Acceptance Errors for Weighted Events

U. Husemann with the help of G. Zech

Definitions:

- Generator weight for event i: w_i
- Acceptance for event i: ε_i
- Number of weighted accepted events: $\sum_i w_i \varepsilon_i$
- Total acceptance: $A = \sum_i w_i \varepsilon_i / \sum_i w_i$

We are interested in how much the number of accepted events varies for a given numer of weighted generated events. Thus the error on the acceptance is only due to variations of the denominator.

For a single generated event i, the mean number of observations is ε_i with variance $\varepsilon_i(1-\varepsilon_i)$ (from binomial distribution). Since the events are independent, we get the error of A from the quadratic sum of the individual contributions:

$$(\delta A)^2 = \frac{\sum_i w_i^2 \varepsilon_i (1-\varepsilon_i)}{\left(\sum_i w_i\right)^2}.$$

To first approximation, the mean number of observations ε_i is constant, resulting in the formula presented by Ilija:

$$(\delta A)^2 = \frac{\varepsilon(1-\varepsilon) \sum_i w_i^2}{\left(\sum_i w_i\right)^2}.$$

In the di-muon analysis, the acceptance ε_i for every event is the product of two efficiencies: the geometrical acceptance $\varepsilon_{i,\mathrm{g}}$ and the trigger efficiency $\varepsilon_{i,\mathrm{t}}$. These quantities can be determined by arranging the events in certain classes: the geometrical acceptances $\varepsilon_{i,\mathrm{g}}$ are functions of x_F and p_T, and the FLT efficiencies $\varepsilon_{i,\mathrm{t}}$ are functions of the global event characteristics and assumed to be different for every event i. Assuming that $\varepsilon_{i,\mathrm{g}}$ and $\varepsilon_{i,\mathrm{t}}$ are uncorrelated, they can be multiplied. The error on acceptance of a given bin j in x_F and p_T are then

$$(\delta A_j)^2 = \frac{\sum_i w_i^2 \varepsilon_{j,\mathrm{g}} \varepsilon_{i,\mathrm{t}} (1-\varepsilon_{j,\mathrm{g}} \varepsilon_{i,\mathrm{t}})}{\left(\sum_i w_i\right)^2}.$$

Alternatively, we could divide the Monte Carlo sample in N (e.g. $N = 100$) subsamples and calculate the acceptances for each of them. The spread of the results, i.e. $\mathrm{RMS}/\sqrt{N}$, is an empirical measure for the acceptance error. The advantage of this method is that all correlations are taken into account automatically.