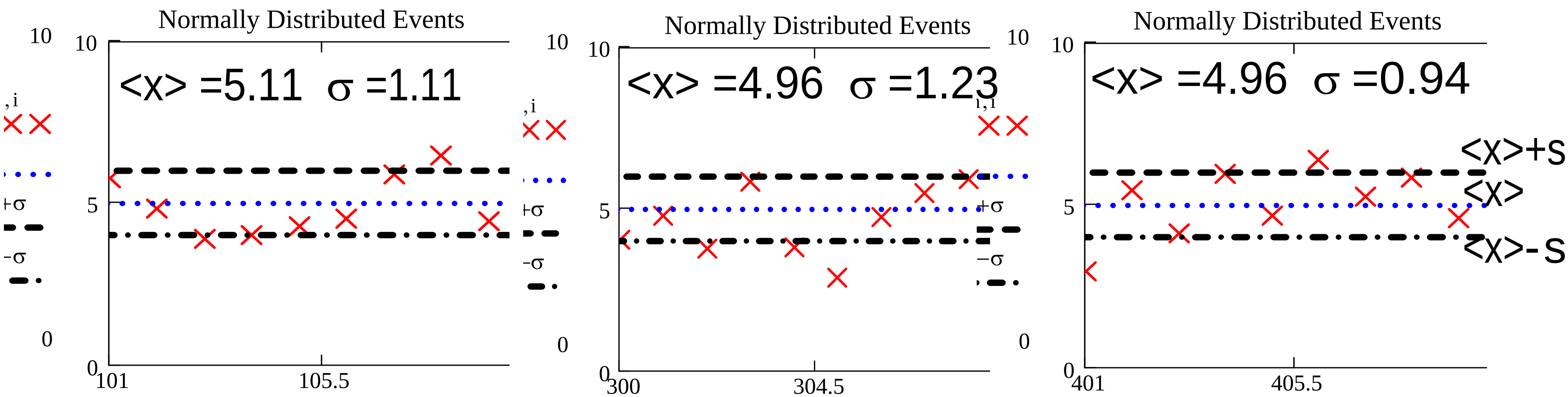


Assume true population distribution for variable x

$$P(x)=\frac{1}{\sqrt{2\pi v_X^2}}\cdot\exp\left\{-\frac{\left(x-\langle x\rangle_{pop}\right)^2}{2v_X^2}\right\}$$

with true ("population") mean $\langle x\rangle_{pop}$ = 5.0, ν_x =1.0

Results of 3 samples (measurements):



Sample average <x>= (5.11+4.96+4.96)/3 = 5.01

Sample variance $s^2=\sigma^2$ = $[(5.11-5.01)^2+2(4.96-5.01)^2]/2$ = 0.01 **s = 0.0075**

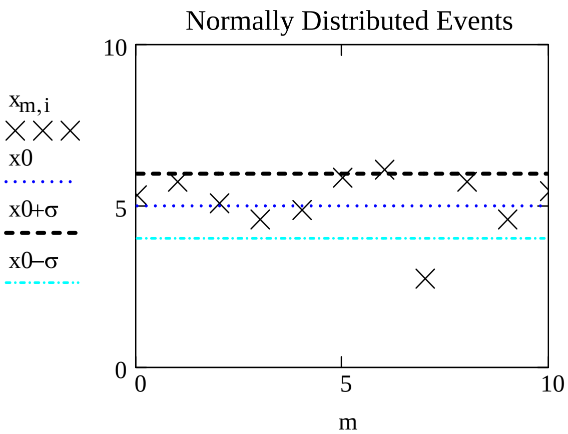
σ_x^2 = 0.0075/3 = 0.0025 σ_x = 0.05 → Result: $\langle x\rangle_{pop} \approx 5.01 \pm 0.05$

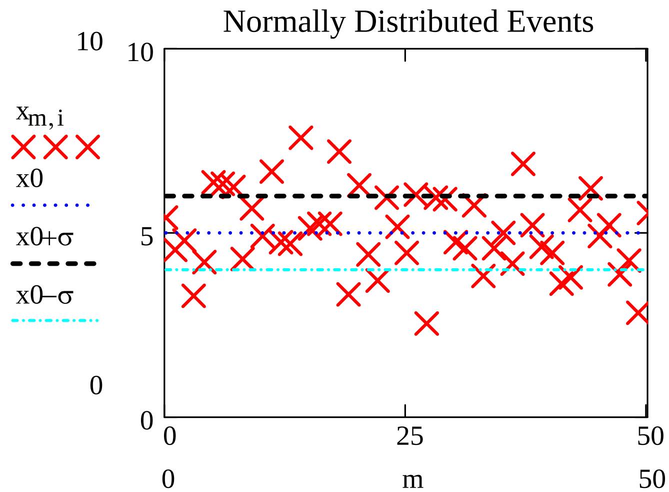
Example

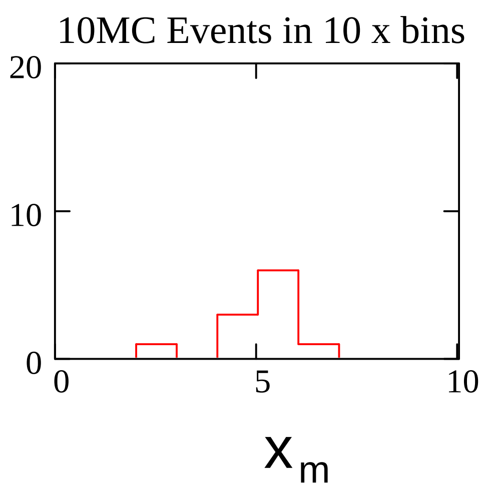
Sample size

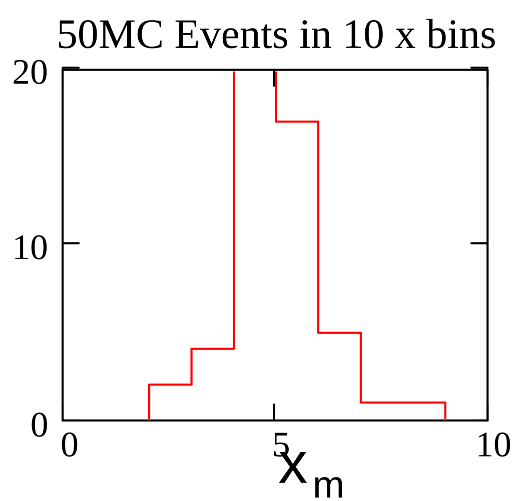
n = 10

n = 50









The larger the sample, the narrower the distribution of x values, the more it approaches the true Gaussian (normal) distribution.

Central-Limit Theorem

The means (averages) of different samples in the previous example cluster together closely. → general property:

The distribution of the sample means approaches a Gaussian normal distribution, if the size n of the sample increases, regardless of the form of the original (population) distribution.

The average of a distribution does not contain information on the shape of the distribution.

The average of any truly random sample of a population is already close to the true population average. Considering many samples, or large samples, narrows the choices. The Gaussian width becomes narrow. The standard error of the mean decreases as the sample size increases.

W. Udo Schröder, 2004

Binomial Distribution

Integer random variable m = number of events, out of N total, of a given type, e.g., decay of m (from a sample of N) radioactive nuclei, or detection of m (out of N) photons arriving at detector.
p = probability for a (one) success (decay of one nucleus, detection of one photon)

Choose an arbitrary sample of m trials out of N trials
$\mathbf{p^m}$ = probability for **at least** m successes
$\mathbf{(1-p)^{N-m}}$ = probability for N-m failures (survivals, escaping detection)
Probability for exactly m successes out of a total of N trials

$$P(m) \propto p^m (1-p)^{N-m}$$

How many ways can m events be chosen out of N ? → Binomial coefficient

$$\binom{N}{m} = \frac{N!}{m!(N-m)!} = \frac{(N-m+1)\cdots N}{1\cdots m}$$

Total probability (success rate) for any sample of m events:

$$P_{binomial}(m) = \binom{N}{m} p^m (1-p)^{N-m}$$

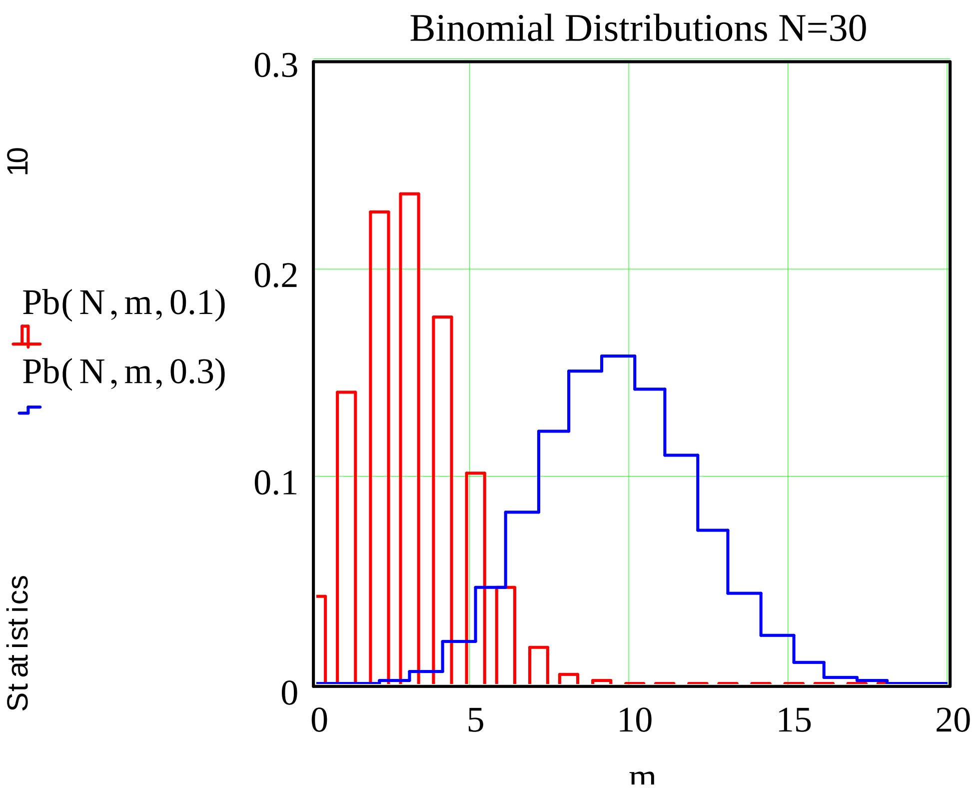
Moments and Limits

$$P_{binomial}(m) = \binom{N}{m} p^m (1-p)^{N-m}$$

Probability for m "successes" out of N trials, individual probability p

Normalization

$$1 = \sum_{m=0}^{N} P_{binomial}(m) = \sum_{m=0}^{N} \binom{N}{m} p^m (1-p)^{N-m}$$



Mean and variance

$$\bar{m} = N \cdot p \quad and \quad \sigma_m^2 = N \cdot p(1-p)$$

$$\frac{\sigma_m}{\bar{m}} = \frac{\sqrt{N \cdot p(1-p)}}{N \cdot p} \sim \frac{1}{\sqrt{N}}$$

Distributions for N=30 and
p=0.1 → p=0.3
Poisson → Gaussian

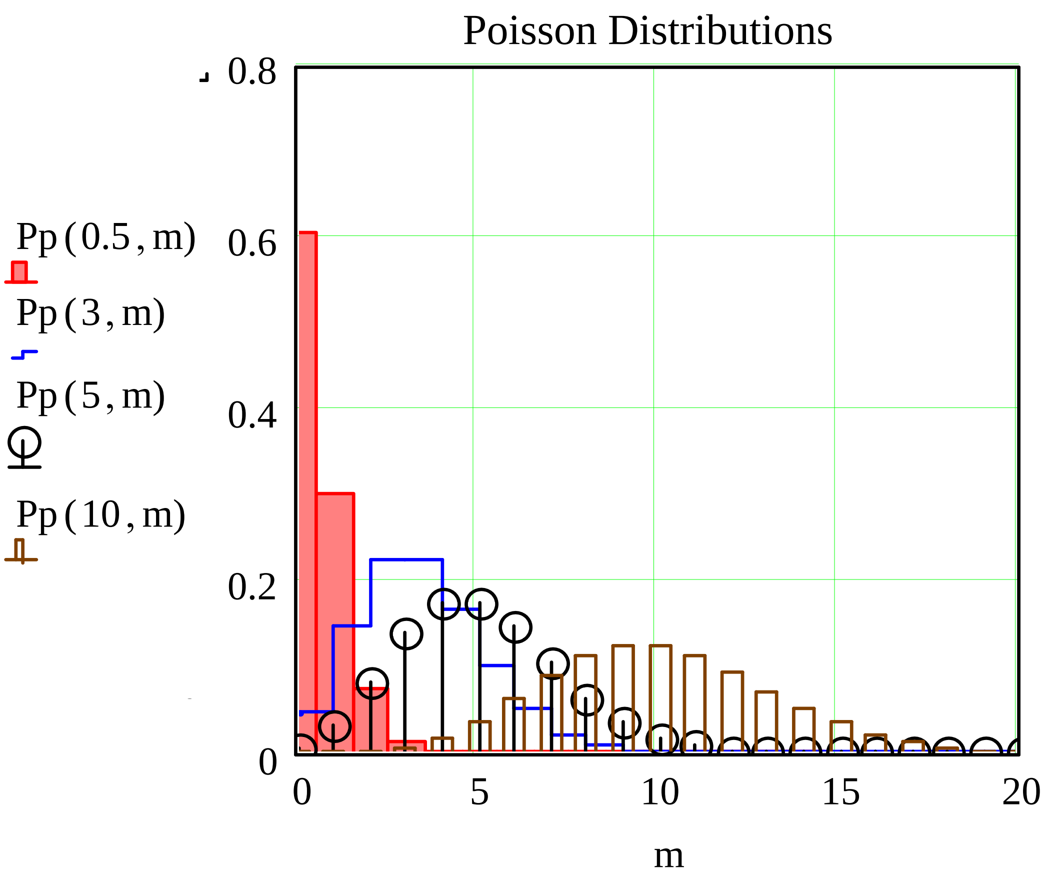
$$\lim_{N\to\infty} P_{binomial}(m) = \frac{1}{\sqrt{2\pi\sigma_m^2}} \cdot \exp\left\{-\frac{\left(x - \langle m \rangle\right)^2}{2\sigma_m^2}\right\}$$

W. Udo Schröder, 2004

Poisson Probability Distribution

Results from binomial distribution in the limit of small p and large N ($N \cdot p > 0$)

$$\lim_{\substack{p\to 0 \\ and\ N\to\infty}} P_{binomial}(N,m) = P_{Poisson}(\mu,m)$$



Probability for observing m events when average is $\langle m \rangle = \mu$

$$P_{Poisson}(\mu,m) = \frac{\mu^m \cdot e^{-\mu}}{m!} \quad m=0,1,2,\ldots$$

$\mu = \langle m \rangle = N \cdot p$ and $\sigma^2 = \mu$

is the mean, the average number of successes in N trials.

Observe N counts (events) → uncertainty is $\sigma = \sqrt{\mu}$

Unlike the binomial distribution, the Poisson distribution does not depend explicitly on p or N !

For large N, p:

Poisson → Gaussian (Normal Distribution)

W. Udo Schröder, 2004

Moments of Transition Probabilities

$$N : 0.25mg = 0.25mg \cdot \frac{6.022 \cdot 10^{23}}{236g} = 6.38 \cdot 10^{17}$$

$$p = \frac{\langle n \rangle}{N} = \frac{3.5946 \cdot 10^{4} \min^{-1}}{6.38 \cdot 10^{17}} = 5.6362 \cdot 10^{-14} \min^{-1}$$

Probability for decay (decay rate per nucleus):

$$p = \lambda = 5.6362 \cdot 10^{-14} \min^{-1}$$

corresponds to "half life" $t_{1/2} = 2.34 \cdot 10^{7}\, a$

Small probability for process, but many trials ($n_0 = 6.38 \cdot 10^{17}$)

→→→ $0 < \mathbf{n_0 \cdot \lambda < \infty}$

Statistical process follows a Poisson distribution: n="random"

Different statistical distributions: Binomial, Poisson, Gaussian

Radioactive Decay as Poisson Process

Useful when only a mean count rate is known: decay, background counts, or reaction.

^{137}Cs unstable isotope, decays

$t_{1/2} = 27$ years $\rightarrow$ **p** $= \ln 2/27 = 0.026/a =$ **$8.2 \cdot 10^{-10} s^{-1}$** $\rightarrow 0$

Sample of 1 μg: N = 10^{15} nuclei (=trials for decay)

How many will decay?

$$\mu = N \cdot p = \mathbf{8.2 \cdot 10^{+5}\, s^{-1}}$$

Count rate estimate $dN/dt = $ (**$8.2 \cdot 10^{+5} \pm 905$**) **$s^{-1}$**

estimated

Probability for m decays $P(\mu, m) =$

$$P_{Poisson}(\mu, m) = \frac{\mu^m \cdot e^{-\mu}}{m!} = \frac{(8.52 \cdot 10^5)^m \cdot e^{-8.52 \cdot 10^5}}{m!}$$

Functions of Stochastic Variables

Random independent variables $N_1, N_2, \ldots, N_n$
corresponding variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$

Function $f(N_1, N_2, \ldots, N_n)$

Gauss' law of error propagation:

$$\sigma_f = \left\{ \underbrace{\left(\frac{\partial f}{\partial N_1}\right)^2 \sigma_1^2}_{\left(\Delta f \mid_{N_2, N_3, ..}\right)^2} + \underbrace{\left(\frac{\partial f}{\partial N_2}\right)^2 \sigma_2^2}_{\left(\Delta f \mid_{N_1, N_3, ..}\right)^2} + \ldots . + \underbrace{\left(\frac{\partial f}{\partial N_n}\right)^2 \sigma_n^2}_{\left(\Delta f \mid_{N_1, N_2, .., N_{n-1}}\right)^2} \right\}^{1/2}$$

$$\left(\Delta f \mid_{N_2, N_3, ..}\right)^2 + \left(\Delta f \mid_{N_1, N_3, ..}\right)^2 + .. + \left(\Delta f \mid_{N_1, N_2, .., N_{n-1}}\right)^2$$

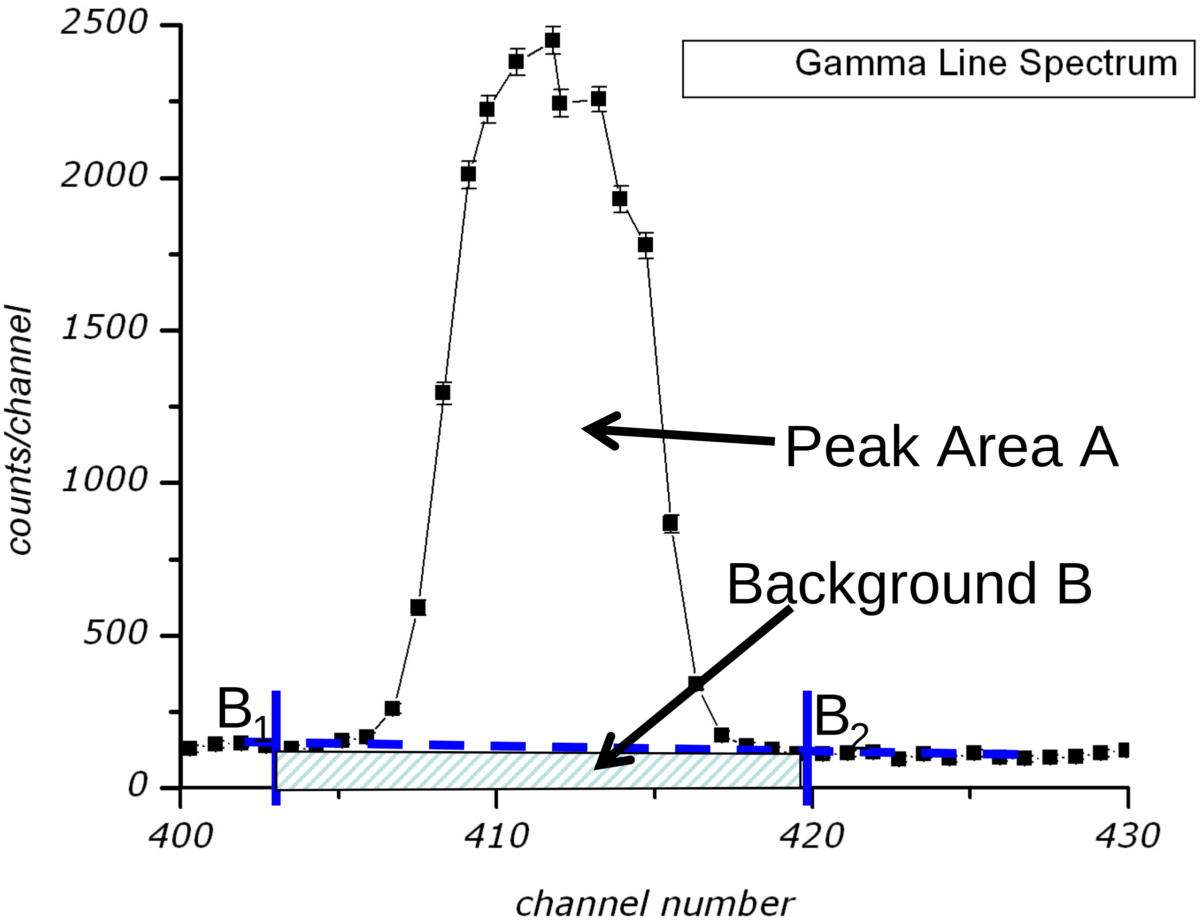
Further terms if N_i not independent (→ correlations)
Otherwise, individual component variances $(\Delta \mathbf{f})^2$ add.

Example: Spectral Analysis

Adding or subtracting 2 Poisson distributed numbers N_1 and N_2:

Variances always add $N := \left[N_1 \pm \sqrt{N_1}\right] \pm \left[N_2 \pm \sqrt{N_2}\right] =$

$$= \left(N_1 \pm N_2\right) \pm \sqrt{N_1 + N_2}$$



Analyze peak in range channels $c_1 - c_2$: beginning of background left and right of peak $n = c_1 - c_2 + 1$.

Total area = N_{12}

$N(c_1)=B_1$, $N(c_2)=B_2$,

Linear background

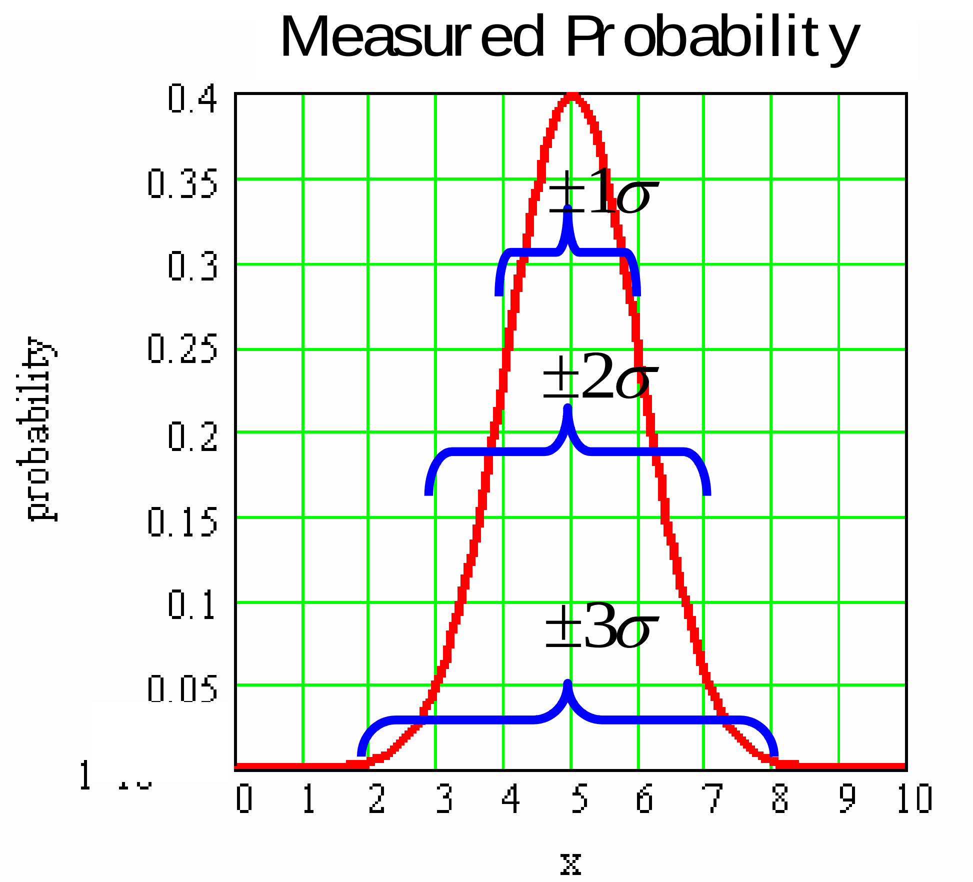
$B = n(B_1+B_2)/2$

Peak Area A:

$$A = N_{12} - n \cdot \left(B_1 + B_2\right)/2$$

$$\sigma_A = \sqrt{N_{12} + n^2 \cdot \left(B_1 + B_2\right)/4}$$

Confidence Level



Assume normally distributed observable x:

$$P(x)=\frac{1}{\sqrt{2\pi v_{pop}^2}}\cdot \exp\left\{-\frac{\left(x-\langle x\rangle_{pop}\right)^2}{2v_{pop}^2}\right\}$$

Sample distribution with data set

→ observed average <x> and std. error σ approximate population.
Confidence level CL
(Central Confidence Interval):

$$P(|<x_{pop}>-<x>|<\delta)\approx\frac{2}{\sqrt{2\pi\sigma^2}}\cdot\int_{<x>}^{<x>+\delta}\exp\left\{-\frac{\left(x-\langle x\rangle\right)^2}{2\sigma^2}\right\}dx=CL$$

With confidence level CL (probability in %), the true value $<x_{pop}>$ differs by less than $\delta = n\sigma$ from measured average.

$CL(\delta=1\sigma)=68.3\%$ $CL(\delta=2\sigma)=95.4\%$

$CL(\delta=3\sigma)=99.7\%$

Trustworthy exptl. results quote ±3σ error bars!

Setting Confidence Limits

Example: Search for rare decay with decay rate λ, observe no counts within time Δt. Decay probability law dP/dt ∝ exp{- λ·t}. Law is symmetric in λ and t : → P(λ,t)

no counts in Δt

$$\Delta P\left(0\,|\,\lambda\right) = \Delta t \cdot e^{-\lambda\cdot\Delta t} \quad with \quad \int_0^\infty \Delta t \cdot e^{-\lambda\cdot\Delta t}\, d\lambda = 1$$ normalized P

$$P(\lambda \le \lambda_0) = \int_0^{\lambda_0} \Delta t \cdot e^{-\lambda\cdot\Delta t}\, d\lambda = 1 - e^{-\lambda_0\cdot\Delta t}$$ normalized P

$$\lambda_0 = \frac{-1}{\Delta t}\cdot \ln[1 - P(\lambda \le \lambda_0)] = \frac{-1}{\Delta t}\cdot \ln[1 - CL] > 0$$ Upper limit

The higher the confidence level CL (0 ≤ CL ≤ 1), the larger the upper limit for a given time Δt inspected. Reduce limit by measuring for longer period.

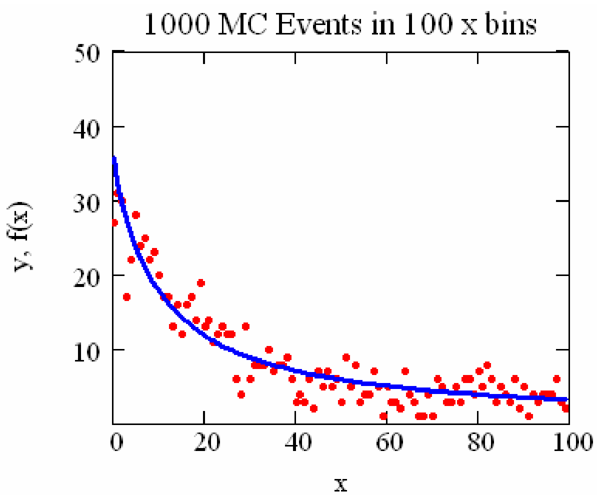
W. Udo Schröder, 2004

Maximum Likelihood

Measurement of correlations between observables y and x: $\{x_i, y_i \mid i=1-N\}$
Hypothesis: $y(x) = f(c_1,\ldots,c_m; x)$. Parameters defining f: $\{c_1,\ldots,c_m\}$
n_{dof}=N-m degrees of freedom for a "fit" of the data with f.

$$P_i(c_1,..,c_m;x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot \exp\left\{ -\frac{\left(y_i - f(c_1,..,c_m;x_i)\right)^2}{2\sigma_i^2} \right\}$$

for every data point



Maximize simultaneous probability

$$P(c_1,..,c_m) = \prod_{i=1}^{N} P_i(c_1,..,c_m;x) =$$

$$= \prod_{j=1}^{N} \frac{1}{\sqrt{2\pi\sigma_j^2}} \cdot \exp\left\{ -\sum_{i=1}^{N} \frac{(\Delta y_i)^2}{2\sigma_i^2} \right\}$$

$$\boxed{\chi^2(c_1,..,c_m) := \sum_{i=1}^{N} \frac{(\Delta y_i)^2}{\sigma_i^2} = \sum_{i=1}^{N} \frac{\left(y_i - f(c_1,..,c_m;x_i)\right)^2}{\sigma_i^2}}$$

Minimize chi-squared by varying $\{c_1,\ldots,c_m\}$: $\partial\chi^2 / \partial c_i = 0$

When is χ^2 as good as can be?

Minimizing χ^2

Example: linear fit f(a,b;x) = a + b·x to data set $\{x_i, y_i, \sigma_i\}$

Minimize: $$\chi^2(a,b) := \sum_{i=1}^{N} \frac{(\Delta y_i)^2}{\sigma_i^2} = \sum_{i=1}^{N} \frac{(y_i - a - bx_i)^2}{\sigma_i^2}$$

$$0 = \frac{\partial}{\partial a}\chi^2(a,b) = \frac{\partial}{\partial a}\sum_{i=1}^{N} \frac{(y_i - a - bx_i)^2}{\sigma_i^2} = -\sum_{i=1}^{N} \frac{2(y_i - a - bx_i)}{\sigma_i^2}$$

$$0 = \frac{\partial}{\partial b}\chi^2(a,b) = -\sum_{i=1}^{N} \frac{2x_i(y_i - a - bx_i)}{\sigma_i^2}$$

Equivalent to solving system of linear equations

$$a\sum_{i=1}^{N}\frac{1}{\sigma_i^2} + b\sum_{i=1}^{N}\frac{x_i}{\sigma_i^2} = \sum_{i=1}^{N}\frac{y_i}{\sigma_i^2}$$

$$a\sum_{i=1}^{N}\frac{x_i}{\sigma_i^2} + b\sum_{i=1}^{N}\frac{x_i^2}{\sigma_i^2} = \sum_{i=1}^{N}\frac{x_i y_i}{\sigma_i^2}$$

$$a d_{11} + b d_{12} = c_1$$
$$a d_{21} + b d_{22} = c_2$$

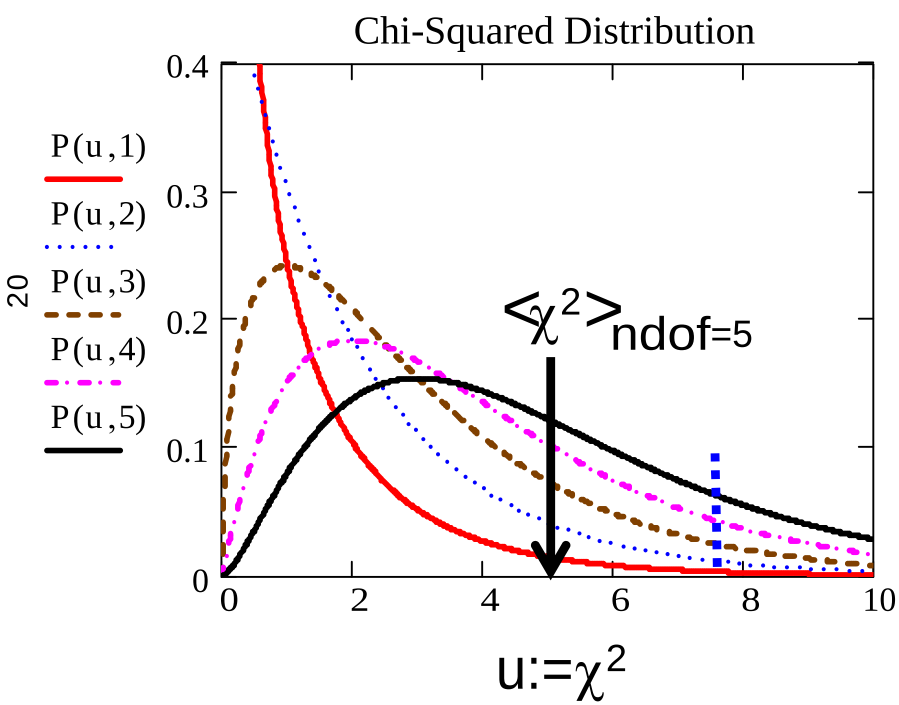
$$D = \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix}$$

$$a = \frac{1}{D}\begin{vmatrix} c_1 & d_{12} \\ c_2 & d_{22} \end{vmatrix} \qquad b = \frac{1}{D}\begin{vmatrix} d_{11} & c_1 \\ d_{21} & c_2 \end{vmatrix}$$

$$\sigma_a^2 = \frac{1}{D}\sum\frac{x_i^2}{\sigma_i^2} \qquad \sigma_b^2 = \frac{1}{D}\sum\frac{1}{\sigma_i^2}$$

Distribution of Chi-Squareds

Distribution of possible χ^2 for data sets distributed normally about a theoretical expectation (function) with n_{dof} degrees of freedom:



$$\boxed{P_{n_{dof}}(\chi^2)d\chi^2 = \frac{\left(\chi^2\right)^{n_{dof}/2-1} e^{-\chi^2/2}}{2^{n_{dof}/2}\Gamma\left(n_{dof}/2\right)}d\chi^2}$$

$$\underline{\left\langle \chi^2 \right\rangle = n_{dof}} \qquad \underline{\sigma^2_{\chi^2} = 2n_{dof}} \qquad n_{dof} \gg 1$$

$\Gamma(n) = (n-1)! =$ Stirling's formula

$$= 2.507 e^{-n} n^{n-1/2}(1 + 0.0833/n)$$

$$P(\chi^2, n_{dof}) = \int_{\chi^2}^{\infty} P_{n_{dof}}(x)dx$$

Should be P $\approx$ 0.5 for a reasonable fit

Reduced χ^2:

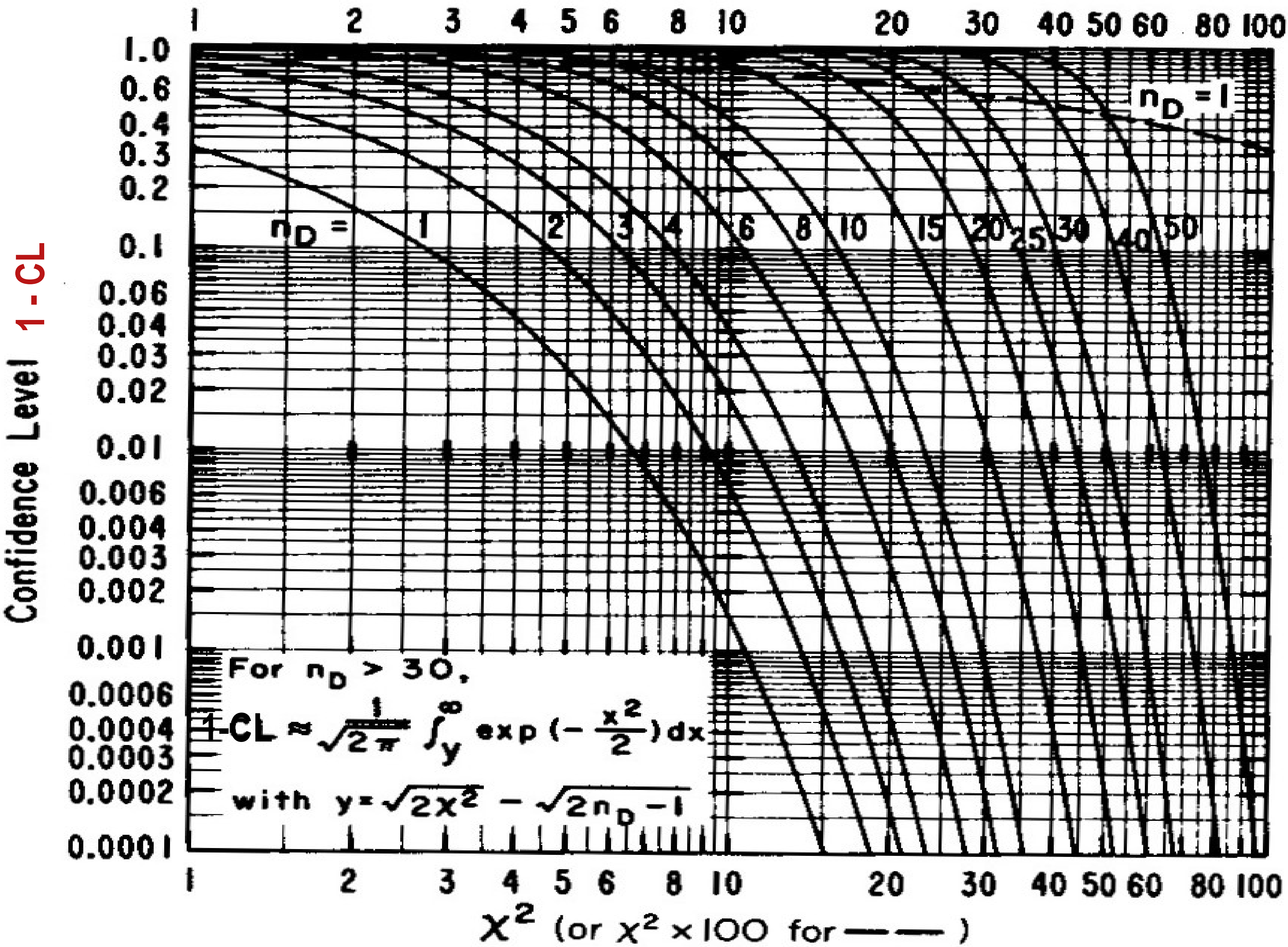
$$\chi_r^2 = \chi^2 / n_{dof} = \chi^2 / (N - m - 1)$$

For

$$0 \le \chi_r^2 < 1.5 \quad \rightarrow \quad Confidence \ge 50\%$$

W. Udo Schröder, 2004

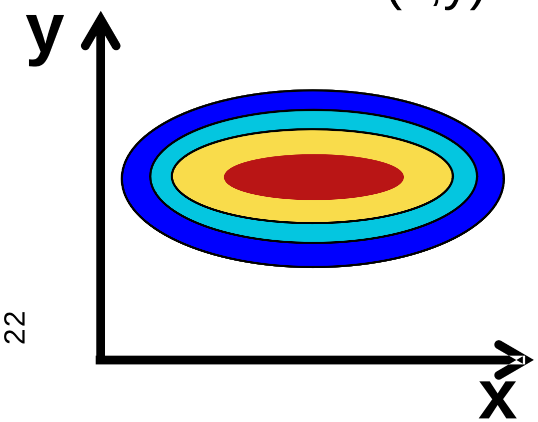
CL for χ^2- Distributions



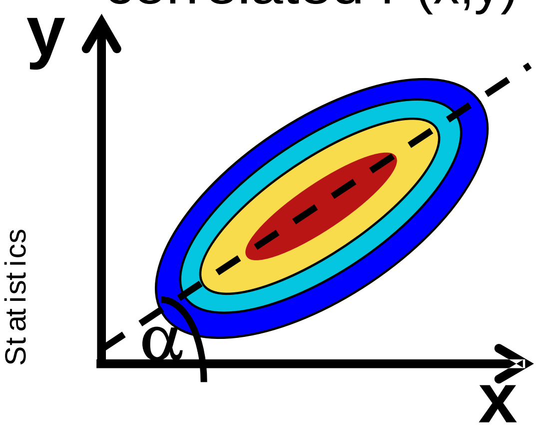
Correlations in Data Sets

Correlations within data set. Example: y_i small whenever x_i small

uncorrel. P(x,y)



correlated P(x,y)



$$P_{unc}(x,y) = P(x)\cdot P(y) =$$

$$= \frac{1}{\sqrt{4\pi^2\sigma_x^2}}\cdot\exp\left\{-\left[\frac{\left(x-\langle x\rangle\right)^2}{2\sigma_x^2}+\frac{\left(y-\langle y\rangle\right)^2}{2\sigma_y^2}\right]\right\}$$

$$P_{corr}(x,y) = \frac{1}{2\pi\sqrt{\sigma_x^2\sigma_y^2-\sigma_{xy}^2}}\cdot$$

$$\cdot\exp\left\{-\frac{\left(x-\langle x\rangle\right)^2\sigma_x^2+\left(y-\langle y\rangle\right)^2\sigma_y^2-2\left(x-\langle x\rangle\right)\left(y-\langle y\rangle\right)}{2\left(\sigma_x^2\sigma_y^2-\sigma_{xy}^2\right)}\right\}$$

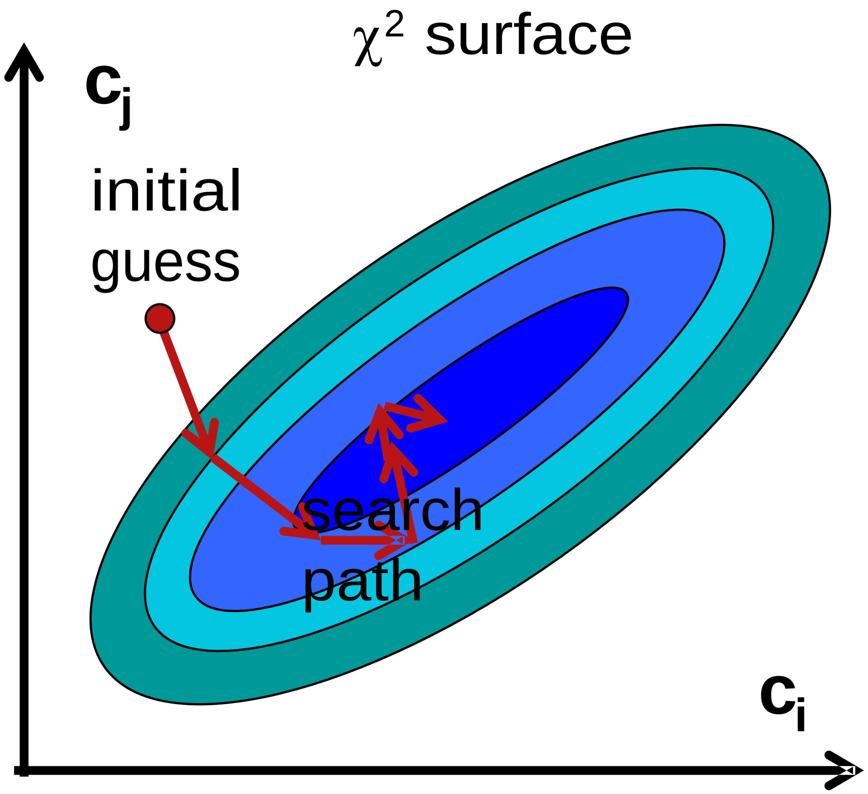
$$\sigma_{xy} = \int\left(x-\langle x\rangle\right)\left(y-\langle y\rangle\right)P(x,y)dxdy \quad covariance$$

$$\cot\alpha = \frac{1}{\sigma_{xy}}\left\{\frac{\sigma_x^2-\sigma_y^2}{2}+\sqrt{\frac{\left(\sigma_x^2-\sigma_y^2\right)^2}{4}+\sigma_{xy}^2}\right\}$$

$$r_{xy} = \sigma_{xy}/\left(\sigma_x\sigma_y\right) \quad correlation\ coefficient \quad -1\le r_{xy}\le 1$$

W. Udo Schröder, 2004

Multivariate Correlations

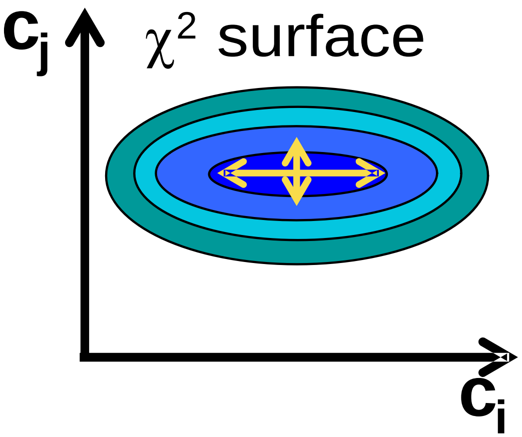


Different search strategies → software packages LINFIT, MINUIT,….

Correlations in Data Sets

uncertainties of deduced most likely parameters c_i (e.g., a, b for linear fit) depend on depth/shallowness and shape of the χ^2 surface

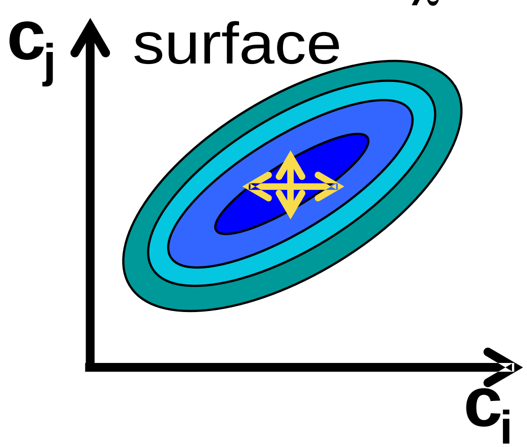
uncorrelated χ^2 surface



$$P(x,y) = P(x)\cdot P(y) =$$

$$= \frac{1}{\sqrt{4\pi^2\sigma_x^2}}\cdot \exp\left\{-\left[\frac{\left(x-\langle x\rangle\right)^2}{2\sigma_x^2}+\frac{\left(y-\langle y\rangle\right)^2}{2\sigma_y^2}\right]\right\}$$

correlated χ^2 surface



$$r_{xy} := \frac{\sum\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)}{\sqrt{\sum\left(x_i-\overline{x}\right)^2\cdot\sum\left(y_j-\overline{y}\right)^2}} \qquad -1 \le r_{xy} \le +1$$

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x\sigma_y};$$

$$\sigma_{xy} := \frac{1}{n-1}\left(\sum x_i y_i - \frac{1}{n}\sum x_i\cdot\sum y_j\right) \; covariance$$