Other Proposed exercises

We perform a cross-section measurement and obtain the following values: N_{cand} =128, $N_b = 14 \pm 2$, $\epsilon = 0.523 \pm 0.002$, $L_{int} = 2.43$ pb$^{-1} \pm 1.8$ %: calculate the resulting cross-section with its uncertainty. In case this is a measurement of $e^+e^- \to \pi^+\pi^-$ at $\sqrt{s} = 1$ GeV, determine the value of the pion time-like form factor with its uncertainty. The formula relating the cross-section to the form factor $F_\pi(s)$ is the following:

$$\sigma(s) = \frac{\pi\alpha^2}{3s}\beta_\pi^3|F_\pi(s)|^2$$

Consider the reaction $e^+e^- \to K^+K^-$ at a Φ-factory. Which fraction of events have at least one kaon decaying within a sphere of R = 20 cm ? In which fraction of events both kaons decay within the same sphere ?

The SM expected semi-leptonic K_S charge asymmetry is 3×10^{-3}. At Dafne we expect to produce a sample of 1.2×10^9 tagged K_Ss. If the BR($K_S \to \pi e\nu$)=BR($K_S \to \pi^+e^-\overline{\nu}$)+BR($K_S \to \pi^-e^+\nu$)=$6.95\times10^{-4}$ which error can we reach on the asymmetry ?

Which average instantaneous luminosity is required to improve by a factor 3 such an uncertainty in one year of data taking (assuming a duty cicle of 50% and a tagging efficiency of 30%) ? [$\sigma(e^+e^- \to \phi)$= 3 μb at the ϕ peak].

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)

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)

$\Rightarrow$ We can use Likelihood ratios as test statistics with known pdf, more general than Pearson χ2, it holds in asymp. limit but whatever is the stat. model.

Connection with the Neyman-Pearson Lemma

$$P(type - I errors) = 1 - \epsilon = \alpha$$

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

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"

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

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

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

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

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.

WILKS THEOREM

5.2.2. *Study of a functional dependence.* A likelihood function can also be easily defined in another context widely used in experimental physics. We consider the case of M measurements z_i all characterized by gaussian fluctuations with uncertainties σ_i done for different values of an independent variable x. If the theory predicts a functional dependence between z and x given by the function $z = f(x/\underline{\theta})$ possibly depending on a set of parameters $\underline{\theta}$, in case of no correlation between the measurements z_i, and completely neglecting possible uncertainties on x, we can build a gaussian likelihood:

$$L_g(\underline{z}/\underline{\theta}) = \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{\left(z_i - f(x_i/\underline{\theta})\right)^2}{2\sigma_i^2}}$$

WILKS THEOREM

Let's now apply the Wilks theorem to this case. For the gaussian measurements we make the identification $\nu_i = E[z_i] = z_i$ and we get:
(113)

$$\chi^2_\lambda = -2\ln\prod_{i=1}^{M}\frac{1}{\sqrt{2\pi}\sigma_i}e^{-\frac{(z_i-f(x_i/\underline{\theta}))^2}{2\sigma_i^2}} + 2\ln\prod_{i=1}^{M}\frac{1}{\sqrt{2\pi}\sigma_i}e^{-\frac{(z_i-z_i)^2}{2\sigma_i^2}} = \sum_{i=1}^{M}\frac{(z_i - f(x_i/\underline{\theta}))^2}{\sigma_i^2}$$

The test statistics obtained here is a χ^2, typically used in the context of the so called **least squares method**. So we have proved essentially that the least square method can be derived through the Wilks theorem by a gaussian likelihood ratio model.

5.2.3. *Unbinned data.* In case we have a limited number N of events so that any binning will bring us to small values of bin contents, a different approach can be used, equally relying on the likelihood method: we can fit the **unbinned** data. In other words we build our likelihood function directly considering the probability of each single event. If we call H our hypothesis (eventually depending on a set of K parameters $\underline{\theta}$), x_i with i=1, ...N the values of the variable x for the N events and $f(x/\underline{\theta})$ the pdf of x given the hypothesis H, the likelihood can be written as:

$$L(\underline{x}/H) = \prod_{i=1}^{N} f(x_i/\underline{\theta}) \tag{114}$$

valid in case the events are not correlated. Notice that in this case the product runs on the events, not on the bins as in the previous case. If N is not fixed but fluctuates we can include "by hand" in the likelihood, the poissonian fluctuation of N around an expectation value that we call N_0 (eventually an additional parameter to be fit)[20]:

$$L(\underline{x}/H) = \frac{e^{-N_0} N_0^N}{N!} \prod_{i=1}^{N} f(x_i/\underline{\theta}) \tag{115}$$

This is called **extended likelihood**.

The - logarithm of the likelihood is used in most cases[21]:

$$-\ln L(\underline{x}/H) = -\sum_{i=1}^{N} \ln f(x_i/\underline{\theta}) \tag{116}$$

Choice of test statistics: correlations

5.2.4. *Fit of correlated data.* By using the product of the probability functions to write down the likelihood, we are assuming no correlation between bins (in case of histograms) or between events (in case of unbinned fits). In general it is possible to take into account properly the correlation between measurements in the definition of a likelihood function. We see how this happens in a simple case. Assume that our gaussian measurements of z_i (see above) are not independent. In this case the likelihood cannot be decomposed in the product of single likelihoods, but a "joint likelihood" $L(\underline{z}, /\underline{\theta})$ is defined, including the covariance matrix V_{ij} between the measurements. The covariance matrix has the parameters variances in the diagonal elements and the covariances in the off-diagonal elements. Starting from the joint likelihood of the measurement, we build the likelihood ratio and in the end we are left with the final χ^2:

$$\chi^2 = \sum_{j,k=1}^{M} (z_j - f(x_j/\underline{\theta}))\, V_{jk}^{-1}\, (z_k - f(x_k/\underline{\theta})) \tag{117}$$

that is still a χ^2 variable with $M - K$ degrees of freedom.