The quest for high Luminosity

- Luminosity formula:
  - f is fixed by the collider radius
  - High N_1 and N_2 and n_b
  - Low σ_x, σ_y
- Integrated Luminosity L_{int}: $[L_{int}] = l^{-2}$ ➔ $\text{nbarn}^{-1} = 10^{33}\ \text{cm}^{-2}$
- Problems:
  - Increase number of particles / bunch ? ➔ beam-beam effects generate instabilities;
  - Increase number of bunches reduces the inter-bunch time T_{BC};
  - Decrease σ_x and σ_y ? (see next slides on beam dynamics).

$$L = n_b f \frac{N_1 N_2}{4\pi\sigma_x\sigma_y} = \frac{I_1 I_2}{4\pi n_b f e^2 \sigma_x \sigma_y}$$

$$L_{\text{int}} = \int_{Trun} L(t)\,dt$$

$$T_{BC} = \frac{1}{n_b f}$$

The pile-up

- How many interactions take place per bunch crossing ? It depends on:
  - Interaction rate that in turns depends on:
    - Luminosity
    - Total Cross-section
  - Bunch crossing rate that depends on
    - Bunch frequency
    - Number of bunches circulating
- Pile-up $\boldsymbol{\mu}$ = average number of interactions per bunch-crossing

$$\mu = \dot{n} T_{BC} = \frac{L\sigma_{tot}}{f n_b}$$

Comparison: e^+e^- vs pp

- DAFNE: e^+e^- @ 1 GeV c.o.m. energy, σ_{tot}=5 μb, $L=10^{33}cm^{-2}s^{-1}$, n_b=120, f=c/100 m = 3 MHz

  ➔ T_{BC}= , μ=

- LHC: pp @ 13 TeV c.o.m. energy, σ_{tot}=70 mb, $L=10^{34}cm^{-2}s^{-1}$, n_b=3000, f=c/27 km = 10 kHz

  ➔ T_{BC}= , μ=

Heavy Ion collisions.

- Lead nuclei @ LHC:
  - Z=82, A=208, M ≈ 195 GeV
  - ΔE_K = ZeV (proton × Z)
  - p= ZeRB (proton × Z)
  - ➔ E_{Pb} = 574 TeV=82 × 7 TeV
  - ➔ E_{Pb}/Nucleon = 574/A = 2.77 TeV
  - $\sqrt{s_{NN}}$=5.54 TeV
- Luminosity: ≈ 10^{27} $cm^{-2}s^{-1}$
- n_b = 600
- $N_1=N_2=7\times10^7$ ions/bunch
- Heavy ions program @ RHIC
  - Au, Cu, U ions up to 100 GeV/nucleon
  - Luminosity ≈$10^{28}\div10^{29}$ $cm^{-2}s^{-1}$
- Cross-sections:
  - $\sigma_{pp}\approx$ 70 mb
  - $\sigma_{pPb}\approx\sigma_{pp}\times A^{2/3}$ $(\approx\sigma_{pp}\times R_{Nuc}^{\ 2})$
  - $\sigma_{PbPb}\approx\sigma_{pp}\times N_{coll}\approx$ 10 barn!
- How much is the pile-up ?

Proposed exercises

Consider the parameters of the three accelerators:

- LHC: protons, R = 4.3 km, E_{max} = 7 TeV, T_{BC} = 25 ns;
- LEP: electrons, R = 4.3 km, E_{max} = 100 GeV, T_{BC} = 22 μs;
- DAFNE: electrons, R = 15 m, E_{max} = 500 MeV, T_{BC} = 2.7 ns;

Evaluate for each accelerator the following quantities: the revolution frequency f; the number of bunches n_b; the minimum value of the magnetic field B_{min} required to hold the particles in orbit. From the luminosity and current profile plots shown as examples in the course slides, determine for DAFNE and LHC, the products $\sigma_x \times \sigma_y$

Design a pp machine at $\sqrt{s}$ = 40 TeV and $L = 10^{36}$ cm^{-2}s^{-1}. Which values of σ_x and σ_y are needed ? The following limits have to be respected:

- B < 5T
- N_1, $N_2 < 10^{11}$/bunch
- $T_{BC} > 10$ ns

Evaluate the maximum $\sqrt{s_{NN}}$ that can be obtained at LHC for Cu-Cu and Pb-Pb collisions respectively.

Evaluate the value of $\sqrt{s_{NN}}$ for Au-Au collisions if the energy of the Au ions is 10.5 TeV. In case these collisions are done at RHIC for which value of the luminosity the pile-up becomes of order 1 ? (RHIC circumference = 3.834 km, n_b=111)

Analysis of event distributions: the fit

(i) to compare the distributions with expectations from theories, and (ii) to extract from them physical quantities of interest like masses, widths, couplings, spins and so on. We call **fit** the method to do both these important things.

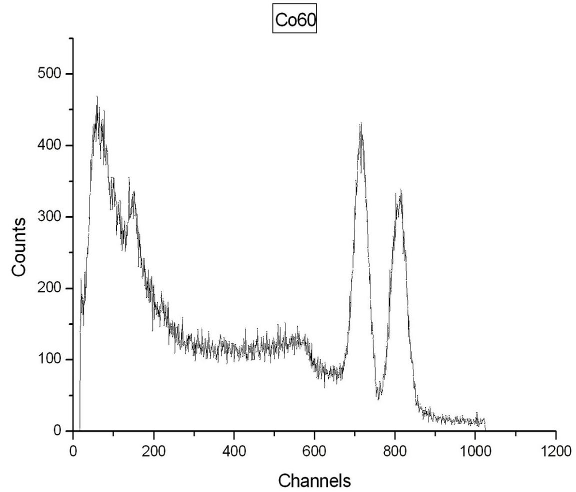
Analysis of event distributions: the fit

(1) First of all we have to define the hypothesis. It can be the theoretical function $y(x/\underline{\theta})$, x being the variable or the set of variables, and $\underline{\theta}$ a set of K **parameters** . K could be even 0, in this case the theory makes an "absolute prediction" and there is no need to adjust parameters to compare it to theory.

(2) Then we have to define a **test statistics** t, that is a variable depending on the data that, if the hypothesis is correct, has a known distribution function (in the following we use **pdf** to indicate probability distribution functions). The meaning of this pdf is the following: if we repeat the experiment many times and if every time we evaluate t, if the hypothesis is correct the histogram of the sample statistics will follow the pdf within the statistical errors of the sample.

(3) Finally we do the experiment. In case the theory depends on few parameters, we adjust the parameters in such a way to get the best possible agreement between data and theory. From this we obtain the **estimates** of the parameters with their uncertainties. We evaluate then the actual value of t, let's call it t^* from the data after parameter adjustment, and see if in the t pdf this value corresponds to a region of high or low probability. In case it is in a region of high probability, it's likely that the theory is correct, so that we conclude that the experiment **corroborates** the theory. In case it corresponds to a region of low probability it's unlikely that the theory is correct, so that we say that the experiment **falsifies** the theory, or, in other words, that we have not found any parameter region that allows an acceptable agreement.

Choice of test statistics: binned data



Histogram: $\sum_{i=1}^{M} n_i = N$

Theory: y=y(x/$\underline{\theta}$) θ_i, i=1….K

Prediction of the theory in bin i:

1) Value of the function at the center $\overline{x}_i$ of the bin multiplied by the bin width δx (note: [y]=[dN/dx])

$$y_i = y(\overline{x}_i/\underline{\theta})\delta x$$

2) or more exactly integrating y over the bin i

$$y_i = \int_{\overline{x}_i-\delta x/2}^{\overline{x}_i+\delta x/2} y(x/\underline{\theta})dx$$

The predicted total number of events is: $\sum_{i=1}^{M} y_i = N_0$

The two definitions are equivalent in the limit of small bin size wrt to the typical scale of variations in the distribution

Choice of test statistics: binned data

Which statistics for the n_i data in the histogram?
two possibilities:

- We repeat the experiment holding the total number of events N fixed. In this case n_i has a multinomial distribution. The joint distribution of the n_i, with i=1,...,M is

$$p(n_1, ..n_M) = N! \prod_{i=1}^{M} \frac{p_i^{n_i}}{n_i!}$$

where p_i is the probability associated to the bin i. Notice that the joint distribution cannot be factorized in a product of single bin probability distributions, since the fixed value of events N determines a correlation between the bin contents.

$$\begin{aligned} E[n_i] &= Np_i \\ Var[n_i] &= Np_i(1-p_i) \\ cov[n_i, n_j] &= -Np_ip_j \end{aligned}$$

Correlation negligible for events distributed over a large number of bins

Which statistics for the n_i data in the histogram?
two possibilities:

- We repeat the experiment holding fixed the integrated luminosity or the observation time of the experiment. In this case N is not fixed and fluctuates in general between an experiment and another. The n_i are independent and have poissonian distributions:

$$p(n_1, ..n_M) = \prod_{i=1}^{M} \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}$$

where λ_i is the expected counting in each bin.

$$\begin{aligned} E[n_i] &= \lambda_i \\ Var[n_i] &= \lambda_i \\ cov[n_i, n_j] &= 0 \end{aligned}$$

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

Definition of the test statistics t :

Neiman χ^2

$$\chi^2_N = \sum_{i=1}^{M} \frac{(n_i - y_i)^2}{n_i}$$

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

Definition of the test statistics t :

Neiman χ^2

$$\chi_N^2 = \sum_{i=1}^{M} \frac{(n_i - y_i)^2}{n_i}$$

Pearson χ^2

$$\chi_P^2 = \sum_{i=1}^{M} \frac{(n_i - y_i)^2}{y_i}$$

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.

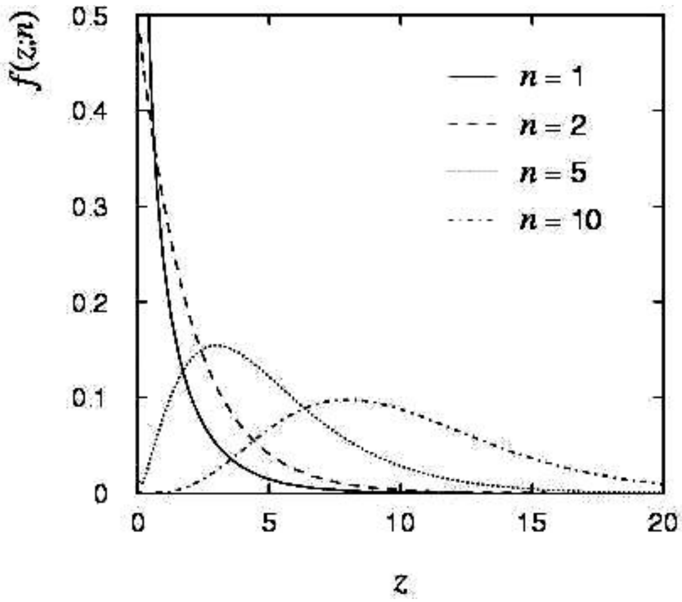
OK example: total deflection due to multiple Coulomb scattering. (Rare large angle deflections give non-Gaussian tail.)

Bad example: energy loss of charged particle traversing thin gas layer. (Rare collisions make up large fraction of energy loss, cf. Landau pdf.)

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.

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

Definition of the test statistics t :

Pearson χ^2

$$\chi^2_P = \sum_{i=1}^{M} \frac{(n_i - y_i)^2}{y_i}$$

In case of n_i being poissonian variables in the gaussian limit, the Pearson χ^2 is a statistics following a χ^2 distribution with a number of degrees of freedom equal to $M - K$. Infact we know that a χ^2 variable is the sum of the squares of standard gaussian variables, so that if eq.102 holds, this is the case for χ^2_P. However we know that the gaussian limit is reached for n_i at least above 10÷20 counts. If we have histograms with few counts, and we are far from the gaussian limit, the pdf of χ^2_P is not exactly a χ^2 so that care is needed in the result interpretation.

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

Definition of the test statistics t :

Neiman χ^2

$$\chi_N^2 = \sum_{i=1}^{M} \frac{(n_i - y_i)^2}{n_i}$$

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

Definition of the test statistics t :

Neiman χ^2

$$\chi^2_N = \sum_{i=1}^{M} \frac{(n_i - y_i)^2}{n_i}$$

The Neyman χ^2 is less well defined. In fact a χ^2 variable requires the gaussian σ in each denominator. By putting n_i we make an approximation[18]. However in case of large values of n_i to a good approximation the Neyman χ^2 has also a χ^2 distribution. A specific problem of the Neyman χ^2 is present when $n_i = 0$. But again, for low statistics histogram a different approach should be considered.

[18] The Neyman χ^2 was widely used in the past, since it makes simpler the calculation, the parameters being only in the numerator of the formula. With the present computing facilities there are no strong motivations to use it.

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

More general the test statistics t : Likelihood

$$(y_i = N_0 p_i)$$

N fixed (multinomial case)
(negligible bin correlation assumed)

$$L_m(\underline{n}/\underline{y}) = N! \prod_{i=1}^{M} \frac{p_i^{n_i}}{n_i!} = N! \prod_{i=1}^{M} \frac{y_i^{n_i}}{n_i! N_0^{n_i}}$$

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

More general the test statistics t : Likelihood

N not fixed (poisson case)

$y_i = \lambda_i$

$$L_p(\underline{n}/\underline{y}) = \prod_{i=1}^{M} \frac{e^{-y_i} y_i^{n_i}}{n_i!}$$

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

More general the test statistics t : Likelihood

N not fixed (poisson case)

$y_i=\lambda_i$

$$L_p(\underline{n}/\underline{y}) = \prod_{i=1}^{M} \frac{e^{-y_i} y_i^{n_i}}{n_i!}$$

$$L_m(\underline{n}/\underline{y}) = N! \prod_{i=1}^{M} \frac{y_i^{n_i}}{n_i! N_0^{n_i}} = \frac{N!}{N_0^N} \prod_{i=1}^{M} \frac{y_i^{n_i}}{n_i!}$$

$$L_p(\underline{n}/\underline{y}) = e^{-N_0} \prod_{i=1}^{M} \frac{y_i^{n_i}}{n_i!} = \frac{e^{-N_0} N_0^N}{N!} L_m(\underline{n}/\underline{y})$$

L_p is essentially L_m multiplied by the poissonian fluctuation of N with mean N_0

Choice of test statistics: binned data

Fit: we impose the condition $y_i = E[n_i]$

More general the test statistics t : Likelihood method

Which test statistics for the Likelihood function?

The pdf of a likelihood function in general depends on the specific problem, and can be evaluated by means of a MonteCarlo simulation of the situation we are considering (TOY MC), i.e. simulations done for different values of the parameters θ_i

Choice of test statistics: binned data

WILKS THEOREM

expectation values $\nu_i = E[n_i]$ of the contents of each bin

$$\chi^2_\lambda = -2 \ln \frac{L(\underline{n}/\underline{y})}{L(\underline{n}/\underline{\nu})}$$

has a χ^2 pdf with $M - K$ degrees of freedom in the asymptotic limit

(ν_i gaussians)

⇒ We can use Likelihood ratios as test statistics with known pdf, more general than Pearson $\chi 2$, it holds in asymp. limit but whatever is the stat. model.

Connection with the Neyman-Pearson Lemma

$$P(type - Ierrors) = 1 - \epsilon = \alpha$$

$$P(type - IIerrors) = \frac{1}{R} = \beta$$

Given the two hypotheses H_s and H_b and given a set of K discriminating variables x_1, x_2,...x_K, we can define the two "likelihoods"

(66) $$L(x_1, ..., x_K/H_s) = P(x_1, ...x_K/H_s)$$

(67) $$L(x_1, ..., x_K/H_b) = P(x_1, ...x_K/H_b)$$

equal to the probabilities to have a given set of values x_i given the two hypotheses, and the **likelihood ratio** defined as

(68) $$\lambda(x_1, ...x_K) = \frac{L(x_1, ..., x_K/H_s)}{L(x_1, ..., x_K/H_b)}$$

Neyman-Pearson Lemma:
For fixed α value, a selection based on the discriminant variable λ has the lowest β value.

=> The "likelihood ratio" is the most powerful quantity to discriminate between hypotheses.

WILKS THEOREM

In the following we evaluate χ^2_λ for the poissonian histogram.

$$\chi^2_\lambda = -2\ln\prod_{i=1}^{M}\frac{e^{-y_i}y_i^{n_i}}{n_i!} + 2\ln\prod_{i=1}^{M}\frac{e^{-\nu_i}\nu_i^{n_i}}{n_i!} \tag{110}$$

Notice that the first term includes the theory (through the y_i), while the second requires the knowledge of the expectation values of the data. If we make the identification $\nu_i = n_i$, we get:

$$\chi^2_\lambda = -2\sum_{i=1}^{M}\left(n_i \ln\frac{y_i}{n_i} - (y_i - n_i)\right) = -2\sum_{i=1}^{M}\left(n_i \ln\frac{y_i}{n_i}\right) + 2(N_0 - N) \tag{111}$$

By imposing $\nu_i = n_i$ eq.109 is the ratio of the likelihood of the theory to the likelihood of the data. The lower is χ^2_λ the better is the agreement between data and theory. For $y_i = n_i$ (perfect agreement) $\chi^2_\lambda = 0$.

If we make the same calculation for the multinomial likelihood we obtain the same expression but without the $N_0 - N$ term that corresponds to the fluctuation of the total number of events. This term is only present when we allow the total number of events to fluctuate, as in the poissonian case.