ON STATISTICAL ESTIMATES OF MEAN FREE PATH MEASUREMENTS

1) Distribution of \( \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \) for constant \( \lambda \)

Let \( x_i \) be the path length of a track before it interacts. If all events are associated with a unique value of \( \lambda \) their distribution will be exponential, i.e. \( e^{-\frac{x_i}{\lambda}} \).

Let us sample \( n \) events out of the universe population. The average is an estimate of \( \lambda \). In fact

\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \Rightarrow \int_0^\infty x e^{-\frac{x}{\lambda}} \frac{d\lambda}{\lambda} = \lambda
\]

Thus \( \bar{x} \) tends asymptotically to \( \lambda \). We want to find out how \( \bar{x} \) fluctuates (for \( n \neq \infty \)) around its asymptotic value.

The probability that the \( n \) events are associated with \( x_i \) - values included between \( x_1, x_1 + dx_1; x_2, x_2 + dx_2 \ldots \ldots x_n, x_n + dx_n \) is

\[
P(\frac{d\lambda_1, \ldots, d\lambda_n}{\lambda}, dx_1, dx_2 \ldots \ldots dx_n) = \hat{A} e^{-\frac{\sum x_i}{\lambda}} dx_1, dx_2 \ldots \ldots dx_n
\]

where \( \hat{A} \) is a normalization coefficient. We can write it also as

\[
P(\frac{d\lambda_1, \ldots, d\lambda_n}{\lambda}, dx_1, \ldots, dx_n) = \hat{A} e^{-\frac{n \bar{x}}{\lambda}} dx_1, \ldots, dx_n
\]

The element of volume can be expressed also in terms of the variation of \( \bar{x} \), i.e. \( d\bar{x} \). We notice that in the \( n \) - dimensional space \( x_1, x_2 \ldots \ldots x_n \), \( \bar{x} = \frac{1}{n} \sum x_i \) for a given \( \bar{x} \), is an hyperplane which cuts the axis at \( x_i = n \bar{x} \). Since all the \( x_i \) can only be positive by definition, a variation of \( \bar{x} \), say \( d\bar{x} \), makes the plane \( n\bar{x} = \sum x_i \) span a volume \( \frac{1}{2} n^2 \bar{x}^{n-1} d\bar{x} \). (See the two - dimensional case in the adjacent figure). Then

PS/4342
\[ \int_0^\infty Pd\bar{x} = A \frac{n^2 \overline{\bar{x}}}{2} \int_0^\infty e^{-\frac{n\bar{x}}{\lambda}} \bar{x}^{n-1} d\bar{x} \]

\[ = A \frac{n^2 \overline{\bar{x}}}{2} \left( \frac{\lambda}{n} \right)^n \Gamma(n) \]

Hence

\[ A \frac{n^2 \overline{\bar{x}}}{2} = \left( \frac{n}{\lambda} \right)^n \frac{1}{\Gamma(n)} \]

and

\[ Pd\bar{x} = \left( \frac{n}{\lambda} \right)^n \frac{1}{\Gamma(n)} e^{-\frac{n\bar{x}}{\lambda}} \bar{x}^{n-1} d\bar{x} \]

The average value of \( \bar{x} \) is the, for a large number of samples each of \( n \) events:

\[ \langle \bar{x} \rangle = \int_0^\infty \bar{x} P(\bar{x}) d\bar{x} = \frac{n}{(n-1)!} \left( \frac{\bar{x}}{\lambda} \right)^n e^{-\frac{n\bar{x}}{\lambda}} \bar{x} \frac{d\bar{x}}{\lambda} = \frac{\lambda n!}{n(n-1)!} = \lambda \]

c.d.d.

PS/4342
The fluctuation of $\bar{x}$ around its asymptotic value is

$$\sigma_n^2(\bar{x}) = \langle (\bar{x} - \lambda)^2 \rangle = \int_0^\infty (\bar{x} - \lambda)^2 \hat{P}(\bar{x}) d\bar{x} = \lambda^2 \left( \frac{n+1}{n} - 2 + 1 \right) = \frac{\lambda^2}{n}$$

Thus

$$\sigma_n(\bar{x}) = \frac{\lambda}{\sqrt{n}}$$

The limitation of this method lies in the fact that it does not allow one to discuss the case in which our sample contains - in unknown proportions - more than one type of particle, each type being associated with a different mean free path $\lambda_i$. It can only determine the consistency of a hypothesis with observations.

2. The maximum likelihood method

The difficulty mentioned above can be overcome using the maximum likelihood method. Let us consider the case of two species only ($\lambda_1$ and $\lambda_2$). Let $x_i$ be the track length of the $i$-th event (between its origin and its interaction) and $\xi_i$ its "potential path". The probability of an event to fall between $x_i$ and $x_i + dx_i$ for a given $\lambda$ is then

$$p_i dx_i = \frac{\lambda^{-1}}{1 - e^{-\xi_i/\lambda}} \frac{dx_i}{\lambda}$$

If the two species are present and $q$ is the (unknown) relative proportion of one of the two (say $\lambda_2$), then we shall consider the combined frequency function

$$\bar{p}(\lambda, \lambda_2 | \xi, x_i) = \frac{1-q}{\lambda_2} \frac{e^{-x_i/\lambda}}{1 - e^{-\xi_i/\lambda}} + \frac{q}{\lambda_2} \frac{e^{-x_i/\lambda_2}}{1 - e^{-\xi_i/\lambda_2}}$$

In the case of interest to us, one of the supposed present species is formed by $\mu$-mesons while the predominant one is of $\pi$'s. We can then assume safety $\lambda_2 = \lambda = \infty$. Then

$$\ell_{\bar{m}} = \frac{1}{\lambda_2} \frac{e^{-x_i/\lambda_2}}{1 - e^{-\xi_i/\lambda_2}} = \frac{1}{\xi_i}$$

PS/4342
Thus, the likelihood function $\mathcal{L}$ reads

$$\mathcal{L} = \prod_{i} \left[ (1-q) \frac{e^{-x_i/\lambda}}{\lambda (1-e^{-x_i/\lambda})} + \frac{q}{\xi_i} \right]$$

or

$$\mathcal{L} = \prod_{i} \left[ \frac{e^{-x_i/\lambda}}{\lambda (1-e^{-x_i/\lambda})} + q \left( \frac{1}{\xi_i} - \frac{e^{-x_i/\lambda}}{\lambda (1-e^{-x_i/\lambda})} \right) \right]$$

where $\lambda$ stands for $\lambda_1 = \lambda_2$

We shall assume that $\lambda$ be known (and, for the moment, constant. Different values of $\lambda$ will be considered later). The unknown parameter to be estimated is then $q$. The likelihood equation then reads

$$0 = \frac{2}{\lambda q} \sum_{i} \mathcal{L} = \sum_{i} \frac{1}{\xi_i} - \frac{e^{-x_i/\lambda}}{\lambda (1-e^{-x_i/\lambda})} \frac{e^{-x_i/\lambda}}{\lambda (1-e^{-x_i/\lambda})} + q \left( \frac{1}{\xi_i} - \frac{e^{-x_i/\lambda}}{\lambda (1-e^{-x_i/\lambda})} \right)$$

Putting $a_i = \frac{1}{\xi_i} - \frac{e^{-x_i/\lambda}}{\lambda (1-e^{-x_i/\lambda})}$, $b_i = \frac{e^{-x_i/\lambda}}{\lambda (1-e^{-x_i/\lambda})}$

$$\sum_{i} \frac{a_i}{b_i + q a_i} = 0$$

The solution of this equation gives the likelihood estimate $q^*$ of the parameter $q$.

The variance of $q^*$ is given by

$$\sqrt{\mathbb{E}\left(-\frac{2}{\lambda q^2} \frac{\partial \mathcal{L}}{\partial q^2}\right)} = \sqrt{\mathbb{E}\left[ \frac{\sum a_i^2}{(a_i q + b_i)^2} \right]}$$

Alternatively one can calculate the Bartlett function

$$S = \frac{\sum a_i/(a_i q + b_i)}{\left\{ \sum a_i^2/(a_i q + b_i)^2 \right\}^{1/2}}$$
which has variance 1. Since $S$ is asymptotically normal, the limits of confidence for $q$, associated with a probability $\Theta$ are given by the roots of the equation

$$S = \pm \Theta (P)$$

where $\Theta (P)$ is defined by the equation

$$\sqrt{\frac{2}{\pi}} \int_{\Theta (P)}^{\infty} e^{-\frac{\eta^2}{2}} d\eta = P (|q - q| > \Theta)$$

***

So far we have assumed a unique value for $\lambda$. In practical cases $\lambda$ is not constant, but varies with the energy of the particle, i.e. along the range. Then the $\mathcal{L}$ function reads

$$\mathcal{L} = \prod_i \sum_j p_j \frac{e^{-\frac{x_j}{\lambda_j}}}{\lambda_j (1 - e^{-\frac{x_j}{\lambda_j}})}$$

which for a continuous distribution of $\lambda$, say $f(\lambda)$ becomes

$$\mathcal{L} = \prod \frac{1}{\lambda_i} \int f(\lambda) \frac{e^{-\frac{x_i}{\lambda}}}{\lambda (1 - e^{-\frac{x_i}{\lambda}})} d\lambda$$

In general $f(\lambda)$ is not expressible in a simple analytical form. Thus we shall use the first expression.

In the case of a $\pi/\mu$ mixture, let the proportion of the observed events falling in the $j$-th interval of energy, corresponding to $\lambda_j$, be $f_j$. Furthermore, let $q$ be the proportion of $\mu$. We normalize the $f_j$ by the equation $\sum f_j = 1$ (That means that $f_j = \frac{\eta_j}{\sum \eta_j}$). Then

$$\mathcal{L} = \prod \left\{ (1 - q) \sum_j f_j \frac{e^{-\frac{x_i}{\lambda_j}}}{\lambda_j (1 - e^{-\frac{x_i}{\lambda_j}})} \right\}$$

FS/4342
and
\[ \frac{\partial \ln L}{\partial q} = 0 = \sum_i \frac{A_i}{A_i q + B_i} \]

where
\[ A_i = \frac{1}{\xi_i} + \sum_j f_j \frac{e^{-x_i/\lambda_j}}{\lambda_j (1 - e^{-x_i/\lambda_j})} \]
\[ B_i = \sum_j f_j \frac{e^{-x_i/\lambda_j}}{\lambda_j (1 - e^{-x_i/\lambda_j})} \]

The Bartlett function reads again
\[ S = \left( \frac{\sum_i \frac{A_i^2}{A_i q + B_i}}{\left( \sum_i \frac{A_i^2}{(A_i q + B_i)^2} \right)^{1/2}} \right)^{1/2} \]

C. Franzinetti

/fv

**Distribution:** (open)

The N.P.A. Scientific Staff

PS/4342