

Noções de Estatística, Simulações e Erros

Alberto Santoro [◊]

*CBPF, Centro Brasileiro de Pesquisas Físicas
Rua Dr. Xavier Sigaud 150, 22290-180 Rio de Janeiro, RJ, Brasil*

Sérgio F. Novaes ^{*}

*Instituto de Física Teórica – UNESP
Rua Pamplona, 145, 01405-900 São Paulo, SP, Brasil*

Vitor Oguri [†]

*UERJ, Universidade do Estado do Rio de Janeiro
Rua São Francisco Xavier, 524, 20559-900 Rio de Janeiro, RJ, Brasil*

[◊] E-mail: santoro@lafex.cbpf.br;

^{*} E-mail: novaes@ift.unesp.br

[†] E-mail: oguri@uerj.br

Índice

1	Introdução	1
1.1	Introdução	1
1.2	Probabilidades e Distribuições	2
1.2.1	Erros e Probabilidades	2
1.2.2	Distribuições de Probabilidades e Parâmetros	4
1.3	Propagação de Erros	11
1.3.1	Estimação de Parâmetros e Erros em Medidas Diretas	14
1.3.2	Correlação no Cálculo de Eficiências em Experimentos de Altas Energias	18
2	Distribuições Básicas	23
2.1	Distribuição Binomial	23
2.2	Distribuição de Poisson	26
2.3	Distribuição Gaussiana ou Normal	27
2.4	Outras Distribuições	28
3	Ajustes de Curvas a um Conjunto de Dados	31
3.1	Método dos Momentos	33
3.2	Método da Máxima Verossimilhança	34
3.2.1	Separação de píons e káons	35
3.3	Método dos Mínimos Quadrados	36
3.4	Teste de χ^2	38

4	Monte Carlo	39
4.1	Porque Utilizar Métodos de Monte Carlo?	41
4.2	Aplicações Simples de Monte Carlo	42
4.2.1	Cálculo de uma Integral	42
4.2.2	Momentum Transversal Médio	43
4.2.3	Espaço de Fase	43
4.2.4	Redução do Erro Estatístico	48
4.2.5	Geração de Eventos	51
5	Limite Superior e Incerteza	57
5.1	Introdução	57
5.2	O caso $n = 0$	58
5.3	Método Integral	59
5.4	Método da Expansão	60
5.5	Subtração de Background	62
6	O Potencial de Descoberta de Física Nova	65
6.1	Introdução	65
6.2	Uma Análise da Flutuação Estatística	68
6.3	Levando em Consideração os Erros Sistemáticos	73
6.4	Limites de Exclusão de Física Nova	75
6.5	A Probabilidade de Descoberta de Física Nova	76
6.6	Conclusão	77

A	Tópicos sobre a Teoria de Probabilidades	79
A.1	Eventos equivalentes	79
A.1.1	pdf de eventos equivalentes	79
A.1.2	Valor médio de função de uma variável aleatória	81
A.2	Variáveis aleatórias multidimensionais	81
A.2.1	Variáveis independentes	82
A.2.2	Funções de várias variáveis aleatórias	82
A.3	Desigualdades e Teoremas	84
A.3.1	Desigualdade de Markov	84
A.3.2	Desigualdade de Chebyshev	85
A.3.3	Lei dos grandes números	85
A.3.4	Função característica	86
A.3.5	Propriedades da função característica	87
A.3.6	Teorema do limite central	88
A.4	Distribuições	89
A.4.1	Distribuição gama	89
A.4.2	Distribuição de χ^2	90
A.4.3	Distribuição de Student	91
B	Ferramentas para Simulação e Análise de Dados	95
B.1	PAW	96
B.1.1	Vetores	96
B.1.2	Histogramas	98
B.1.3	N-úplas	99
B.2	ROOT	100
B.3	Exemplos	100

C	Cinemática Relativística	107
C.1	Relatividade Restrita	107
C.1.1	Transformação de Lorentz	107
C.1.2	Sistemas de Referência	112
C.2	Decaimento	116
C.2.1	Decaimento de Dois Corpos	116
C.2.2	Decaimento de Três Corpos	116
C.3	Espalhamento $2 \rightarrow 2$	119
C.3.1	Região Física em s , t e u	121
C.4	Seção de Choque e Largura de Decaimento	122
C.5	Espaço de Fase	124
C.5.1	Espaço de Fase de Uma Partícula	126
C.5.2	Espaço de Fase de Duas Partículas	126
C.5.3	Espaço de Fase de Três Partículas	129

Prefácio

Nessas notas, reunimos alguns tópicos e técnicas estatísticas que utilizamos em nosso trabalho em Física Experimental de Altas Energias, quer seja em análise de dados ou na simulação de experimentos, como os Métodos de Monte Carlo, que utilizamos desde a simulação dos processos de colisões à simulação dos próprios detectores de partículas.

Após a definição de alguns conceitos estatísticos, necessários à descrição de qualquer conjunto de dados, apresentamos as principais Distribuições de Probabilidades, sobre as quais se baseiam as análises de erros e a extração de resultados a partir dos conjuntos de medidas.

Apesar de estar dirigido, principalmente, aos estudantes de Física de Altas Energias, os métodos apresentados são gerais e, portanto, aplicáveis em outros contextos.

Além de serem aquelas que mais nos influenciaram na construção dessas notas, as referências bibliográficas são para que o leitor interessado consulte-as e aprofunde os tópicos apresentados.

Finalmente, gostaríamos de salientar que essas notas de aula não constituem um trabalho original nem um trabalho completo sobre os tópicos tratados. Constituem apenas um guia de aula para o estudante e o professor. Assim, esperamos que com a utilização dessas notas em cursos de Física Experimental tenhamos sugestões que nos permitam ampliar e dar uma redação mais homogênea ao texto.

Capítulo 1

Introdução

1. Introdução
2. Probabilidades e Distribuições
3. Propagação dos Erros

1.1 Introdução

Inicialmente, abordaremos algumas das noções com as quais trabalhamos no cotidiano de uma análise de dados em Física Experimental de Altas Energias. Para isso, resumimos alguns tópicos baseados em referências clássicas sobre o tema, que são apresentadas na bibliografia e, que incluem livros como o de Bevington [1], até textos mais recentes como os de Lyons [2] e Cowan [3].

Todo experimento em Física de Altas Energias, em suas tomadas de dados (*run*) iniciais, produz resultados que diferem dos inicialmente esperados. Essa é uma constatação da observação de vários experimentos. Daí a necessidade de um período de testes (*test-beam*), em que são realizadas algumas tomadas para calibração dos instrumentos, que antecipa a obtenção de dados. Somente então, os dados são trabalhados com métodos estatísticos apropriados que permitem a extração e apresentação correta dos resultados e suas respectivas incertezas ou erros.

Entendemos o **erro** ou a **incerteza** associada a um resultado¹ como sendo a diferença entre um valor calculado ou observado e o valor esperado ou de referência da grandeza.

¹Como a determinação do valor de uma grandeza.

A seguir, vamos definir alguns dos termos estatísticos mais comumente usados e calculados em experimentos de Física de Altas Energias.

1.2 Probabilidades e Distribuições

1.2.1 Erros e Probabilidades

1. Erros Sistemáticos

Resultam da falta de calibração, da resolução limitada dos instrumentos de medição ou de algum efeito ou critério tendencioso (*bias*) na observação do sistema. Ou seja, o erro sistemático é a própria reprodução da imprecisão instrumental ou dos métodos de observação, de calibração ou de outras técnicas usadas em um experimento. Uma outra fonte de erros sistemáticos decorre da própria modelagem do fenômeno ou processo que se descreve.

2. Erros Aleatórios ou Estatísticos

Resultam de variações aleatórias, devido a fatores incontroláveis do fenômeno em si ou do processo de medição. Desse modo, acarretam flutuações dos resultados mesmo após a repetição do experimento.

3. Probabilidade

Pode ser definida, por exemplo, em termos da teoria de conjuntos, como foi por Kolmogorov em 1933. Dado um conjunto S , denominado espaço amostral, a probabilidade associada a cada sub-conjunto A de S é um número real $P(A)$ definido a partir de três axiomas:

(i) Para todo A em S ,

$$P(A) \geq 0$$

(ii) Para quaisquer dois sub-conjuntos A e B disjuntos (*i.e.* $A \cap B = \emptyset$),

$$P(A \cup B) = P(A) + P(B)$$

(iii) A probabilidade associada ao espaço amostral:

$$P(S) = 1$$

Uma variável que assume um valor específico para cada elemento de S é chamada de **Variável Aleatória**.

4. Probabilidade Condicional

Supondo que $S \supset A$ e B e $P(B) \neq 0$, define-se $P(A | B)$, *i.e.* a probabilidade de A dado B , como

$$P(A | B) \equiv \frac{P(A \cap B)}{P(B)} \quad (1.1)$$

Exercício 1.1 Usando a definição de $P(B | A)$ e $A \cap B = B \cap A$ derive o teorema de Bayes:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} \quad (1.2)$$

ou se $S = \cup_i A_i$ com $A_i \cap A_j = \emptyset$ para $i \neq j$ e $P(A_i) \neq 0$,

$$P(A | B) = \frac{P(B | A)P(A)}{\sum_i P(B | A_i)P(A_i)} \quad (1.3)$$

5. Estatística Clássica

O significado do conceito de probabilidade num experimento decorre principalmente de duas interpretações ligadas a escolas filosóficas distintas.

Para a Estatística Clássica, os elementos de S correspondem aos possíveis resultados de uma medida. A probabilidade de um elemento de S fornecer A é definida como a fração de vezes que A ocorre no limite em que a medida é repetida um número infinito de vezes, *i.e.*

$$P(A) = \lim_{N \rightarrow \infty} \frac{\text{número de ocorrências de } A \text{ em } N \text{ medidas}}{N}$$

O objetivo principal é estimar essas probabilidades dado um número finito de dados experimentais e testar até que ponto uma determinada teoria que prediz essas probabilidades é compatível com os dados observados.

6. Estatística Bayesiana

Para a Estatística Bayesiana, os elementos de S correspondem a hipóteses que podem ser falsas ou verdadeiras. A probabilidade associada à uma hipótese A é interpretada como uma medida do grau de convicção nessa hipótese,

$$P(A) = \text{grau de confiança de que a hipótese } A \text{ seja verdadeira}$$

e o teorema de Bayes (1.2) pode ser escrito como

$$P(\text{teoria} \mid \text{dado}) \propto P(\text{dado} \mid \text{teoria}) \times P(\text{teoria})$$

onde,

$P(\text{teoria})$ é a probabilidade *a priori* de que a teoria seja verdadeira,

$P(\text{dado} \mid \text{teoria})$, chamado de Verossimilhança (*Likelihood*), é a probabilidade de que, utilizando as hipóteses da teoria, observe-se os dados que foram realmente obtidos e,

$P(\text{teoria} \mid \text{dado})$ é a probabilidade *a posteriori* de que a teoria seja correta em vista dos dados do experimento realizado.

Uma vez que $P(\text{teoria})$ seja dado, a Estatística Bayesiana nos mostra como se deve mudar nosso grau de confiança no modelo, em vista dos dados experimentais.

1.2.2 Distribuições de Probabilidades e Parâmetros

1. Função Densidade de Probabilidade (pdf) ou Distribuição de Probabilidades

É a função que fornece a probabilidade de se observar um valor x de uma determinada variável contínua dentro de um intervalo infinitesimal $[x, x + dx]$, ou seja:

$$\text{Probabilidade de observar } x \text{ em } [x, x + dx] = f(x)dx$$

Toda pdf deve ser normalizada, *i. e.* $\int f(x)dx = 1$, de tal forma que a probabilidade total seja 1.

A figura 1.1 mostra um exemplo de uma distribuição genérica.

Para uma variável discreta x_i , as probabilidades são,

$$\text{Probabilidade de se observar } x_i = P(x_i) = f_i$$

$$\text{e } \sum_i f_i = 1.$$

Segundo a interpretação de frequência relativa (Estatística Clássica) $f(x)dx$ fornece a fração de vezes que x é observado no intervalo $[x, x + dx]$ no limite de um número infinito de medidas.

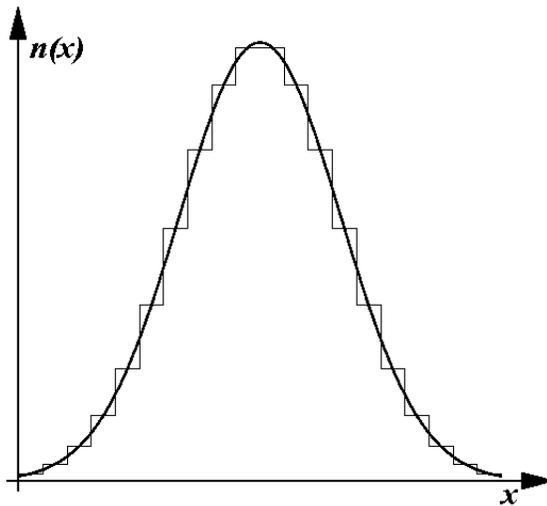


Figura 1.1: Exemplo de uma distribuição genérica $f(x)$. O histograma representado pelos segmentos de retas pode ser substituído pela curva que é o resultado de um ajuste com os seus pontos médios em cada “bin”.

2. Distribuição Original

É a distribuição de um conjunto infinito hipotético de valores da qual os dados experimentais são supostos ser uma amostra aleatória. Ou seja, é a **pdf** que melhor descreve as sucessivas medidas realizadas em um experimento. O conjunto de valores da distribuição original chama-se **População Original**.

3. Média

Dado um conjunto de N medidas x_i , define-se a **média** (\bar{x}) das medidas por

$$\bar{x} = \frac{1}{N} \sum_i^N x_i \quad (1.4)$$

4. Parâmetros

São quantidades numéricas (média, desvio padrão,...) associadas a uma dada distribuição, cujas mudanças não alteram a natureza da distribuição.

Uma das hipóteses básicas da estatística ² é que, quando o número N de medi-

²O Teorema do Limite Central (apêndice A).

das for muito grande, os valores dos parâmetros experimentais se aproximam dos valores originais. Daí poderemos afirmar que os parâmetros originais são iguais aos parâmetros experimentais no limite do número de medidas infinitas.

$$\boxed{\text{parâmetro experimental} \xrightarrow{N \rightarrow \infty} \text{parâmetro original}}$$

Assim, a **média** μ da população original é estimada por

$$\mu = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i^N x_i \quad (1.5)$$

(Observação : Sempre que estivermos nos referindo a um conjunto de dados e seus parâmetros característicos, estaremos nos referindo a população original, salvo referência explícita do contrário).

5. Distribuição Cumulativa

Representa a probabilidade de uma variável aleatória assumir um valor menor ou igual a um determinado x , *i.e.*

$$F(x) = \int_{-\infty}^x f(x') dx'$$

6. Mediana

A **mediana** de x , denotada por $x_{1/2}$, representa a fronteira de 50% das medidas menores que $x_{1/2}$ e 50% das medidas maiores que $x_{1/2}$, ou seja, existe igual probabilidade de x ser observado menor ou maior do que $x_{1/2}$ (figura 1.2),

$$F(x_{1/2}) = \int_{-\infty}^{x_{1/2}} f(x) dx = \frac{1}{2} \implies x_{1/2} = F^{-1}(1/2)$$

7. Moda

A **moda** x_M é o valor da variável aleatória x para o qual a função densidade de probabilidade é máxima. Este é o valor mais provável da população original, *i.e.* o valor da medida mais comumente observado (figura 1.2).

Para distribuições com uma única moda,

$$\mu \leq x_{1/2} \leq x_M \quad \text{ou} \quad x_M \leq x_{1/2} \leq \mu$$

Para uma distribuição simétrica vale a igualdade.

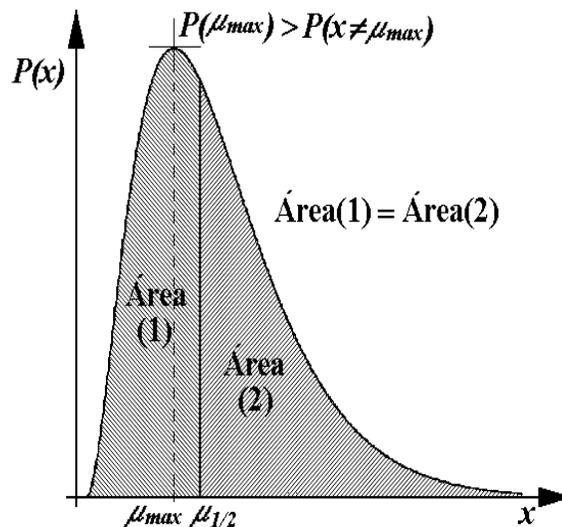


Figura 1.2: Mediana e moda numa distribuição genérica.

Exercício 1.2 Suponha que uma variável aleatória x possua a pdf $f(x)$. Dada uma função contínua desta variável, $a(x)$, mostre que a pdf $g(a)$ que descreve a distribuição de a é,

$$g(a) = f[x(a)] \left| \frac{dx}{da} \right|$$

ou, no caso de várias variáveis,

$$g(a_1, \dots, a_n) = f(x_1, \dots, x_n) \left| J \left(\begin{matrix} x_1, \dots, x_n \\ a_1, \dots, a_n \end{matrix} \right) \right|$$

onde J é o Jacobiano da transformação.

8. Valor Esperado ou Médio

O **valor esperado** $\langle x \rangle$ ou μ de uma variável aleatória x distribuída de acordo com a função densidade de probabilidade $f(x)$ é

$$\mu \equiv \langle x \rangle = \int_{-\infty}^{+\infty} x f(x) dx . \quad (1.6)$$

Para uma função $a(x)$ (vide exercício anterior)

$$\langle a(x) \rangle = \int_{-\infty}^{+\infty} a g(a) da = \int_{-\infty}^{+\infty} a(x) f(x) dx .$$

ou, para várias variáveis,

$$\langle a(x_1, \dots, x_n) \rangle = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} a(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \dots dx_n \equiv \mu_a \quad (1.7)$$

9. Variância

A **variância** σ^2 associada a um conjunto de medidas x_i é definida como o limite da média dos quadrados dos desvios dos x_i 's da média μ ,

$$\sigma^2 = \lim_{N \rightarrow \infty} \left[\frac{1}{N} \sum (x_i - \mu)^2 \right] = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum x_i^2 \right) - \mu^2 \quad (1.8)$$

onde σ é chamado de **desvio padrão**. Assim,

$$\sigma^2 = \langle (x - \mu)^2 \rangle = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \langle x^2 \rangle - \langle x \rangle^2 = \langle x^2 \rangle - \mu^2 \quad (1.9)$$

Para uma função de várias variáveis (cf. Eq. 1.7)

$$\sigma_a^2 = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} [a(x_1, \dots, x_n) - \mu_a]^2 f(x_1, \dots, x_n) dx_1 \dots dx_n$$

10. Covariância

Podemos definir a **covariância** de duas variáveis aleatórias como,

$$V_{xy} = \langle (x - \mu_x)(y - \mu_y) \rangle = \langle xy \rangle - \mu_x \mu_y \quad (1.10)$$

A matriz cujos elementos são $V_{xy} = \text{cov}[x, y]$ é chamada de matriz de covariância ou matriz de erro. Note que $V_{xx} = \sigma_x^2$. Se as duas variáveis forem independentes, $f(x, y) = X(x)Y(y)$, temos $\langle xy \rangle = \langle x \rangle \langle y \rangle$ e, portanto, $V_{xy} = 0$

11. Coeficiente de Correlação

O **coeficiente de correlação** ρ_{xy} fornece uma medida do nível de correlação entre duas variáveis aleatórias:

$$\rho_{xy} = \frac{V_{xy}}{\sigma_x \sigma_y}$$

com $-1 \leq \rho_{xy} \leq 1$.

Suponhamos, por exemplo, que quando $x > \mu_x$ a probabilidade de encontrar $y > \mu_y$ aumente. Neste caso $V_{xy} > 0$ e as variáveis possuem correlação positiva. Caso contrário, *i.e.* quando $x > \mu_x$ a probabilidade de encontrar $y < \mu_y$ aumenta, $V_{xy} < 0$ e as variáveis possuem correlação negativa.

Exercício 1.3 Meça a largura de uma mesa, 10, 50 e 100 vezes separadamente, com uma simples régua.

- (1) Construa uma tabela com seis colunas onde voce coloca em cada coluna os títulos:
 1. Comprimento L em centímetros (x);
 2. Frequência do Valor Medido (f);
 3. Produto dos itens (a) e (b) ($x \times f$);
 4. Valor do desvio ($x - \bar{x}$)
 5. Quadrado do desvio anterior ($(x - \bar{x})^2$);
 6. Produto de $f \times (x - \bar{x})^2$.
- (2) Calcule a média ponderada

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N f_i x_i$$

- (3) Calcule a variância da amostra

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N f_i (x_i - \bar{x})^2$$

Como um exemplo, para ilustrar os parâmetros de uma distribuição, vamos tomar a chamada Distribuição Gaussiana . Sua forma geral é dada pela expressão (figura 1.3):

$$G(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right] \quad (1.11)$$

Algumas das propriedades da Distribuição Gaussiana são:

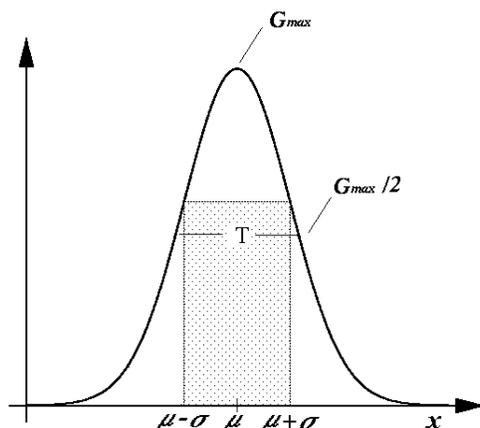


Figura 1.3: Um exemplo de Distribuição Gaussiana.

1. O fator de normalização $1/(\sigma\sqrt{2\pi})$ garante que

$$\int_{-\infty}^{+\infty} G(x; \mu, \sigma) dx = 1 .$$

2. O valor médio ou média é μ .
3. A função é simétrica em torno da média μ . Valor para o qual G assume seu valor máximo.
4. O valor máximo é

$$G_{max} = G(x = \mu; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}$$

5. O desvio padrão da distribuição, que caracteriza a largura da gaussiana, é σ .
6. A altura da curva para $x = \mu \pm \sigma$ é $(1/\sqrt{e})$ do valor máximo:

$$G(x = \mu \pm \sigma; \mu, \sigma) = \frac{1}{\sqrt{e}} G_{max}$$

7. Um outro parâmetro característico é a largura a meia altura Γ (FWHM – *full width at half maximum*), definida pelo intervalo limitado por valores de x para os quais o valor de G cai à metade de seu valor máximo, ou seja,

$$G(\mu \pm \Gamma/2) = G_{max}/2$$

8. As áreas compreendidas entre

$$\mu - N\sigma \leq x \leq \mu + N\sigma$$

são 0.6827 ($N = 1$), 0.9545 ($N = 2$) e 0.9973 ($N = 3$). Isto quer dizer que o valor medido de x está 68.27 % das vezes contido no intervalo $\mu \pm \sigma$. Essas áreas são mostradas na figura 1.4

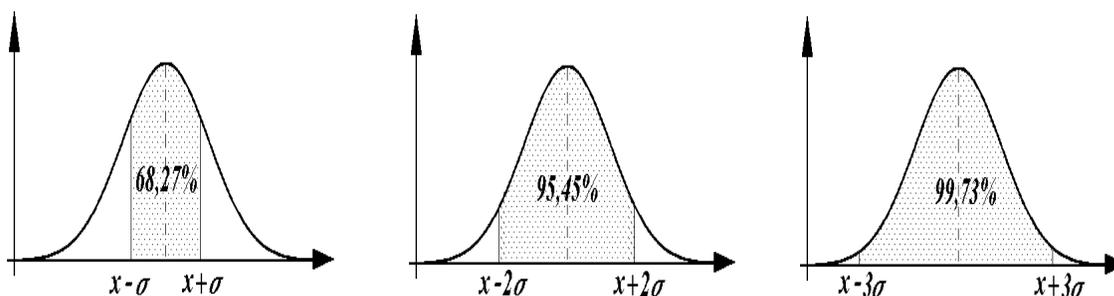


Figura 1.4: Intervalos e áreas características de uma Distribuição Gaussiana

1.3 Propagação de Erros

Se a determinação do valor de uma grandeza x se dá por um processo indireto, por exemplo, a partir das medidas de n outras grandezas (y_1, \dots, y_n) , como combinar as incertezas de variáveis individuais para estimar a incerteza no resultado?

Assim, se o valor de x que é dado pela função

$$x = \phi(y_1, \dots, y_n). \quad (1.12)$$

mesmo quando apenas a média (μ_i) e as covariâncias (V_{ij}) de y_i são conhecidas (e não a distribuição específica dessas variáveis) podemos estimar o valor esperado de x e sua variância.

Façamos agora uma expansão em série de Taylor na expressão (1.12) em torno da média de y_i , até a primeira ordem:

$$\begin{aligned} x &\simeq \phi(\mu_1, \dots, \mu_n) + \left(\frac{\partial \phi}{\partial y_1} \right)_\mu (y_1 - \mu_1) + \dots + \left(\frac{\partial \phi}{\partial y_n} \right)_\mu (y_n - \mu_n) \\ &= \phi(\mu_1, \dots, \mu_n) + \sum_{i=1}^n \left(\frac{\partial \phi}{\partial y_i} \right)_\mu (y_i - \mu_i) \end{aligned} \quad (1.13)$$

onde $(\partial\phi/\partial y_i)_\mu$ é a derivada calculada no ponto sobre o qual foi feita a expansão *i.e.*, $(\mu_1, \mu_2, \dots, \mu_n)$.

Podemos ver que a média de x em primeira ordem é

$$\langle x \rangle \simeq \phi(\mu_1, \dots, \mu_n) \quad (1.14)$$

já que $\langle (y_i - \mu_i) \rangle = 0$.

O valor esperado de $\langle x^2 \rangle$ será,

$$\begin{aligned} \langle x^2 \rangle &\simeq \phi^2(\mu_1, \dots, \mu_n) + 2\phi(\mu_1, \dots, \mu_n) \sum_{i=1}^n \left(\frac{\partial\phi}{\partial y_i} \right)_\mu \langle y_i - \mu_i \rangle \\ &+ \left\langle \left[\sum_{i=1}^n \left(\frac{\partial\phi}{\partial y_i} \right)_\mu (y_i - \mu_i) \right] \left[\sum_{j=1}^n \left(\frac{\partial\phi}{\partial y_j} \right)_\mu (y_j - \mu_j) \right] \right\rangle \\ &= \phi^2(\mu_1, \dots, \mu_n) + \sum_{i,j=1}^n \left(\frac{\partial\phi}{\partial y_i} \frac{\partial\phi}{\partial y_j} \right)_\mu V_{ij} \end{aligned} \quad (1.15)$$

onde $V_{ij} = \langle (y_i - \mu_i)(y_j - \mu_j) \rangle$ é a matriz de covariância das variáveis y .

De (1.14) e (1.15) obtemos a variância de x ,

$$\sigma_x^2 \equiv \langle x^2 \rangle - \langle x \rangle^2 \simeq \sum_{i,j=1}^n \left(\frac{\partial\phi}{\partial y_i} \frac{\partial\phi}{\partial y_j} \right)_\mu V_{ij} \quad (1.16)$$

Para um conjunto de m funções $\phi_1(y_1, \dots, y_n), \dots, \phi_m(y_1, \dots, y_n)$ a matrix de covariância será,

$$U_{kl} \equiv \text{cov}[\phi_k, \phi_l] \simeq \sum_{i,j=1}^n \left(\frac{\partial\phi_k}{\partial y_i} \frac{\partial\phi_l}{\partial y_j} \right)_\mu V_{ij} = \sum_{i,j=1}^n \left(\frac{\partial\phi_k}{\partial y_i} \right)_\mu V_{ij} \left(\frac{\partial\phi_l}{\partial y_j} \right)_\mu = \sum_{i,j=1}^n A_{ki} V_{ij} (A^T)_{jl} \quad (1.17)$$

onde definimos a matrix $A_{ij} = (\partial\phi_i/\partial y_j)_\mu$

Essas são as expressões mais gerais da propagação dos erros. Supondo que as variáveis y_i não sejam correlacionadas, *i.e.* $V_{ii} = \sigma_i^2$ e $V_{ij} = 0$ para $i \neq j$, temos

$$\sigma_x^2 \simeq \sum_{i=1}^n \left(\frac{\partial\phi}{\partial y_i} \right)_\mu^2 \sigma_i^2 \quad (1.18)$$

e

$$\text{cov}[\phi_k, \phi_l] \simeq \sum_{i=1}^n \left(\frac{\partial \phi_k}{\partial y_i} \frac{\partial \phi_l}{\partial y_j} \right)_{\mu} \sigma_i^2 \quad (1.19)$$

Podemos obter relações específicas para algumas funções de interesse, aplicando a relação (1.18):

1. Adição : $x = y_1 + y_2 + \dots + y_n$

neste caso $(\partial \phi / \partial y_i)^2 = 1$ e

$$\sigma_x^2 = \sum_{i=1}^n \sigma_i^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2$$

Assim para um conjunto de medidas que se somam ou estão dentro de uma soma, o resultado final se expressa como $x \pm \sigma_x$ onde x e σ_x são obtidos da forma acima.

2. Subtração : $x = y_1 - y_2$

$$\sigma_x^2 = \sigma_1^2 + \sigma_2^2$$

3. Multiplicação : $x = y_1 \cdot y_2 \cdot \dots \cdot y_n$

como

$$\left(\frac{\partial \phi}{\partial y_i} \right)_{\mu} = (y_1 \cdot \dots \cdot y_{i-1} \cdot y_{i+1} \cdot \dots \cdot y_n)_{\mu} = \left(\frac{\phi}{y_i} \right)_{\mu} = \left(\frac{\langle x \rangle}{\mu_i} \right)$$

temos

$$\sigma_x^2 = \langle x \rangle^2 \left[\left(\frac{\sigma_1}{\mu_1} \right)^2 + \left(\frac{\sigma_2}{\mu_2} \right)^2 + \dots + \left(\frac{\sigma_n}{\mu_n} \right)^2 \right]$$

4. Divisão : $x = \frac{y_1}{y_2}$

como

$$\left(\frac{\partial \phi}{\partial y_1} \right)_{\mu} = \frac{1}{\mu_2} = \frac{\langle x \rangle}{\mu_1} \quad \text{e} \quad \left(\frac{\partial \phi}{\partial y_2} \right)_{\mu} = \frac{-\mu_1}{\mu_2^2} = \frac{-\langle x \rangle}{\mu_2}$$

temos

$$\sigma_x^2 = \langle x \rangle^2 \left[\left(\frac{\sigma_1}{\mu_1} \right)^2 + \left(\frac{\sigma_2}{\mu_2} \right)^2 \right]$$

1.3.1 Estimação de Parâmetros e Erros em Medidas Diretas

No sentido lato, os objetivos de qualquer *experimento em física* estão associados a/ao:

- Determinação de um “*valor de referência*” para uma grandeza associada a um fenômeno ou sistema físico, segundo algum modelo pré-concebido ou Teoria Física existente ou proposta;
- Estabelecimento de uma Lei Física particular, de relações entre grandezas ou de “parâmetros” de uma teoria, a partir de um conjunto de medidas.

Na linguagem estatística, aos objetivos dos experimentos do primeiro grupo chamam-se problemas de *estimação de parâmetros* e, aos do segundo, de *testes de hipóteses*. A partir desses procedimentos e da correlação de seus resultados com algum modelo físico é que, ordinariamente, procura-se estabelecer os fundamentos e as teorias que constituem a *Física*.

É claro que esses problemas não são independentes. Pois, ao se determinar um parâmetro, a partir de um conjunto de N medidas diretas ³ $\{x_i\}$ de uma grandeza x , pressupõe-se que os desvios ou erros das medidas $x_i - \mu$, em relação a um valor esperado de referência (μ), estarão distribuídos segundo uma Distribuição Gaussiana.

$$\exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma_x} \right)^2 \right]$$

O que constitui uma hipótese a qual se pode atribuir, posteriormente, um nível de significância estatístico. E assim, na maioria dos casos, os teste estatísticos requerem o conhecimento de parâmetros que devem ser determinados previamente.

A hipótese básica da teoria dos erros, na estimativa do valor de uma grandeza a partir de um conjunto de medidas diretas, é que reduzindo-se os erros sistemáticos a níveis insignificantes, os erros aleatórios experimentais se distribuem normalmente e que a média aritmética das medidas é a melhor estimativa para o valor esperado da grandeza.

$$\bar{x} = \frac{1}{N} \sum_i^N x_i$$

³Resultados de medições realizadas pelo mesmo experimentador, com um mesmo instrumento de leitura direta e nas mesmas condições .

Uma vez que, segundo a Distribuição Gaussiana, 68% das medidas estarão centradas em torno do intervalo $(\mu - \sigma_x, \mu + \sigma_x)$, o desvio padrão σ_x é interpretado como a incerteza ou erro em cada uma das medidas.

A incerteza nessa estimativa ou, o erro na média, pode ser avaliado a partir da propagação de erros na adição .

Considerando-se a expressão da média como uma função linear de N variáveis aleatórias independentes $(x_1/N, x_2/N, \dots, x_N/N)$, cada uma com a mesma incerteza σ_x , de acordo com a fórmula da propagação de erros, a incerteza na média será dada por

$$\sigma_{\bar{x}}^2 = \sum_{i=1}^N \left(\frac{\partial \bar{x}}{\partial x_i} \right)^2 \sigma_x^2 = \frac{\sigma_x^2}{N}$$

ou seja,

$$\sigma_{\bar{x}} = \Delta \bar{x} = \frac{\sigma_x}{\sqrt{N}}$$

Assim,

A melhor estimativa para o valor esperado de um grandeza x , extraídas de uma grande amostra de N medidas $\{x_1, x_2, \dots, x_N\}$, diretas e independentes, é dada por:

$$\bar{x} \pm \Delta \bar{x}$$

onde

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (\text{média das medidas})$$

$$\Delta \bar{x} = \frac{\sigma_x}{\sqrt{N}} \quad (\text{incerteza na média})$$

$$\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (\text{desvio padrão})$$

• Intervalo de Confiança

De um outro ponto de vista estatístico, a estimativa para o valor esperado de uma grandeza pode ser expressa em termos de um **intervalo de confiança** (Neyman/Pearson - 1928). Ou seja, um intervalo ao qual se atribui uma determinada expectativa, ou **nível de confiança**, de que contenha o parâmetro valor esperado de uma grandeza.

Se um outro conjunto de N medidas $\{x'_1, x'_2, \dots, x'_N\}$, diretas e independentes, com o mesmo desvio padrão σ_x é determinado, uma outra média será obtida. Entretanto, no limite de um grande número de conjuntos de N medidas⁴, a média distribuir-se-á também segundo uma Gaussiana, com desvio padrão σ_x/\sqrt{N} ⁵.

Desse modo, diz-se que o **nível de confiança** de que o intervalo $(\bar{x} - \sigma_x/\sqrt{N}, \bar{x} + \sigma_x/\sqrt{N})$ contenha o valor esperado da grandeza é de 68.3%. Ou que ele é um intervalo de confiança de 68.3%.

⁴Ver apêndice A

⁵Se a variância também for desconhecida, pode-se admitir que \bar{x} tenderá a ser distribuída segundo

$$\mathcal{N}_{\bar{x}}(\mu, \Delta\bar{x}) = P(\bar{x}) = \frac{1}{\Delta\bar{x}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\bar{x}-\mu}{\Delta\bar{x}}\right)^2}$$

onde $\Delta\bar{x} = \frac{\hat{\sigma}_x}{\sqrt{N}}$ é um estimador para σ_x/\sqrt{N} e $\hat{\sigma}_x = \sqrt{\sum_i \frac{(x_i - \bar{x})^2}{N-1}}$ é um estimador para σ_x .

No caso de grandes amostras, a troca de σ_x por $\hat{\sigma}_x$ não acarreta grandes distorções na estimativa de um intervalo de confiança. Entretanto, para pequenas amostras, Gousset (1908), determinou que a variável reduzida

$$t = \frac{\bar{x} - \mu}{\hat{\sigma}_x/\sqrt{N}}$$

em vez de obedecer a uma distribuição normal padrão $\mathcal{N}_t(0,1)$, obedecia a uma distribuição, agora, chamada **Distribuição de Student**. Por isso, para pequenas amostras, as estimativas por intervalos de confiança deveriam ser corrigidas segundo a nova distribuição de t .

A tabela 1.1 mostra alguns intervalos de confiança típicos para uma variável x , cujas medidas são distribuídas normalmente, e os correspondentes níveis de confiança.

intervalo de confiança	nível de confiança
$(\bar{x} - \sigma_x/\sqrt{N}, \bar{x} + \sigma_x/\sqrt{N})$	68.3 %
$(\bar{x} - 2\sigma_x/\sqrt{N}, \bar{x} + 2\sigma_x/\sqrt{N})$	95.4 %
$(\bar{x} - 3\sigma_x/\sqrt{N}, \bar{x} + 3\sigma_x/\sqrt{N})$	99.7 %

Tabela

1.1: Intervalos de confiança relativos a Distribuição de Gauss.

Assim, alternativamente, pode-se sintetizar a estimativa, por um intervalo de confiança, do resultado da medição direta de uma grandeza como

A partir da amostra $\{x_1, x_2, \dots, x_N\}$ de um número N grande de medidas, diretas e independentes, de uma grandeza x , o nível de confiança de que o intervalo $(\bar{x} - \Delta\bar{x}, \bar{x} + \Delta\bar{x})$

$$\text{onde } \begin{cases} \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \\ \Delta\bar{x} = \frac{1}{N} \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \end{cases}$$

contenha o seu valor esperado é de 68.3%

1.3.2 Correlação no Cálculo de Eficiências em Experimentos de Altas Energias

Experimentos em Física de Altas Energias como aqueles que são realizados em grandes aceleradores de partículas, como o **Tévatron**⁶ do Fermilab⁷ nos EUA, ou o LEP⁸ no CERN⁹ na Suíça, constituem ambientes nos quais os métodos estatísticos são utilizados em toda sua plenitude, desde a simulação de eventos, de detectores, de sistemas de detecção até a extração e a validação de resultados experimentais.

Os exemplos em Física de Altas Energias citados nessas notas refletem a participação e as experiências dos autores na montagem e implementação de sistemas de detecção e aquisição de dados de um complexo de detectores chamado DØ¹⁰, construído por um grupo de cerca de 500 físicos, engenheiros e técnicos, durante os anos 80, no Tévatron do Fermilab.

A taxa de colisões próton-antipróton no Tévatron é tal que um dos componentes básicos de um experimento típico em altas energias é um sistema de seleção de eventos denominado “*Sistema de Trigger*”. Esse é um conjunto de circuitos eletrônicos, capaz de selecionar “*eventos interessantes*”, segundo certos critérios, em um nível adequado à capacidade do sistema de aquisição de dados do experimento.

No Tévatron, a ocorrência de um múon, resultante de colisões inelásticas de próton-antipróton, é o que se denomina de “**evento**” e, está associado a diversos atributos físicos como:

- momentum
- energia
- posição de ocorrência nos diversos setores do detector¹¹

Alguns dos critérios, com relação aos atributos de um múon, que definem “eventos interessantes” são:

⁶Anel de colisão de feixes de prótons (p) e antiprótons (\bar{p}), onde a estrutura e a criação de partículas são estudadas através da detecção de outras partículas resultantes de colisões inelásticas entre os feixes a energias da ordem de 1.0 TeV

⁷Fermi National Accelerator Laboratory.

⁸Anel de colisão de feixes de elétrons e pósitrons.

⁹European Laboratory for Particle Physics.

¹⁰Dentre os resultados obtidos por esse experimento destaca-se a observação e descoberta do quark TOP, um dos constituintes fundamentais do modelo atual (padrão) de classificação das partículas.

¹¹o que permite a reconstrução de sua trajetória ao longo do detector.

- energia acima de determinado valor
- parâmetro de impacto dentro de certos limites
- energia detectada próxima ao ponto de colisão

Tais critérios de seleção podem ser aplicados durante a aquisição de dados ou em um período posterior de análise de dados. Quando aplicados na análise, esses critérios são denominados genericamente de “cortes” e, permitem a obtenção de eventos realmente procurados, ou seja, definem a qualidade dos dados.

Desde que os cortes estão associados a medições realizadas por detectores e sistemas de aquisição não-ideais, torna-se necessário o conhecimento das ineficiências de detecção associadas a cada um dos cortes utilizados na obtenção de um bom conjunto de dados.

O método básico para a determinação das ineficiências consiste na geração simulada de dados, que devem ser submetidos também a uma simulação dos diversos detectores envolvidos no experimento. A seguir, a eficiência (ϵ) associada a um dado corte específico é calculada a partir do número de eventos que passaram (N_p) por todos os cortes aplicados ao dados excepto o corte específico e dos que não passaram (N_n) pelos cortes, segundo a razão

$$\epsilon = \frac{N_p}{N_p + N_n}.$$

O cálculo da eficiência associada a um determinado corte é um bom exemplo para ilustrar o problema da correlação .

Uma vez que o número total ($N_t = N_p + N_n$) de eventos é conhecido a priori, bastaria a determinação do número de eventos que passaram (N_p) pela totalidade de cortes excepto o corte específico para o cálculo da eficiência, segundo

$$\epsilon_t = \frac{N_p}{N_t}.$$

É claro que o valor médio da eficiência resultante das duas expressões é o mesmo. Entretanto, no primeiro caso, as quantidades N_p e N_n não são correlatas, enquanto no segundo, existe uma correlação entre N_p e N_t que deve ser levada em conta no cálculo dos erros estatísticos associados à eficiência.

Desde que,

$$\begin{cases} \frac{\partial \epsilon}{\partial N_p} = \frac{N_n}{N_t^2} \\ \frac{\partial \epsilon}{\partial N_n} = -\frac{N_p}{N_t^2} \end{cases}$$

e o número de eventos obedece a uma distribuição de Poisson, as incertezas são dadas por (Capítulo 2)

$$\begin{cases} \Delta N_p = \sqrt{N_p} \\ \Delta N_n = \sqrt{N_n} \\ \Delta N_t = \sqrt{N_t} \end{cases}$$

A incerteza ($\Delta\epsilon$) associada ao valor médio da eficiência será dada por

$$\begin{aligned} (\Delta\epsilon)^2 &= \left(\frac{N_n}{N_t^2}\right)^2 (\Delta N_p)^2 + \left(\frac{-N_p}{N_t^2}\right)^2 (\Delta N_n)^2 \\ &= \frac{N_p N_n}{N_t^3} = \frac{\epsilon(1-\epsilon)}{N_t} \end{aligned}$$

Entretanto, assumindo-se que não há correlação entre N_p e N_t .

$$\begin{cases} \frac{\partial \epsilon_t}{\partial N_p} = \frac{1}{N_t} \\ \frac{\partial \epsilon_t}{\partial N_t} = -\frac{N_p}{N_t^2} \end{cases}$$

implica

$$\begin{aligned} (\Delta\epsilon_t)^2 &= \left(\frac{1}{N_t}\right)^2 (\Delta N_p)^2 + \left(\frac{-N_p}{N_t^2}\right)^2 (\Delta N_t)^2 \\ &= \frac{N_p N_n + 2N_p^2}{N_t^3} \end{aligned}$$

ou seja, $\Delta\epsilon_t > \Delta\epsilon$, a incerteza pode ser superestimada.

Levando-se em conta que existe a correlação entre N_p e N_t ,

$$\begin{aligned}
 \sigma_{N_p N_t} &= \langle N_p N_t \rangle - \langle N_p \rangle \langle N_t \rangle \\
 &= \langle N_p (N_p + N_t) \rangle - \langle N_p \rangle \langle N_p + N_t \rangle \\
 &= \langle N_p^2 \rangle - \langle N_p \rangle^2 + \underbrace{\langle N_p N_n \rangle - \langle N_p \rangle \langle N_t \rangle}_{\sigma_{N_p N_n} = 0} \\
 &= \sigma_{N_p N_p} = (\Delta N_p)^2
 \end{aligned}$$

ou seja, tendo-se em conta as hipóteses corretas, a expressão

$$\begin{aligned}
 (\Delta \epsilon_t)^2 &= \underbrace{\left(\frac{\partial \epsilon_t}{\partial N_p} \right)^2 (\Delta N_p)^2 + \left(\frac{\partial \epsilon_t}{\partial N_t} \right)^2 (\Delta N_t)^2}_{(\Delta \epsilon)^2} + 2 \left(\frac{\partial \epsilon_t}{\partial N_p} \right) \left(\frac{\partial \epsilon_t}{\partial N_t} \right) \sigma_{N_p N_p} \\
 &= \frac{N_p N_n + 2N_p^2}{N_t^3} - 2 \frac{N_p^2}{N_t^3} = \frac{N_p N_n}{N_t^3}
 \end{aligned}$$

mostra que a incerteza no valor da eficiência pode ser calculada corretamente por ambos os modos.

Capítulo 2

Distribuições Básicas

1. Distribuição Binomial
2. Distribuição Poisson
3. Distribuição Gaussiana ou Normal
4. Outras Distribuições

Apresentaremos as distribuições mais usadas em estatística, suas definições e algumas de suas propriedades. Detalhes matemáticos deixaremos para o leitor interessado consultar a bibliografia [1, 2, 3].

2.1 Distribuição Binomial

Tomaremos o jogo das moedas como exemplo para nosso desenvolvimento nesse capítulo. A probabilidade para que uma moeda caia em um dos lados é de 50% ($1/2 + 1/2=1$). Temos então um caso de duas configurações. Se forem duas moedas teremos 4 configurações.

Vejam os quais são as permutações e combinações possíveis nessa configuração de n moedas e x resultados de um tipo. Se jogarmos n moedas, existem 2^n diferentes configurações que elas podem apresentar. Vejam caso a caso.

- 2^1 moeda = 2 configurações :

Cara (Ca) e Coroa (Co)

- 2^2 moedas = 4 configurações :

$$\begin{array}{ll} \text{Ca}^1 + \text{Ca}^2 & \text{Co}^1 + \text{Co}^2 \\ \text{Ca}^1 + \text{Co}^2 & \text{Co}^1 + \text{Ca}^2 \end{array}$$

portanto $(1/4 + 1/4 + 1/4 + 1/4) = 1$.

- 2^3 moedas = 8 configurações :

$$\begin{array}{ll} \text{Ca}^1 \text{Ca}^2 \text{Ca}^3 & \text{Co}^1 \text{Ca}^2 \text{Ca}^3 \\ \text{Ca}^1 \text{Co}^2 \text{Ca}^3 & \text{Co}^1 \text{Co}^2 \text{Ca}^3 \\ \text{Ca}^1 \text{Co}^2 \text{Co}^3 & \text{Co}^1 \text{Co}^2 \text{Co}^3 \\ \text{Ca}^1 \text{Ca}^2 \text{Co}^3 & \text{Co}^1 \text{Ca}^2 \text{Co}^3 \\ \dots & \dots \end{array}$$

- 2^n moedas = 2^n configurações :

e assim por diante.

Sabendo-se que elas são igualmente prováveis, as probabilidades para cada caso são:

$$\begin{array}{ll} \frac{1}{2^1} = \frac{1}{2} & \implies 50 \% \text{ para cada configuração.} \\ \frac{1}{2^2} = \frac{1}{4} & \implies 25 \% \text{ para cada configuração.} \\ \frac{1}{2^3} = \frac{1}{8} & \implies 12,5 \% \text{ para cada configuração.} \\ \dots & \dots \\ \frac{1}{2^n} = \frac{1}{2^n} & \implies \frac{100}{2^n} \% \text{ para cada configuração.} \end{array}$$

É possível construir arranjos em diversas configurações com n objetos (moedas ou qualquer outro objeto). Um arranjo qualquer de n objetos é chamado de permutação. Pode-se também denominar de permutação a troca de posição de tais objetos de um arranjo. Podemos também construir arranjos de r -objetos (com $r \leq n$) dos n -objetos e a ele chamarmos de permutação r . O número de permutações de n objetos tomando r objetos de cada vez é dado pela expressão:

$$P(n, r) = n(n-1)(n-2) \cdots (n-r+2)(n-r+1) = \frac{n!}{(n-r)!}$$

Se tivermos repetição de uma configuração ou mais, então o número de permutações possíveis para os n objetos, com n_1 repetições ou n_2 repetições é:

$$PR(n, r) = \frac{n!}{n_1!n_2! \cdots nr!}$$

O número de diferentes combinações $C(n, r)$ é dada pelo número de permutações enumeradas $P(n, r)$ dividida pelo fator de degenerescência $r!$ das permutações.

$$C(n, r) = \frac{P(n, r)}{(r)!} = \frac{n!}{r!(n-r)!} \equiv \binom{n}{r} \quad (2.1)$$

A probabilidade $P(r, n)$ para observar r moedas com o lado da coroa para cima e $n-r$ com o lado da cara para cima, é o produto do número de diferentes combinações $C(n, r)$ que contribuiriam para esse conjunto de observações vezes a probabilidade para cada uma das combinações ocorrerem.

A probabilidade para obtermos moedas idênticas com cara para cima não é igual a probabilidade $q = 1 - p$ de não obtê-las ou de obtê-las com coroa para cima. A probabilidade para cada uma das combinações de r moedas de cara para cima e $n-r$ moedas com coroa para cima é $p^r q^{n-r}$. Então a probabilidade $P_B(r, n, p)$ para observar r dos n itens com probabilidade p é dada pela Distribuição Binomial :

$$P_B(r, n, p) = \binom{n}{r} p^r q^{n-r} = \frac{n!}{r!(n-r)!} p^r (1-p)^{(n-r)} \quad (2.2)$$

onde $q = 1 - p$. Então, a Distribuição Binomial $P_B(r, n, p)$ nos dá a probabilidade de observar r dos n itens com probabilidade p . No nosso caso, das moedas, poderia se dizer que a Distribuição Binomial B é dada pela probabilidade P de observar r moedas das n moedas que têm a probabilidade p de ser observada com cara para cima.

A média de uma Distribuição Binomial é dada por

$$\mu = \sum_{r=0}^n \left[r \frac{n!}{r!(n-r)!} p^r (1-p)^{(n-r)} \right] = np \quad (2.3)$$

A variância σ^2 da Distribuição Binomial é dada por:

$$\sigma^2 = \sum_{r=0}^n \left[(r - \mu)^2 \frac{n!}{r!(n-r)!} p^r (1-p)^{(n-r)} \right] = np(1-p) \quad (2.4)$$

2.2 Distribuição de Poisson

A Distribuição de Poisson é uma aproximação da Distribuição Binomial quando o número médio de eventos observados é muito menor que o número real de eventos. Se aplica portanto a muitos casos especiais de análise, principalmente quando procuramos um sinal de alguma coisa ainda não detectada, ou seja, num processo de descoberta.

Um exemplo é o número de partículas de um feixe que incide sobre um alvo qualquer ($\sim 10^{13}$ eventos/segundo), que é muito maior que aquele registrado pelos detectores ($\sim 10^4$ partículas incidentes por segundo). Considerando então a Distribuição Binomial dada pela expressão (2.2) e o limite de n muito grande ($n \rightarrow \infty$), número de eventos $p \rightarrow 0$ (ou $p \ll 1$ número de sucessos ou acertos), $\lambda = np = \mu$ constante, temos,

$$P_B(r, n, p) = \frac{1}{r!} \frac{n!}{r!(n-r)!} p^r (1-p)^{-r} (1-p)^n \quad (2.5)$$

$$\lim_{p \rightarrow 0} (1-p)^r = \lim_{p \rightarrow 0} [(1-p)^{1/p}]^\lambda = \left(\frac{1}{e}\right)^\lambda = e^{-\lambda} \quad (2.6)$$

$$\lim_{p \rightarrow 0} P_B(r, n, p) = P_P(r, \lambda) = \left(\frac{\lambda^r}{r!}\right) e^{-\lambda} \quad (2.7)$$

Assim a Distribuição de Poisson é dada pela expressão,

$$P_P(r, \lambda) = \left(\frac{\lambda^r}{r!}\right) \exp(-\lambda) \quad (2.8)$$

A expressão para a média de uma Distribuição de Poisson é dada pelo mesmo parâmetro $\lambda = \mu$ que aparece na expressão da probabilidade (2.8). Pois, avaliando-se o valor esperado de r temos,

$$\langle r \rangle = \sum_{r=0}^{\infty} \left[r \left(\frac{\lambda^r}{r!}\right) e^{-\lambda} \right] = \lambda e^{-\lambda} \sum_{r=0}^{\infty} \left[\frac{\lambda^{r-1}}{(r-1)!} \right] = \quad (2.9)$$

$$= \lambda e^{-\lambda} \sum_{y=1}^{\infty} \frac{\lambda^y}{y!} = \lambda \quad (2.10)$$

A variância é dada por

$$\sigma^2 = \lim_{p \rightarrow 0} [np(1-p)] = \lambda = \mu \quad (2.11)$$

Uma das mais importantes características da Distribuição de Poisson é a sua reprodutibilidade, ou seja, se dois ou mais processos independentes de Poisson, como a contagem de partículas que decaem e o *background* associado, com parâmetros λ_a e λ_b , independentemente produzem r_a e r_b eventos, ocorrerem simultaneamente, os eventos resultantes estarão distribuídos segundo uma Distribuição de Poisson de parâmetro $\lambda = \lambda_a + \lambda_b$.

Uma vez que a probabilidade de ocorrência combinada de um total de r eventos é dada por

$$P(r) = \sum_{r_a=0}^r P(r_a; \lambda_a) P(r - r_a; \lambda_b) \quad (2.12)$$

$$= e^{-\lambda_a} e^{-\lambda_b} \sum_{r_a=0}^r \frac{\lambda_a^{r_a} \lambda_b^{r-r_a}}{r_a! (r-r_a)!} \quad (2.13)$$

$$= e^{-(\lambda_a + \lambda_b)} \frac{(\lambda_a + \lambda_b)^r}{r!} \underbrace{\sum_{r_a=0}^r \frac{r!}{r_a! (r-r_a)!} \left(\frac{\lambda_a}{\lambda_a + \lambda_b} \right)^{r_a} \left(\frac{\lambda_b}{\lambda_a + \lambda_b} \right)^{r-r_a}}_1 \quad (2.14)$$

$$= e^{-\lambda} \frac{\lambda^r}{r!} \quad (2.15)$$

onde $\lambda = \lambda_a + \lambda_b$.

2.3 Distribuição Gaussiana ou Normal

Essa é a distribuição mais utilizada em aplicações estatísticas. Do ponto de vista matemático, é uma aproximação da Distribuição Binomial, no caso em que o número de eventos é muito grande e, portanto, a probabilidade de se observar o evento desejado é também muito grande, *i.e.*, $np \gg 1$.

Ela é definida como:

$$P_G(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right] \quad (2.16)$$

onde média e o desvio padrão são dados por μ e σ e, suas principais características já foram apresentadas na sub-seção 1.2.2 do capítulo 1,

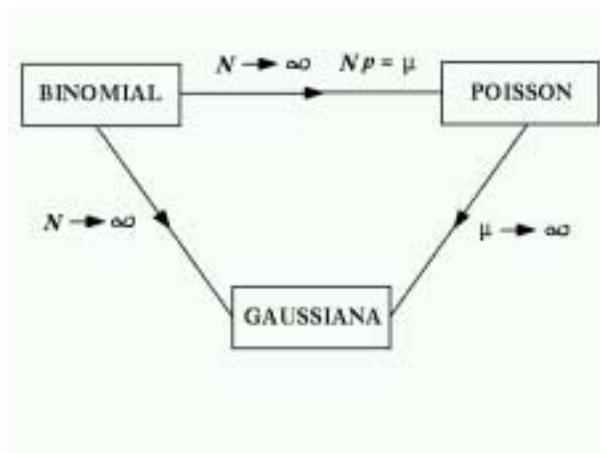


Figura 2.1: As ligações entre as principais distribuições: Binominal, Poisson e Gaussiana.

A Fig. 2.1 mostra como se relacionam as três distribuições mais usadas.

E, na tabela a seguir resumimos as três distribuições apresentadas e suas principais propriedades.

Distribuições			
	Binomial	Poisson	Gaussiana
Definição	$\frac{n!}{r!(n-r)!} p^r (1-p)^{(n-r)}$	$\left(\frac{\lambda^r}{r!}\right) \exp(-\lambda)$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$
Média	$\mu = Np$	$\mu = \lambda$	μ
Variância	$\sigma^2 = Npq$	$\sigma^2 = \lambda$	σ^2
Desvio Padrão	$\sigma = \sqrt{Npq}$	$\sigma = \sqrt{\lambda}$	σ

2.4 Outras Distribuições

Muitas outras distribuições são utilizadas, principalmente, no ajuste de parâmetros de teorias aos dados de um experimento. Uma muito usada para descrever o comportamento de processos ressonantes é a chamada Distribuição Lorentziana, Breit-Wigner ou de Cauchy,

$$P_L(x, \mu, \Gamma) = \frac{1}{\pi} \frac{\Gamma/2}{(x - \mu)^2 + (\Gamma/2)^2} \quad (2.17)$$

onde μ é a média e Γ é a sua FWHM (*full width at half maximum*).

Essa distribuição tem a particularidade de que decai tão lentamente em torno de seu valor médio μ que não possui uma variância. Entretanto, sua largura pode ser facilmente calculada por seus valores a meia altura.

A figura 2.2 mostra a comparação de uma Lorentziana (linha pontilhada) e uma Gaussiana (linha sólida) de mesma largura Γ .

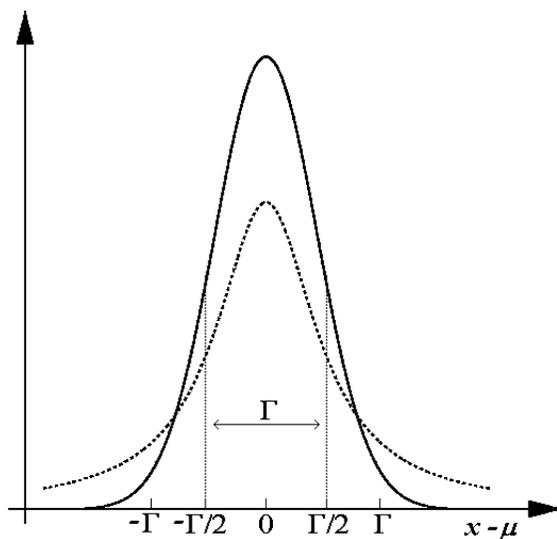


Figura 2.2: Distribuições de Lorentz (linha pontilhada) e Gauss (linha sólida) de mesma largura Γ .

Capítulo 3

Ajustes de Curvas a um Conjunto de Dados

1. Método dos Momentos
2. Método da Máxima Verossimilhança
3. Método dos Mínimos Quadrados
4. Teste de χ^2

O ajuste de uma curva, ou genericamente, de uma função que possui alguns parâmetros a um conjunto de dados é sempre realizado em uma região física bem delimitada de tal forma que se pode inicializar o processo com certos valores dos parâmetros. A determinação desses parâmetros constitui um objeto específico de estudo e, dada a complexidade que aparece no processo de ajuste da função aos dados experimentais, torna-se necessário o entendimento dos principais métodos estatísticos que podem ser utilizados num ajuste. Seguiremos de perto o exposto na referência [2].

Consideremos o exemplo de uma distribuição angular consistindo de um conjunto de valores $\{\cos \theta\}$ para cada interação, onde θ_i é o ângulo observado que as partículas fazem com uma determinada direção. (Usamos $\cos \theta$ e não de θ como variável já que uma distribuição no espaço nos dá uma distribuição uniforme em $\cos \theta$ e não em θ).

Cada $\cos \theta_i$ seria associado a um valor peso $w_i \geq 1$ para corrigir as ineficiências de detecção. Chamamos à atenção para algumas questões que deverão ser levantadas quando tratamos um problema de ajuste de curvas com dados:

1. Testes de hipóteses:

Os dados são consistentes com uma distribuição angular da forma?

$$\frac{dn}{d \cos \theta} = a + b \cos^2 \theta \quad (3.1)$$

Qual a significância estatística desse acordo?

2. Estimativas de parâmetros:

Supondo que a equação (3.1) seja razoável para testar nosso experimento, que valores de a e b seriam escolhidos para fornecer a melhor descrição dos dados?

Quão precisos são a e b ?

3. Normalização:

As normalizações dos dados e da teoria têm que ser as mesmas. Por exemplo, a distribuição angular dada pela equação (3.1) deve obedecer à condição

$$N = \int_{-1}^{+1} (a + b \cos^2 \theta) d(\cos \theta) = 2(a + b/3) \quad (3.2)$$

onde N é o número total de eventos observados.

Logicamente, um teste de hipótese precede a estimativa de parâmetros, entretanto, é praticamente impossível estabelecer uma opinião segura sobre um teste das hipóteses sem fazer uma avaliação dos parâmetros e, portanto, fazer um ajuste com os dados.

Os principais métodos estatísticos para a estimativa de parâmetros num ajuste de funções são:

- dos Momentos,
- da Máxima Verossimilhança
- dos Mínimos Quadrados.

3.1 Método dos Momentos

O método dos momentos consiste em calcular o valor médio dos dados de uma certa quantidade. Esta média é analiticamente relacionada ao parâmetro de interesse. Retomando nosso exemplo, com a distribuição angular dada pela equação (3.1), achamos a média que resulta em:

$$\overline{\cos^2 \theta} = \frac{5 + 3b/a}{5(3 + b/a)} \quad (3.3)$$

ou ainda, isolando b/a

$$\frac{b}{a} = \frac{5(3\overline{\cos^2 \theta} - 1)}{3 - 5\overline{\cos^2 \theta}} \quad (3.4)$$

E, para avaliar a média diretamente de novos dados fazemos:

$$\overline{\cos^2 \theta} = \frac{1}{n} \sum_{i=1}^n \cos^2 \theta_i$$

e o erro associado a essa média é :

$$\delta = \left[\frac{1}{n(n-1)} \sum_{i=1}^n \left(\cos^2 \theta_i - \overline{\cos^2 \theta} \right)^2 \right]^{1/2} \quad (3.5)$$

Façamos alguns comentários sobre o método:

- **Simplicidade:** O método é muito fácil de ser aplicado e não exige nenhum procedimento de maximização.
- **Estimativa dos Erros:** Os erros deste método são obtidos pelas diferenças sobre os valores observados $\cos^2 \theta_i$ ou, alternativamente, dos ajustes da distribuição usada. Pode-se calcular como função de b/a . O erro é simétrico na variável $\cos^2 \theta_i$: Nós devemos então transformar essa variável para a variável b/a que é em geral assimétrica.
- **Eventos Individuais:** O procedimento de momentos faz uso de eventos individuais e já que não existe a necessidade de se fazer uma escolha de compromisso, de tamanho do bin com o qual se plota um histograma.

3.2 Método da Máxima Verossimilhança

Esse é o método mais poderoso para encontrar-se valores de parâmetros desconhecidos no processo de ajuste de uma curva a um conjunto de dados. Para entendermos como funciona o método tomemos a equação (3.1) de uma distribuição angular muito comum em Física de Partículas e reescrevamo-la como

$$y = A \left[1 + \left(\frac{b}{a} \right) \cos^2 \theta \right] \quad (3.6)$$

onde A é o fator de normalização tal que

$$\int_{-1}^{+1} y d \cos \theta = 1$$

isto é

$$A = \frac{1}{2[1 + (b/3a)]} \quad (3.7)$$

Temos então a função y que representa a distribuição normalizada. Para a utilização do método é fundamental que a distribuição seja normalizada pois, sem o fator A na equação de y , não teríamos um máximo absoluto. Além disso, há que se levar em consideração outros pontos importantes no processo de normalização, tal como a redução do número de parâmetros.

O objetivo agora é determinar o parâmetro b/a na equação (3.6).

Para o i -ésimo evento calculamos

$$y_i = A(1 + (b/a) \cos^2 \theta_i)$$

que é a densidade de probabilidade para observação deste evento i , e é uma função de b/a que desejamos determinar. Define-se então a função de verossimilhança \mathcal{L} como o produto de y_i para todos os eventos da amostra, i.e.,

$$\mathcal{L}(b/a) \equiv \prod_{i=1}^n y_i$$

Então para qualquer valor de b/a , \mathcal{L} é a densidade de probabilidade para obter um particular conjunto de valores de $\cos \theta_i$ que tenha sido observado no experimento. Esse valor de $\mathcal{L}(b/a)$ é maximizado como função de b/a para encontrar-se o melhor valor de b/a . Vejamos um exemplo.

3.2.1 Separação de píons e káons

Esse é um problema que aparece quando tentamos determinar a composição de partículas de um dado conjunto de traços. Por simplicidade suporemos que nossa amostra contém somente uma mistura de píons e káons. No enfoque do Método da Máxima Verossimilhança vamos maximizar $\log \mathcal{L}$ onde,

$$\mathcal{L} = \prod \left\{ f_\pi \exp \left[-\frac{(x_i - x_\pi)^2}{2\sigma^2} \right] + (1 - f_\pi) \exp \left[-\frac{(x_i - x_k)^2}{2\sigma^2} \right] \right\} \quad (3.8)$$

- Nessa expressão x_π e x_k são posições esperadas dos picos do pión e do káon para a variável que estamos usando para distinguir as partículas (por exemplo, dE/dx);
- $(x_i \pm \sigma)$ é o valor medido para o i -ésimo traço;
- f_π é a fração de píons que esperamos determinar;
- O produto é estendido sobre os N traços da amostra e, omitimos as constantes irrelevantes de nossa definição de \mathcal{L} .

A chamada função de verossimilhança estendida inclui um fator extra para a probabilidade de obtenção de uma amostra do tamanho N de uma distribuição de Poisson de média ϕ , *i.e.*,

$$\mathcal{L}_E = \frac{e^{-\phi} \phi^N}{N!} \mathcal{L}$$

Nossas variáveis são agora f_π e ϕ ou equivalentes a:

$$\begin{aligned} \phi_\pi &= \phi f_\pi \\ \phi_k &= \phi(1 - f_\pi) \end{aligned}$$

onde $\phi_{\pi(k)}$ é o número de píons e kaons.

Em termos de ϕ_π e ϕ_k a verossimilhança estendida pode equivalentemente ser escrita como:

$$\mathcal{L}_E = \frac{e^{-(\phi_\pi + \phi_k)}}{N!} \prod \left\{ \phi_\pi \exp \left[-\frac{(x_i - x_\pi)^2}{2\sigma^2} \right] + \phi_k \exp \left[-\frac{(x_i - x_k)^2}{2\sigma^2} \right] \right\} \quad (3.9)$$

Estimando os números de píons e káons, temos 3 relevantes fatores:

1. Qual a ambiguidade na resolução de píons e káons experimentalmente? Isso depende crucialmente da grandeza de $(x_\pi - x_k)/\sigma$;
2. Qual é a nossa estimativa do número de píons em nossa amostra particular? Isto é respondido pelo método da máxima verossimilhança;
3. Qual é a nossa estimativa do número de píons para experimentos semelhantes ao nosso? Aqui o método deve ser estendido por necessidade de se obter melhores resultados.

3.3 Método dos Mínimos Quadrados

Vamos supor que dispomos de uma distribuição de dados, num histograma, das observações y_i de uma variável x , como, por exemplo, o número de partículas em intervalos (*bins*) de momentum.

Denotaremos por y_i^{obs} as entradas do histograma, onde i indica o particular *bin* do mesmo. Como habitualmente, tentaremos ajustar os dados por uma forma funcional $y_i^{\text{th}}(\alpha_j)$, onde α_j são os parâmetros da expansão teórica. Então, definiremos uma função como,

$$S = \sum_{i=1}^{\text{bins}} \left[\frac{y_i^{\text{obs}} - y_i^{\text{th}}(\alpha_j)}{\sigma_i} \right]^2 \quad (3.10)$$

onde σ_i são as incertezas esperadas em cada *bin*. Se a teoria e os dados experimentais estão de acordo, o valor de S é muito pequeno.

O Método dos Mínimos Quadrados pode ser derivado a partir do Método da Máxima Verossimilhança, se o conjunto de medidas obedecem à Distribuição de Gauss. Entretanto, de modo alternativo, a minimização da função S pode ser tomada como a hipótese fundamental do processo de ajuste.

Observe que é altamente indesejável ter menos de 5 eventos em qualquer *bin*. Se temos uma única medida $y_i^{\text{obs}} \pm \sigma$ de uma particular quantidade y^{th} , contaríamos

$$S = \left(\frac{y^{\text{obs}} - y^{\text{th}}}{\sigma} \right)^2 \quad (3.11)$$

e essa tem um mínimo com $S = 0$ para $y^{\text{obs}} = y^{\text{th}}$ e $S = 1$ quando $y^{\text{obs}} = y^{\text{th}} \pm \sigma$.

Se temos 2 medidas y_1 e y_2 de uma só quantidade e com erros iguais a essa o melhor valor estimado é:

$$S = \left(\frac{y_1 - y^{\text{th}}}{\sigma} \right)^2 + \left(\frac{y_2 - y^{\text{th}}}{\sigma} \right)^2$$

minimizando com $dS/dy^{\text{th}} = 0$ temos $y^{\text{th}} = (y_1 + y_2)/2$ e o erro é $\sigma/\sqrt{2}$, deduzido de $S = S_{\text{min}} + 1$.

Do ponto de vista prático, uma vez que os σ_i 's, em geral, dependem dos parâmetros α_j 's a se determinar, o processo de minimização torna-se muito complicado e, usualmente, substitui-se as incertezas esperadas pelas observadas, que são iguais aos próprios conteúdos de cada *bin*, admitindo-se que esses conteúdos se distribuem segundo a Distribuição de Poisson. Ou seja, minimiza-se a função

$$S = \sum_{i=1}^{\text{bins}} \left[\frac{y_i^{\text{obs}} - y_i^{\text{th}}(\alpha_j)}{y_i^{\text{obs}}} \right]^2 \quad (3.12)$$

Há inúmeras funções apropriadas para receber o tratamento dos mínimos quadrados no ajuste de funções a um conjunto de dados. Uma lista dessas funções é a seguinte:

1. Reta: $Y = a_0 + a_1X$
2. Parábola: $Y = a_0 + a_1X + a_2X^2$
3. Cúbica: $Y = a_0 + a_1X + a_2X^2 + a_3X^3$
4. Polinômio de Grau n : $Y = a_0 + a_1X + a_2X^2 + \dots + a_nX^n$
5. Curva Geométrica: $Y = aX^b$
6. Geométrica Modificada: $Y = aX^b + c$
7. Hipérbole: $Y = \frac{1}{a_0 + a_1X}$
8. Exponencial: $Y = ab^X$
9. Exponencial Modificada: $Y = ab^X + c$

3.4 Teste de χ^2

A utilização de qualquer um dos métodos para estimação de parâmetros não estabelece um critério para julgar a significância estatística ou qualidade de um determinado ajuste.

No caso do Método dos Mínimos Quadrados, o procedimento mais comum para julgar a qualidade de um ajuste de função é denominado Teste de χ^2 , que pode ser estabelecido a partir da distribuição de probabilidades da variável S .

Se a função de ajuste a N dados tem m parâmetros, a variável S , também escrita como χ^2 , obedece a uma Distribuição de χ^2 com $\nu = N - m$ graus de liberdade (apêndice A).

Uma vez que o valor esperado de χ^2 é igual ao número de graus de liberdade da distribuição, o critério usual para avaliar o bom acordo entre a função de ajuste e os dados é que χ^2/ν seja próximo da unidade.

Apresentaremos a seguir o Método de Monte Carlo, no qual é implícito o conhecimento das noções dadas no capítulo.

Capítulo 4

Monte Carlo

1. Porque Utilizar Métodos de Monte Carlo?
2. Aplicações Simples de Monte Carlo
 - Cálculo de uma Integral
 - Momentum Transversal Médio
 - Espaço de Fase
 - Geração de Eventos

Vamos primeiramente tentar responder a questão: O que é um Método de Monte Carlo? São algoritmos numéricos que se utilizam da geração de números aleatórios. Quaisquer situações que envolvam processos aleatórios, como a simulação de amostras aleatórias, ou mesmo aqueles que não têm caráter probabilístico como uma simples integração numérica, podem se utilizar de um Método de Monte Carlo. Não há experimento em Física de Altas Energias que não seja projetado, realizado e analisado à luz de programas de computação que utilizam-se de Métodos de Monte Carlo. Uma série de exemplos seguirão para ilustrar o método.

Inicialmente, vamos mostrar como funciona o Método de Monte Carlo no cálculo de uma integral do tipo:

$$I = \int_a^b y(x)dx .$$

onde $y(x)$ é uma função contínua qualquer. Usando o método numérico mais simples, como mostra a figura 4.1, divide-se o intervalo $[a, b]$ em partes iguais e calcula-se o valor

estimado de I por

$$\frac{b-a}{n} \sum_{i=1}^n y(x_i) \quad (4.1)$$

onde

$$x_i = a + \frac{(i-1/2)(b-a)}{n}.$$

Esse talvez não seja, o melhor método para o cálculo da integral I mas é o que mais se aproxima de um Método de Monte Carlo.

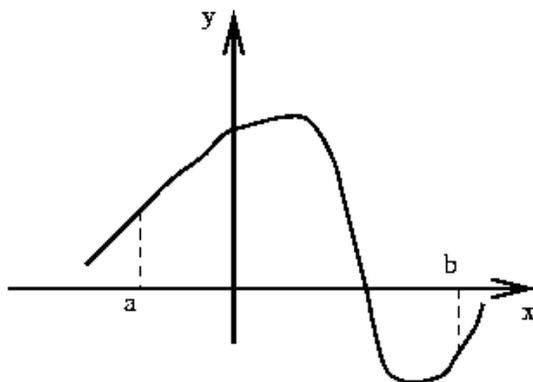


Figura 4.1: Exemplo de curva para cálculo da integral e divisão do intervalo de integração em partes iguais.

O correspondente Método de Monte Carlo também consiste no cálculo da soma (4.1), mas no lugar de se determinar $y(x_i)$ em intervalos iguais, calcula-se aleatoriamente os números x_i , no intervalo $[a, b]$, através de,

$$x_i = a + r_i(b-a) \quad (4.2)$$

onde r_i 's são membros de uma série de números aleatórios uniformemente distribuídos entre 0 e 1, que podem ser obtidos por qualquer gerador disponível. Esses números aleatórios estão embutidos em todo programa de Monte Carlo e é importante conhecer suas propriedades em relação ao problema tratado. Deixamos como exercício a verificação das propriedades de alguns geradores.

Exercício 4.1 Examine o seu gerador de números aleatórios, comparando-o graficamente com outros geradores. Que conclusões voce tira?

4.1 Porque Utilizar Métodos de Monte Carlo?

Se usarmos n números aleatórios para obter um valor para a integral I , a resposta tem uma precisão proporcional a $1/\sqrt{n}$, enquanto que para um dos métodos numéricos mais simples é de $1/n^2$. Qual é então o ponto que nos leva a usar um Monte Carlo? Se tivéssemos somente integrais em uma dimensão não necessitaríamos de um Método de Monte Carlo. Mas a realidade é outra e nos deparamos frequentemente com integrais múltiplas. Nesses casos, a precisão dos métodos numéricos mais simples é de $n^{-2/d}$, onde d é a dimensão das integrais, e no Método de Monte Carlo continua a ser $n^{-1/2}$.

Cada vez que se calcula a mesma integral se obtém um número ligeiramente diferente, no entanto, pode-se associar a essa diferença um erro. Na prática, o erro obtido pode ser muito pequeno, satisfazendo plenamente os objetivos do problema.

Uma questão interessante é que a integral converge para o resultado mais rapidamente do que em qualquer outro método. Por exemplo, para uma integração tríplice pelo método de Gauss, usando-se apenas 10 pontos para cada integral, teríamos 1000 iterações ou cálculos da função, no entanto, usando o Método de Monte Carlo com um número menor de pontos chegaremos mais rapidamente ao resultado e com melhor precisão.

Procura-se o máximo de precisão no método olhando-se para:

- Extratificação: dividindo-se o intervalo em x , por exemplo, em duas partes, o número de pontos gerados em cada uma das regiões é reduzido;
- Amostragem: transformando-se o integrando em uma função que varie mais lentamente.

Para integrar no caso de n dimensões, precisamos de n conjuntos de números aleatórios para gerar o ponto no qual o integrando é avaliado. Muitas regiões de integração não têm limites independentes uns dos outros. Portanto, para avaliar

$$\int z(x, y) dx dy$$

para a região de integração triangular bidimensional, como mostra a figura 4.2 temos que escolher os intervalos de variação de cada variável, corretamente, por exemplo, x escolhido entre 0 e 1 e uma condição para o valor de y , tipo $y < x$.

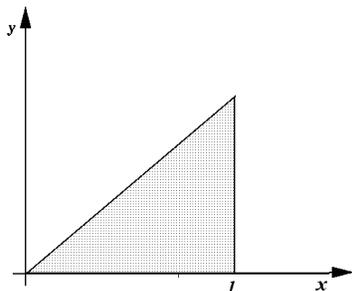


Figura 4.2: Exemplo de integração sob uma reta.

4.2 Aplicações Simples de Monte Carlo

São inúmeras as aplicações de Métodos de Monte Carlo. Escolhemos algumas mais próximas de nossas atividades em Física de Altas Energias.

4.2.1 Cálculo de uma Integral

Vamos continuar a mostrar como funciona o Método de Monte Carlo agora com o exemplo explícito da seguinte integral:

$$I = \int_0^1 x^3 dx$$

Usaremos N números gerados aleatoriamente, que por sua vez dão origem ao conjunto $\{x_i\}$ de pontos, os quais usados para avaliar a integral. Repete-se esse procedimento um certo número de vezes e a precisão requerida é obtida facilmente. Vejamos cada passo:

1. Usa-se a expressão (4.1) para obter-se o valor da integral e a expressão (4.2) para o cálculo dos pontos x_i 's, o que significa:

$$x_i = r_i \quad \text{e} \quad I = \frac{1}{N} \sum_{i=1}^N x_i^3;$$

2. Constrói-se a tabela de N valores de x_i e calcula-se para cada um dos x_i os seus correspondentes x_i^3 ;

3. Faz-se a soma de todos os x_i^3 .
4. Faz-se a comparação entre o valor dado pela (4.1) e o valor obtido pelo cálculo exato que é 0.25.

Exercício 4.2 Faça este cálculo com 10, 50 e 100 pontos e verificar como se aproxima cada resultado obtido com o Monte Carlo e o valor verdadeiro.

4.2.2 Momentum Transversal Médio

Na Física de Altas Energias é observado que o momentum transversal médio das partículas produzidas em interações aumenta lentamente quando a energia incidente aumenta. Um gráfico de particular interesse é o do $\langle p_T \rangle$ como função de p_L das partículas. Para estudar esse gráfico e seus problemas, supomos que as partículas são produzidas no c.m. com momenta p , uniforme no intervalo de 0–1 GeV/c, e são também uniformes em $\cos \theta$. Nenhuma dessas hipóteses é realista mas servem para simplificar nosso problema e tornar simples nossa aplicação. O enfoque via Monte Carlo para esse problema consiste em gerar um grande número (~ 10.000) de traços cujo momentum é dado por números aleatórios, enquanto o ângulo de produção é definido em termos do segundo número aleatório r por:

$$\cos \theta = 2r - 1$$

com p_L e p_T calculados como:

$$p_L = p \cos \theta$$

e

$$p_T = p \sin \theta$$

obtemos o $\langle p_T \rangle$ para o conjunto de eventos com intervalo de p_L especificado.

4.2.3 Espaço de Fase

É possível encontrar aplicações para o o Método de Monte Carlo em quase todas as áreas do conhecimento humano. Para projetar um experimento, toda a sua instrumentação antes de ser construída deve passar por uma fase de simulação. Um exemplo é o cálculo do momentum dos traços que representam a passagem da partícula pelo detector, com o

auxílio dos campos magnéticos das regiões respectivas, que pode ser obtida, medindo-se o ângulo de curvatura. Introduce-se as características das camadas de cada detector por onde passou a partícula e seu desvio causado por múltiplas interações. Compara-se os resultados obtidos a modelos existentes. Testa-se programas especiais de computação dos experimentos. Faz-se estudos e estimativas de contaminação faz-se correções geométricas e calcula-se o espaço de fase para cada interação.

Os Métodos de Monte Carlo são os mais eficientes, precisos e fáceis de trabalhar e, portanto, os que melhores se adaptam ao tipo de integração que temos. A sua aplicação em Física de Altas Energias encontrou um grande sucesso com o programa FOWL [4] seguido por outros programas mais acabados e mais complexos como o GEANT do CERN. Vejamos então algumas das características de um Método de Monte Carlo.

- (a) Mais rápida convergência que qualquer outro método: o erro decresce com $1/(\sqrt{N})$, onde N é o número de pontos gerados na amostragem.
- (b) Eficiente e versátil: O método se aplica aos problemas que aparecem na Física de Partículas.
- (c) A implementação de programas é simples.
- (d) O método trata os eventos exatamente da mesma forma que são tratados os dados de um experimento, simulando o próprio experimento.

Vamos dar uma olhada no Espaço de Fase ou a região física onde atua a dinâmica do choque de partículas. Para isso nos basearemos no FOWL, um dos primeiros programas de Monte Carlo para partículas [4]. O programa gerador de eventos de Monte Carlo FOWL, gera eventos segundo uma lei. Vejamos qual. O programa como está concebido, tem suas variáveis de massa e dos ângulos fornecidas por uma lei uniforme, linear, e limitada pelas fronteiras naturais da Física. Isto quer dizer que:

- Dispomos de números aleatórios x , uniformemente distribuídos;
- As variáveis aleatórias são: as massas M_i , os ângulos θ e ϕ e a posição no espaço de cada vetor, fornecido segundo os valores de x , então uniformemente limitado pelos limites físicos. Exemplo

$$0 \leq \theta \leq \pi$$

$$0 \leq \phi \leq 2\pi$$

$$m_1 + m_2 \leq \sqrt{s_1} \leq \sqrt{s} - m_3$$

$$M_i = M_{12} = \sqrt{s_1}$$

Isso corresponde, no programa de Monte Carlo, a colocar “contadores”, por todos os possíveis locais, sem se preocupar em ter ou querer observar um tipo específico de evento, por exemplo, querer ver somente o aspecto difrativo da interação, ou da pionização, etc. Fazendo uma experiência hipotética, com o Monte Carlo FOWL, podemos traduzir nossa necessidade, a de querer observar somente o aspecto difrativo, sem que o espaço de fase seja tocado. Isto quer dizer, colocar um contador “difrativo”. Ou ainda criar um enriquecimento nas amostras desejadas. Tudo se passa como se fizéssemos a integração por Gauss fazendo as partições dentro do intervalo de integração para termos o maior número possível de eventos onde nossa função pode ter uma grande variação. Passemos ao problema fazendo um sumário da descrição do método segundo a referência [5].

Integração Numérica no Espaço de Fase

A integral no espaço de fase (EF) é um dos muitos problemas de cálculo numérico que se depara a Física de Partículas. A integral no EF é definida por

$$I_n = \int_V \cdots \int \prod_{i=1}^n \frac{d^3 p_i}{2E_i} \delta^4(p - p_1 - \cdots - p_n) T \quad (4.3)$$

onde,

- T é o elemento de matriz ao quadrado ou a representação acabada do processo dinâmico em estudo,
- p_i, E_i são os momenta e as energias das partículas envolvidas no processo,
- V representa o volume do espaço de fase como um todo.
- $(3n - 4)$ é o número de variáveis independentes, onde n é o número de partículas do estado final. Por exemplo, uma reação do tipo $a + b \rightarrow 1 + 2$ ($n = 2$) possui 2 variáveis independentes.

A avaliação da função $f_n(\Phi)$ é feita para um conjunto de pontos gerados aleatórios. A densidade dos eventos gerados aleatoriamente é dada por $g(\Phi)$ no Espaço de Fase. Um evento aqui é um conjunto de n vetores momenta $\vec{p}_1, \cdots, \vec{p}_n$ em qualquer sistema de

referência satisfazendo a conservação de quadrimomentos e as condições de camada de massa dada por $\vec{p}_i^2 = m_i^2, i = 1, \dots, n$.

A densidade $g(\Phi)$ é definida como a usual densidade de probabilidade tal que a probabilidade para que um evento aleatório apareça em uma vizinhança infinitesimal de um ponto Φ do EF é $g(\Phi)d\Phi$ onde $d\Phi$ é o elemento do volume. Os invariantes e os ângulos serão usados como variáveis, além de componentes dos momentos em alguns casos.

Definindo como eficiência a quantidade de trabalho gasta para obter-se um resultado de boa precisão estatística, todo programa de Monte Carlo deve apresentar uma certa eficiência dentro das múltiplas escolhas para a geração de eventos. Qualitativamente a eficiência é tão melhor quanto mais $g(\Phi)$ se aproxima de $f_n(\Phi)$ para ser integrada. Essa função $f_n(\Phi)$ é definida por

$$I_n = \int_V d\Phi \rho_n(\Phi) T(\Phi) = \int d\Phi f_n(\Phi) \quad (4.4)$$

portanto, $f_n(\Phi) = \rho_n(\Phi)T(\Phi)$ e $\rho_n(\Phi)$ é a densidade do Espaço de Fase. Note que, do ponto de vista da geração de eventos existem 3 tipos de densidades.

1. O gerador de eventos ideal, para o Monte Carlo é uma interação de partículas.
2. Qualquer modelo envolve a especificação do elemento de matriz ou a densidade de eventos no Espaço de Fase. Estes elementos de Matriz são muitas vezes complicados e não podem ser usados no computador.
3. Programas de computador do tipo Monte Carlo geram eventos com uma densidade $g(\Phi)$ as quais são algumas funções simples de Φ .

Vamos agora tentar explicar como pode-se gerar eventos e que tipos de densidade $g(\Phi)$ podem surgir na prática. Para isto tomamos a seguinte integral,

$$m = \int_0^1 dx f(x) \quad (4.5)$$

onde m , o valor da integral, é também valor médio da função $f(x)$. Assim, a variância é definida por

$$\sigma^2 = \int_0^1 dx [f(x) - m]^2 \quad (4.6)$$

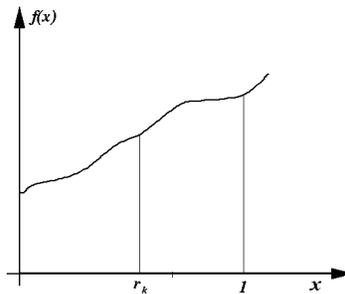


Figura 4.3: Exemplo de uma curva representando uma função qualquer sob integração Monte Carlo.

A variância mede as flutuações de $f(x)$ no intervalo $(0, 1)$. Quanto maior for σ^2 maior é a flutuação de $f(x)$ em torno do seu valor médio.

Suponha agora que estamos prontos para gerar os números aleatórios r_k com $0 \leq r_k \leq 1$, onde $k = 1, 2, \dots$, tal que os valores entre 0 e 1 sejam igualmente prováveis. Isto quer dizer que os r_k são uniformemente distribuídos no intervalo de 0 a 1. Vamos considerar os números r_1, \dots, r_n como uma amostra para o cálculo da integral (4.5). A aproximação de m é uma média aritmética dos N números resultantes de $f(r_k)$ e dada por

$$\bar{m} = \frac{1}{N} \sum_{k=1}^N f(r_k) \quad (4.7)$$

Esse é um Monte Carlo da integral (4.5). Usando a linguagem da Física de Partículas, r_k é chamado de evento e $f(r_k)$ é o peso associado ao evento. Como r_k é um número aleatório, então \bar{m} também é uma variável aleatória, tal que se considerarmos outro conjunto $\{r_k\}$ então obtemos outro valor para \bar{m} , daí então ser necessário calcular o desvio do valor médio.

O desvio de \bar{m} do seu valor mais provável é medido por,

$$\bar{\sigma}^2 = \frac{1}{N-1} \sum_{k=1}^N [f(r_k) - \bar{m}]^2 \quad (4.8)$$

o qual se aproxima da (4.7) assim como \bar{m} de m . Vamos supor que as variáveis aleatórias x_i $i = 1, \dots, N$ são distribuídas de tal forma que suas médias são m_i e as variâncias σ_i^2 .

Sob condições bem gerais, a distribuição da soma $x = (x_1 + \dots + x_N)/N$ para grande N se aproxima da distribuição normal,

$$N(x; m, \sigma) = \frac{1}{\sigma\sqrt{\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma}\right] \quad (4.9)$$

com média,

$$m = \frac{1}{N} \sum_{i=1}^N m_i \quad (4.10)$$

e variância,

$$\sigma^2 = \frac{1}{N^2} \sum_{i=1}^N \sigma_i^2 \quad (4.11)$$

No presente caso, os números aleatórios são os números $f_k = f(r_k)$ os quais são distribuídos com média m e variância σ^2 . Então, para N grande, \bar{m} converge para m e a variância da distribuição em \bar{m} é $(1/N)\sigma^2$. Assim, o resultado da integração por Monte Carlo é expresso por

$$m = \bar{m} \pm \frac{\bar{\sigma}}{\sqrt{N}} \quad (4.12)$$

4.2.4 Redução do Erro Estatístico

O erro estatístico do Método de Monte Carlo descrito depende de 2 fatores, da variância σ^2 da função a ser integrada e do número de pontos aleatórios para os quais a função é avaliada. O caminho mais simples para diminuir o erro é aumentar N . O erro é proporcional a $1/\sqrt{N}$ tal que decrescer o erro de 10 corresponde a um aumento de 100 para N . Esse caminho é limitado pelo tempo de computação que fica cada vez mais proibitivo com o aumento de N . Existem muitos outros caminhos para decrescer a variância. O problema geral pode-se chamar de "enriquecimento da amostra" de acordo com o problema tratado. Isto consiste em modificar ligeiramente a função $f_n(\Phi)$ por uma aproximação e/ou mudar a distribuição de eventos totalmente aleatórios. Vamos dar alguns exemplos. Pode-se melhorar a taxa de convergência tal que o erro estatístico venha a ser proporcional a $1/N$ usando números quasi-aleatórios.

Com números pseudo-aleatórios não temos idéia de quais pontos do integrando $f_n(\Phi)$ são avaliados. O FOWL usa os geradores pseudo-aleatórios. Um segundo caminho para

melhorar o Método de Monte Carlo consiste em reduzir a variância da função do integrando. De acordo com (4.8) vimos que é possível transformar a integral de tal forma que a função seja a mais constante possível. Vejamos um exemplo.

Consideremos a integral

$$m = \int_0^1 dx a e^{ax} = e^a - 1 \quad (4.13)$$

onde a é constante. Integrando desse tipo aparece em Física de Altas Energias quando temos dependências explícitas com o momento transferido do tipo e^{bt} . Um simples cálculo nos dá,

$$\sigma^2 = \frac{a}{2} \left(1 - \frac{2}{a} \right) e^{2a} + \frac{4}{a} e^a - \frac{2}{a} - 1 \quad (4.14)$$

Há um crescimento do erro relativo $\sigma/(m\sqrt{N})$ da integração por M.C. como $\sqrt{a/(2N)}$. Fazemos uma mudança de variável

$$y = e^{ax}$$

o que resulta em

$$m = \int_1^{e^a} dy$$

e podendo estão gerar pontos uniformemente em y no intervalo $(1, e^a)$. Como o integrando é uma constante a variância foi reduzida a zero.

Desse exemplo, é óbvio que se $f(x)$ varia consideravelmente, o método torna-se mais eficiente quando os pontos aleatórios são gerados de tal modo que sua densidade é próxima a $|f(x)|$. Este método é chamado de "importância da amostragem". Para aplicarmos este método é preciso gerar números aleatórios de acordo com a densidade $g(x)$. Esta densidade $g(x)$ é definida de tal forma que a probabilidade que um valor entre x e $x + dx$ seja obtido é dado por $(1/G)g(x)dx$ onde $G = G(+\infty)$ e

$$G(x) = \int_{-\infty}^x g(t) dt \quad (4.15)$$

Consideremos então a integral,

$$I = \int_{x_-}^{x_+} dx f(x) \quad (4.16)$$

onde os limites de integração serão introduzidos mais adiante. Tomemos como nova variável,

$$\begin{cases} r = \frac{G(x) - G(x_-)}{G(x_+) - G(x_-)} \\ dr = \frac{g(x)dx}{G(x_+) - G(x_-)} \end{cases} \quad (4.17)$$

variando entre 0 e 1. Então a equação (4.17) pode ser escrita como,

$$I = \int_{x_-}^{x_+} f(x) dx = \int_{x_-}^{x_+} g(x) \frac{f(x)}{g(x)} dx = [G(x_+) - G(x_-)] \int_0^1 \frac{f(x)}{g(r)} dr \quad (4.18)$$

onde

$$x = G^{-1} \{G(x_-) + r[G(x_+) - G(x_-)]\} \quad (4.19)$$

Quando aplicamos a equação (4.7)

$$\bar{m} = \frac{1}{N} \sum_{k=1}^N f(r_k)$$

na equação (4.17) obtemos para I :

$$I = \int_{x_-}^{x_+} dx f(x) = \bar{I} = \frac{[G(x_+) - G(x_-)]}{N} \sum_{k=1}^N \frac{f(x_k)}{g(x_k)} \equiv \frac{1}{N} \sum_{k=1}^N w_k \quad (4.20)$$

onde x_k é definido em termos de r_k pela equação (4.19) e w_k é o peso associado ao k -ésimo evento. Segundo (4.17) as variáveis aleatórias x_k são agora distribuídas de acordo com a densidade

$$\frac{dN(x)}{dx} = \left[\frac{dN(r)}{dr} \right] \left(\frac{dr}{dx} \right) = \frac{g(x)}{G(x_+) - G(x_-)} \quad (4.21)$$

a qual é normalisada a unidade no intervalo de integração.

Na integral sobre r , na equação (4.18), somente a razão f/g aparece e portanto a variância de f/g é reduzida se g se aproxima de f melhor que uma constante. Usando as equações (4.7), (4.8), (4.12) o resultado da integração por Monte Carlo de (4.20) é

$$I = \bar{I} \pm \delta \bar{I} \quad (4.22)$$

onde,

$$\bar{I} = \frac{1}{N} \sum_{k=1}^N w_k \quad (4.23)$$

$$\delta \bar{I}^2 = \frac{1}{N(N-1)} \left\{ \sum_{k=1}^N w_k^2 - \frac{1}{N} \left[\sum_{k=1}^N w_k \right]^2 \right\} \quad (4.24)$$

e

$$w_k = [G(x_+) - G(x_-)] \frac{f(x_k)}{g(x_k)} \quad (4.25)$$

Esse método para gerar números aleatórios nem sempre é aplicável, exige que todas as funções envolvidas $g(x)$, $y = G(x)$, $x = G^{-1}(y)$ sejam elementares. Um exemplo de aplicabilidade do método são as funções muito usadas em Física de Altas Energias, tais como a exponencial e a Breit-Wigner. Veja a tabela:

Tipo de variação	$g(x)$	$G(x)$	$G^{-1}(y)$
Exponencial	ae^{ax}	e^{ax}	$\frac{1}{a} \log y$
Breit-Wigner	$\frac{b^2}{(x-a)^2 + b^2}$	$\frac{\pi}{2} b \arctan [(x-a)/b]$	$a + b \tan [y - (\pi/2)b]/b$

No caso de uma Gaussiana $(-\infty, +\infty)$ ou de funções que não são integráveis de uma forma fechada o procedimento não se aplica.

4.2.5 Geração de Eventos

Nessa aplicação veremos como se pode fazer a geração aleatória dos eventos de uma interação entre partículas. O tratamento dos eventos propriamente ditos serão objeto de outro desenvolvimento.

Para calcular a integral (4.3) sobre o Espaço de Fase, devemos primeiramente eliminar as funções δ a 4 dimensões. E aí então escrever I_n explicitamente em termos do conjunto Φ de $3n - 4$ variáveis como na equação (4.4). Os valores de Φ definem o evento de forma única, tal que pode ser construída uma configuração de momentum p_1, \dots, p_n quando Φ é dada. Definimos um conjunto

$$\Phi' = (r^1, \dots, r^{3n-4})$$

tal que as $3n - 4$ variáveis formam o volume do hipercubo de integração V' ,

$$0 \leq r^l \leq 1, l = 1, \dots, 3n - 4$$

e a região física correspondente V .

Se definimos agora o Jacobiano de Φ para Φ' como $\frac{\partial \phi'}{\partial \phi}$ a integral (4.4) resulta,

$$I_N = \int_{V'} d\Phi' \frac{\rho_n[\phi(\phi')] T'(\phi')}{\partial \phi' / \partial \phi} = \frac{\partial \phi'}{\partial \phi} \int_{V'} d\Phi' \frac{T'(\phi')}{g'_n(\phi')} \quad (4.26)$$

com

$$g'_n(\phi') = \frac{\partial \phi' / \partial \phi}{\rho_n(\phi)} \quad (4.27)$$

Comparando com a equação (4.23) podemos escrever,

$$\left\{ \begin{array}{l} \bar{I}_n = \frac{1}{N} \sum_{k=1}^N w_k \\ w_k = \frac{T'(\phi'_k)}{g'_n(\phi'_k)} \end{array} \right. \quad (4.28)$$

para estimar o $\delta \bar{I}_n$ do erro estatístico usa-se a equação (4.24). Da equação (4.26) vemos que a unidade do evento é

$$\rho'_n(\phi') = \frac{1}{g'_n(\phi')} \quad (4.29)$$

As variáveis ϕ consistem de:

- $(n - 2)$ massas invariantes $M_i, M_i^2 = k_i^2 = (p_{1+}, \dots, +p_i)^2, i = 1, \dots, .$

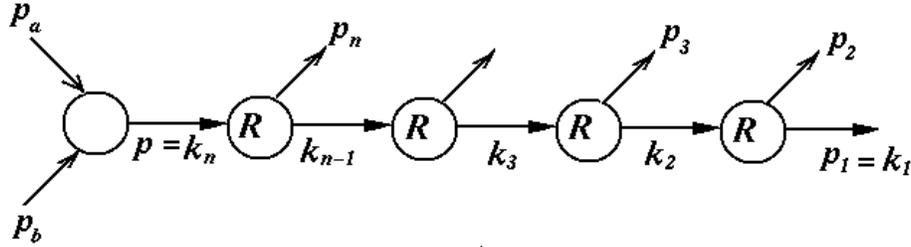


Figura 4.4: A reação $p_a + p_b \rightarrow p_1 + \dots + p_n$ expressa uma seqüência em cadeia de decaimentos em dois corpos.

- $(n - 1)$ massas dos sistemas intermediários, *i.e.* partículas que decaem como na figura 4.4.
- $2(n - 1)$ ângulos $\Omega_i = (\cos \theta_i, \phi_i)$, $i = 2, \dots, n$ definidos na figura 4.5, onde $k_{i+1} = p_1 + \dots + p_{i+1} = (M_{i+1}, \vec{0})$ *i.e.*, no sistema de repouso da partícula que se desintegra.

Com estas variáveis, a densidade do espaço de fase, fica,

$$\rho_n(\phi) = \frac{1}{2\sqrt{s}} \prod_{i=2}^n \frac{1}{2} P_i \quad (4.30)$$

Os ângulos θ_i , e ϕ podem ser gerados no correspondente sistema de repouso $k_{i+1} = (M_{i+1}, \vec{0})$ escrevendo-se,

$$\phi_i = 2\pi r^{(i)} \quad (4.31)$$

$$\cos \theta_i = (2\bar{r}^{(i)} - 1) \quad (4.32)$$

onde $r^{(i)}$ e $\bar{r}^{(i)}$ são números aleatórios no intervalo (0 e 1).

Vejam agora como gerar as massas invariantes. Sempre que a equação

$$\mu_i \leq M_i \leq M_{i+1} - m_{i+1} \quad (4.33)$$

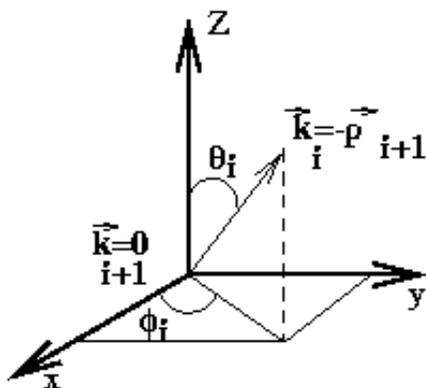


Figura 4.5: Sistema de referencia que define os ângulos.

onde, $i = 2, \dots, n - 1$, estiver satisfeita, o decaimento da figura 4.4 é físico.

No espaço $n - 2$ dimensões descrito por M_2, \dots, M_{n-1} , as desigualdades (4.33) formam um "simplex" que é, um intervalo para $n = 3$, um triângulo para $n = 4$, uma pirâmide para $n = 5$, etc. como mostra a figura 4.6.

Vejamos alguns exemplos de gerar M_i com um simplex.

(1) O método mais direto é o que transforma o simplex em um hipercubo de unidade $n - 2$ dimensões pela relação linear,

$$M_i = \mu_i + \bar{r}^{(i)}(M_{i+1} - \mu_{i+1}) \quad (4.34)$$

com $i = 2, \dots, n - 1$. Cada integração em M_i é transformada em uma integral em $\bar{r}^{(i)}$ de 0 a 1. Incluindo o Jacobiano vindo de (4.31), (4.32), (4.33) a integral por Monte Carlo I_n pode então ser escrita na forma (4.28) com

$$\frac{1}{g_n(\phi)} = \frac{1}{2\sqrt{s}} \prod_{i=2}^n 2\pi P_i \prod_{i=3}^n (M_i - \mu_i) \quad (4.35)$$

(2) Um outro caminho para gerar as massas M_i consiste em usar o fato de que o simplex (4.33) é parte de um hipercubo de dimensão $(n - 2)$,

$$\mu_i \leq M_i \leq \sqrt{s} - \mu_n + \mu_i \quad (4.36)$$

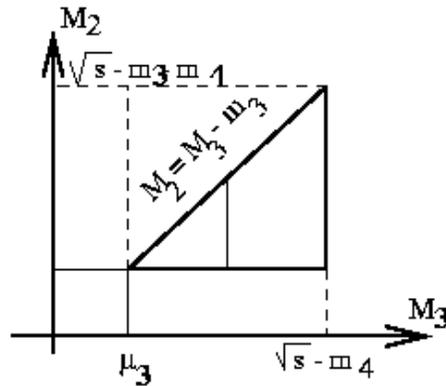


Figura 4.6: Regiões Físicas das Massas Invariantes.

com $i = 2, \dots, n - 1$. Se for imposto uma ordenação adicional temos,

$$M_i \leq M_{i+1} - m_{i+1} \quad (4.37)$$

então a região estudada é idêntica a (4.33). São assim geradas uniformemente no hiper-cubo como,

$$M_i = \mu_i + \bar{r}^{(i)}(\sqrt{s} - \mu_n) \quad (4.38)$$

tendo-se que são ordenadas inicialmente os $(n-2)$ números $\bar{r}^{(i)}$ como

$$\bar{r}^2 \leq \bar{r}^3 \leq \dots \leq \bar{r}^{n-1}.$$

São estas as duas formas de gerar eventos para os problemas apresentados acima. Maiores detalhes consulte outras referências indicadas no nosso livro texto indicado na referência [5].

Para n integrações numéricas temos vantagens e desvantagens no uso de Métodos de Monte Carlo, as quais podemos enumerar.

- Desvantagens:
 - Lenta convergência para poucas dimensões .
 - Erro $\propto 1/N$ enquanto que a
 - Regra do trapézio: erro $\propto 1/N^2$
 - Regra de Simpson: erro $\propto 1/N^4$

- Vantagens:
 - Rápida convergência em muitas dimensões .
 - Regiões de integração arbitrárias.
 - Uma estimativa é obtida com poucos pontos.
 - A precisão cresce com cada ponto adicionado.
 - Fácil estimativa de erro.

Em outro contexto analisaremos e descreveremos os processos estatísticos que são submetidos os dados de um experimento.

Capítulo 5

Limite Superior e Incerteza

1. Introdução
2. O Caso $n = 0$
3. Método da Integral
4. Método da Expansão
5. Subtração de *Background*

5.1 Introdução

Nesse capítulo [6] mostraremos o problema de como incorporar a incerteza na sensibilidade experimental dos cálculos de um limite sobre um *Branching Ratio* ou quantidade similar. Se o número de eventos é pequeno ou mesmo zero, mas sem *background*, a correção ao resultado usual é dada por uma fórmula simples e de fácil aplicação. A precisão dos resultados também será discutida.

Será considerado aqui o problema que se apresenta nos experimentos de decaimentos raros nos quais observam-se poucos ou nenhum evento. Como se pode incluir a incerteza sistemática da sensibilidade experimental no **Limite Superior** que é encontrado para o *branching ratio*? Problema similar se coloca nas medidas das secções de choque de produção de partículas raras, e assim por diante. Esses problemas foram desprezados até recentemente em todos os trabalhos publicados sobre medidas similares. Isso deve ter sido baseado na hipótese de que incertezas sistemáticas devem ser menos importantes que as flutuações de Poisson. Vamos começar analisando o caso no qual não se observa nenhum evento ($n = 0$) nas condições que desejamos para o problema.

5.2 O caso $n = 0$

Consideremos o caso de um experimento para determinar o valor de um *branching ratio* (BR) pela observação de n contagens em um detector. Seja R seu verdadeiro valor e S o valor do fator de sensibilidade, uma combinação de fluxo do feixe, aceitação do detector etc., definido de tal forma que a média do número de eventos esperado é

$$\mu = R S.$$

O número de eventos na verdade, observados em um experimento será uma amostra desenhada por uma distribuição de Poisson $P(m, \mu, S)$. Para propor uma inferência sobre o valor R_t exige-se mais informação.

Esta informação é obtida através de medidas sobre o valor de S_t subsidiárias que dão uma estimativa \hat{S} para S_t com uma incerteza σ .

Considerando que desejamos especificar um limite superior \bar{R}_0 para R_t no caso que o experimento não observa nenhum evento, devemos entender muito bem duas situações:

- (i) Um número significativo de eventos é observado ($n \gg 1$). O ponto estimado de R_t é então $\hat{R} = n/\hat{s}$. Usando uma aproximação padrão, um experimental atribuirá normalmente uma incerteza de $\hat{R}\sqrt{1/n + \sigma_t^2}$ como uma medida do intervalo central de confiança, onde $\sigma_t = \sigma/\hat{s}$ é a incerteza relativa em \hat{s} .
- (ii) Nenhum evento é observado ($n = 0$), mas não existe incerteza no valor de S_t ($\sigma = 0$). O nível de confiança de 90% para o limite superior sobre o parâmetro μ_t é $\ln 10 = 2.30$ e desde que o limite superior sobre R_t é $\bar{R}_{00} = 2.30/\hat{s}$. ($n = 0$ e $\sigma = 0$). O limite superior clássico é dado pelo valor de μ que faz $P(0; \mu) = 0.1$ i.e., o valor para o qual uma observação tão pequena ou menor que o presente é não mais que 10% provável.

O problema que estamos considerando é do tipo (ii), mas com uma variância $S_t \neq 0$. A probabilidade para se observar nenhum evento é dada por $p_0(R_t, S_t) = P(0; R_t, S_t)$. O valor R_t é desconhecido e é o que estamos tentando determinar ou pelo menos limitar. O valor de S_t é conhecido no sentido probabilístico de uma amostra de medidas subsidiárias. Supondo que o \hat{S} é uma estimativa não *biased*.

Uma generalização razoável da definição de um limite superior de 90% de nível de confiança (*confidence level* - CL) neste caso é explorar a função $p_0(R, S)$ e achar o valor \bar{R}_0 que faz $\langle p_0 \rangle = 0.1$, onde $\langle \rangle$ indica uma média sobre $w(S, \hat{S}, \sigma)$. Isto pode ser feito por Monte Carlo, mas veremos do ponto de vista analítico.

5.3 Método Integral

Comecemos com o caso de nenhum evento observado, $n = 0$,

$$p_0(R) = \langle p_0(R, S) \rangle = \int_0^x \exp(RS) w(S, \hat{S}, \sigma) ds \quad (5.1)$$

Nos casos mais complexos essa integral deve ser feita por Monte Carlo. Em alguns casos pode-se simplificar. É muito comum considerar w como uma Gaussiana, tal que,

$$p_0(R) = \int_0^x \exp(RS) \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(S - \hat{S})^2}{2\sigma^2}\right] ds \quad (5.2)$$

Essa forma para w exige $\sigma \ll \hat{S}$, e podemos estender o limite inferior da integral para $-\infty$ com uma mudança desprezível no resultado. (Se este não é o caso, a manutenção de uma forma Gaussiana para w é duvidosa, e uma distribuição log-normal pode nos fornecer um modelo melhor.). Desenvolvendo (5.2),

$$p_0(R) = \exp(-R\hat{S} + R^2\sigma^2/2) = \exp\left[-R\hat{S} + (RS\sigma_t)^2/2\right] \quad (5.3)$$

Achamos o limite superior de \bar{R}_0 de 90% de CL desenvolvendo a equação $p_0(\bar{R}_0) = 0.1$,

$$\bar{R}_0 = \left[1 - \sqrt{1 - 2\sigma^2/2 \ln 10}\right] / (\hat{S}\sigma_t^2) \quad (5.4)$$

Expandindo-se a raiz quadrada para pequeno σ_t temos,

$$\bar{R}_0 = 2.30 (1 + 2.30 \sigma^2/2) / \hat{S} \quad (5.5)$$

Como mostramos abaixo num segundo enfoque, essa equação é bastante geral e não se restringe a Gaussiana. Se existem eventos, $n > 0$, o enfoque é o mesmo mas a quantidade de computação aumenta neste enfoque. Para a Gaussiana w , a probabilidade de observar um evento é

$$p_t(R) = R(\hat{S} - R\sigma^2) \exp(-R\hat{S} + R^2\sigma_t^2/2) = (R\hat{S} - (R\hat{S}\sigma)^2) p_0(R). \quad (5.6)$$

Dado um evento, o limite superior é encontrado resolvendo-se numericamente para R_t na equação

$$p_0(\bar{R}_1) + p_1(\bar{R}_1) = 0.1$$

Podemos também expandir esta equação em σ_t , ou talvez fazer uma série de expansões no integrando.

5.4 Método da Expansão

Nesse segundo enfoque trabalharemos com os momentos de $w(S; \hat{S}\sigma)$ para derivar limites superiores sem integração explícita. Esse enfoque tem a vantagem de ser facilmente generalizável. Observemos inicialmente que devido a probabilidade de zero eventos $p_0(R, S) = e^{-RS}$ ser uma função côncava para cima, o valor da média sobre um intervalo de S é maior que o valor no centro do intervalo. Então R deve ser ligeiramente maior para dar o mesmo valor $\langle p_0 \rangle$ que o obtido sem a média.

Façamos a expansão de Taylor de $p_0(R, S)$ sobre o ponto que teríamos para o limite superior se não fosse a incerteza em S . Esse é o ponto $\bar{R} = \bar{R}_{00}$, $S = \hat{S}$ como no caso (ii) acima.

$$\begin{aligned} \Delta p_0 &= \left(\frac{\partial p}{\partial S} \right) \Delta S + \left(\frac{\partial p}{\partial R} \right) \Delta R \\ &+ \left(\frac{\partial^2 p}{\partial S^2} \right) (\Delta S)^2 + 2 \left(\frac{\partial^2 p}{\partial S \partial R} \right) \Delta S \Delta R + \left(\frac{\partial^2 p}{\partial R^2} \right) \frac{(\Delta R)^2}{2} \end{aligned} \quad (5.7)$$

Inserindo as derivadas de $p_0(R, S)$ obtemos,

$$\begin{aligned} \Delta p_0 &= \left[-\bar{R}_{00} \Delta S - \hat{S} \Delta R \right. \\ &+ \left. \bar{R}_{00}^2 \frac{(\Delta S)^2}{2} + (\bar{R}_{00} \hat{S} - 1) \Delta S \Delta R + \hat{S}^2 \frac{(\Delta R)^2}{2} \right] \exp(-\bar{R}_{00} \hat{S}) \end{aligned} \quad (5.8)$$

e

$$\langle \Delta p_0 \rangle = \left[-\hat{S} \Delta R + \frac{\bar{R}_{00}^2 \sigma^2}{2} \right] \exp(-\bar{R}_{00} \hat{S}) \quad (5.9)$$

onde usamos a definição $\sigma^2 = \langle \Delta S^2 \rangle$ e a hipótese de que \hat{S} é *unbiased*, da qual segue que $\langle \Delta S \rangle = 0$. Desprezamos também a segunda ordem em ΔR quando comparada com a primeira ordem na hipótese de que $\hat{S} \Delta R \ll 2$.

A probabilidade média é deslocada de

$$\langle p_0(\bar{R}_{00} + \Delta R, S) \rangle = p_0(\bar{R}_{00} \hat{S}) + \langle \Delta p_0 \rangle \quad (5.10)$$

Ambos os cálculos de p_0 nesta equação são iguais a 0.1 por definição porque o primeiro termo é a probabilidade média de zero eventos para o limite superior redefinido $\bar{R}_0 =$

$\bar{R}_{00} + \Delta R$ e o segundo tipo é a probabilidade no ponto da expansão da variância que é o caso (ii), isto é nenhum evento é observado $n = 0$. Então

$$\langle \Delta p_0 \rangle = 0 \quad \implies \quad \Delta R = \frac{\bar{R}_{00}^2 \sigma^2}{2\hat{S}}$$

No caso geral a probabilidade para a observação de n ou poucos eventos é:

$$p_n(R, S) = \sum_{j=0}^n \frac{(RS)^j}{j!} \exp(-RS).$$

Outra vez, desejamos expandir o ponto de limite superior (\bar{R}^{n_0}, \bar{S}) que adotaríamos no caso de variância zero. Tomando-se a primeira derivada de p_n notamos que todos os termos permanecem mas um se cancela,

$$\frac{\partial p_n}{\partial R} = -\frac{R^n S^{n+1}}{n!} \exp(-RS).$$

Com uma expressão simétrica a $\partial p_n / \partial R$. As derivadas que ficam na equação (5.7) são realmente avaliadas tal que a média sobre S nos dá um resultado semelhante ao da equação (5.9).

$$\begin{aligned} \langle \Delta p_n \rangle &= \left[-(\bar{R}^{n_0} \hat{S})^n \hat{S} \Delta R + (\bar{R}^{n_0} \hat{S})^{n-1} (\bar{R}^{n_0} \hat{S} - n) \frac{(\hat{S} \Delta R)^2}{2} \right. \\ &\quad \left. + (\bar{R}^{n_0} \hat{S})^{n+1} (\bar{R}^{n_0} \hat{S} - n) \frac{\sigma^2}{2} \right] \exp\left(\frac{-\bar{R}^{n_0} \hat{S}}{n!}\right) \end{aligned}$$

Dessa vez os termos de segunda ordem em ΔR foram mantidos. O valor da equação é zero pelo mesmo tipo de razão que o usado até o presente. Vamos simplificar a equação por:

$$U_n \equiv \bar{R}_n S$$

obtemos

$$\Delta U_n = U_{n_0} \frac{1}{E_n} [1 - (1 - \sigma_t^2 E_n^2)^{1/2}] \quad (5.11)$$

onde definimos $E_n = U_{n_0} - n$. E_n representa o excesso do limite superior de um parâmetro de Poisson sobre o número n de eventos observados para um especificado nível de confiança.

O valor de U é

$$U_n = U_{n_0} \left\{ 1 - \frac{1}{E_n} \left[1 - (1 - \sigma_t^2 E_n^2)^{1/2} \right] \right\} \quad (5.12)$$

Como na outra seção simplificamos a equação, expandindo a raiz quadrada e restando somente termos de segunda ordem em σ_t^2 .

$$U_n = U_{n_0} \left\{ 1 + E_n \frac{\sigma_t^2}{2} \left[1 + \left(\frac{E_n \sigma_t}{2} \right)^2 \right] \right\} \quad (5.13)$$

ou somente a primeira ordem:

$$U_n = U_{n_0} \left(1 + E_n \frac{\sigma_t^2}{2} \right) \quad (5.14)$$

Essa equação (5.14) é uma generalização da equação (5.5) e para n arbitrário se reduz a isso. Para 90% de CL os valores de U_{n_0} são familiares 2.30, 3.89, \dots para $n = 0, 1, \dots$ e $E_n = 2.30, 2.89, \dots$, etc. U_n é o valor modificado levando em conta a incerteza sistemática. Segundo a definição de U_n o limite superior sobre o *Branching Ratio* é $\bar{R}_n = U_n/S$. A raiz quadrada em (5.12) mostra que o resultado não pode ser correto para valores de σ_t os quais são também grandes e que o valor limite de σ_t de ficar menor quando n aumenta e exige que

$$\sigma_t \leq \frac{1}{E_n} \quad (5.15)$$

Claramente esse é um resultado da aproximação de Taylor. As avaliações de U_n são obtidas por Monte Carlo supondo uma incerteza sistemática do tipo Gaussiana. O resultado simples e intuitivo da equação (5.14) é adequado para muitos propósitos práticos. O caso não usual com $\sigma_t > 0.3$ provavelmente seria mostrado pela integração numérica e se uma observação dada $n > 3$ eventos normalmente daria uma estimativa do limite.

5.5 Subtração de Background

Se existe um *background* que deve ser subtraído, a determinação de um limite superior fica mais difícil [7]. No caso em que a taxa média do *background*, B , é conhecida com incerteza

desprezível, o Particle Data Group, recomenda a adoção do procedimento bayesiano sugerido por [7]. Em nossa notação a equação para a probabilidade de observar n ou poucos eventos é:

$$p_n = K \exp(-RS) \sum_{j=0}^n \frac{(B + RS)^j}{j!} \quad (5.16)$$

onde a normalização K depende de B e n através de

$$k^{-1} = \sum_{j=0}^n \frac{B^j}{j!} \quad (5.17)$$

Para incluir a incerteza sistemática, seria necessário um tratamento semelhante aquele dado na seção 3. Mas por simplicidade trataremos somente de termos lineares em ΔR , dado por,

$$\Delta U_n = \left[\frac{(U_{n_0} + B - n)}{(U_{n_0} + B)} \right] \frac{U_{n_0}^2 \sigma_t^2}{2} \quad (5.18)$$

Nestas instâncias U_{n_0} representa o limite superior que seria retirado das soluções gráficas fornecidas por esta análise. Para $B = 0$ este resultado concorda com a equação (5.11) para todo B , uma consequência do tratamento bayesiano da subtração do *background*. Seria útil introduzir uma incerteza no valor de B nesta análise, mas a dependência complexa das equações (5.16), (5.17) em B parece não permitir qualquer outro enfoque que o numérico sugerido acima.

Capítulo 6

O Potencial de Descoberta de Física Nova

1. Introdução
2. Uma Análise da Flutuação Estatística
3. Levando em Consideração os Erros Sistemáticos Relacionados ao Desconhecimento Parcial das Seções de Choque do Background e do Sinal
4. Estimativa dos Limites de Exclusão de Física Nova
5. A Probabilidade de Descoberta de Física Nova
6. Conclusão

6.1 Introdução

Nesse capítulo vamos descrever praticamente todo o artigo da referência [9] por ser um dos artigos mais claros sobre o assunto e conter a inclusão dos erros sistemáticos nas estimativas apresentadas, comparando com outros resultados existentes e mostrando diferenças essenciais. Decidimos incluir este capítulo por acharmos que esse é um assunto que todo físico deve se preocupar em vista da constante preparação de novos experimentos e novas condições experimentais.

As simulações em Física de Altas Energias têm usado frequentemente uma definição para a *significância* estatística de uma certa amostra que tem um número médio $\langle N_s \rangle$ de eventos de sinal e $\langle N_b \rangle$ de eventos de *background* ou simplesmente *background*.

O propósito desse estudo é mostrar como se deve levar em conta os erros sistemáticos relacionados a um desconhecimento do sinal e do *background* das seções de choque.

Se olharmos para os principais experimentos em preparação no Tevatron, no LHC e alguns outros poucos aceleradores de altas energias, observamos um objetivo comum dos experimentos em preparação que é a procura de “Física Nova” (daqui para frente referida sómente por FN), além do Modelo Padrão. A Supersimetria, os Higgs, etc., (toda física que está integral ou parcialmente fora do quadro estabelecido pelo Modelo Padrão) fazem parte dos assuntos a serem estudados.

Os exemplos usarão as condições de experimentação a serem usadas pelo CMS do LHC/CERN, para facilitar as aplicações nestas notas, o que não impede que sejam aplicadas as formulações aqui apresentadas a outros futuros experimentos.

A seção de choque do *background* é calculada para uma dada luminosidade média \mathcal{L} dos eventos de *background* como,

$$\langle N_b \rangle = \sigma_b \cdot \mathcal{L} \quad (6.1)$$

Vamos supor que a existência de FN nos dá uma seção de choque não nula σ_s com a mesma assinatura do *background* o que resulta na predição de um número médio adicional de eventos de sinal,

$$\langle N_s \rangle = \sigma_s \cdot \mathcal{L} \quad (6.2)$$

O número médio total dos eventos é dado por,

$$\langle N_{ev} \rangle = \langle N_s \rangle + \langle N_b \rangle = (\sigma_s + \sigma_b) \cdot \mathcal{L} \quad (6.3)$$

Em um experimento a probabilidade de produção de n eventos é dada pela distribuição de Poisson como foi visto na seção 2.2:¹

$$P(n, \langle n \rangle) = \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle} \quad (6.4)$$

¹Se fizermos $\langle n \rangle = \lambda$ e $n = r$ obteremos exatamente a expressão (2.8).

Como foi visto em (2.11) a *variância* para a distribuição de Poisson é dada por

$$\sigma^2 = \langle n \rangle \quad (6.5)$$

Para fazermos uma estimativa da probabilidade para descobrir FN temos que comparar com a estatística de Poisson as definições aqui descritas,

$$\langle n \rangle = \langle N_b \rangle \implies \text{quando } \underline{\text{n\~{a}o}} \text{ existe Física Nova} \quad (6.6)$$

$$\langle n \rangle = \langle N_b \rangle + \langle N_s \rangle \implies \text{quando existe Física Nova} \quad (6.7)$$

Em Física Experimental de Altas Energias, para estimar as possibilidades de descoberta de FN usa-se a definição de significância,

$$S_1 = \frac{N_s}{\sqrt{\langle N_b \rangle}} \quad (6.8)$$

$$S_2 = \frac{N_s}{\sqrt{\langle N_b \rangle + \langle N_s \rangle}} \quad (6.9)$$

Um valor convencional para confirmar uma descoberta de FN é $S_1(S_2) \geq 5$, ignorando-se os erros sistemáticos. Para $N_b \gg N_s$ $S_1 = S_2$ ² enquanto para $N_b \simeq N_s$, $S_1 \neq S_2$.

Mas estarão corretas estas definições de significância? Cabe então estabelecer a diferença entre um "experimento futuro" e um "experimento corrente ou real". No experimento real o número de eventos N_{ev} é um número dado, conhecido, sendo medido. As comparações se fazem com $\langle N_b \rangle$ para testar a validade da física corrente. Então o número de eventos de sinal é determinado por

$$N_s = N_{ev} - \langle N_b \rangle \quad (6.10)$$

e é comparado com o número de eventos do *background* $\langle N_b \rangle$.

A flutuação do *background* é $\sigma_{fb} = \sqrt{N_b}$ quando então podemos usar S_1 para apreciar a medida da distribuição da física corrente. Em um experimento futuro só conhecemos os números médios dos eventos de sinal e de *background*. Então fazemos uso da distribuição de Poisson $P(n, \langle n \rangle)$ e $P(n, \langle N_b \rangle + \langle N_s \rangle)$ para encontrar a probabilidade de achar FN no experimento futuro.

A definição da estimativa de FN é dada por $\langle N_b \rangle (\langle N_s \rangle) \gg 1$ e a significância seria,

$$S = \sqrt{\langle N_s \rangle + \langle N_b \rangle} - \sqrt{\langle N_b \rangle} \quad (6.11)$$

²Este é o caso por exemplo da assinatura $H \rightarrow \gamma \gamma$ dos bósons de Higgs.

6.2 Uma Análise da Flutuação Estatística

Como dissemos na sessão anterior, para se determinar a probabilidade de descobrir nova física em futuros experimentos, temos que comparar as distribuições de Poisson (6.4), (6.7), (6.7) tomando-se os valores

$$\begin{aligned} \langle n \rangle &= \langle N_b \rangle && \text{para Física Padrão} \\ \langle n \rangle &= \langle N_b \rangle + \langle N_s \rangle && \text{para Física Nova.} \end{aligned}$$

Vamos considerar primeiramente o caso $\langle N_b \rangle$ ($\langle N_s \rangle$) $\gg 1$. Nesse caso a distribuição de Poisson é aproximada por uma Gaussiana,

$$P_G(n, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(n - \mu)^2}{2\sigma^2}\right] \quad (6.12)$$

Com $\mu = \sigma^2$, $\mu = \langle N_b \rangle$ ou $\mu = \langle N_b \rangle + \langle N_s \rangle$ Utilizando a expressão (2.16) para $n = x$ obtemos (6.12).

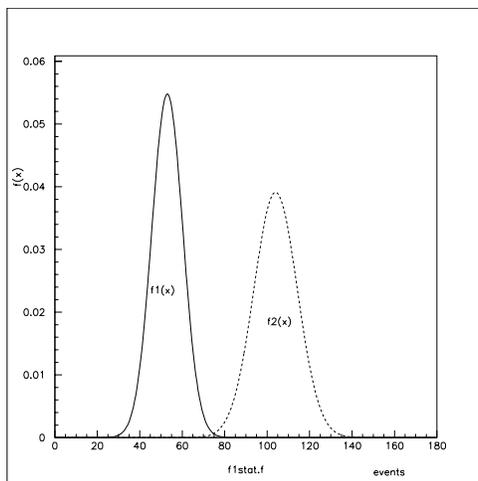


Figura 6.1: $f_1(x)$, $f_2(x)$ representam as densidades de probabilidades, dadas pela Gaussiana $P_G(x, \mu_{1,2}, \sigma^2)$ para $\mu_1 = \langle N_b \rangle = 53$ e $\mu_2 = \langle N_b \rangle + \langle N_s \rangle = 104$.

Dado o número médio de eventos $\langle n \rangle$ a serem observados no experimento futuro, a distribuição Gaussiana descreve a densidade de probabilidade para realizar n eventos. Na figura 6.1 as distribuições Gaussianas P_G foram obtidas dando dois valores para $\langle n \rangle$ como indicado no comentário da própria figura. Observe que existe uma área comum entre as duas curvas. Esta área representa a probabilidade de que a FN seja simulada

pela Física Padrão corrente (FP). Se supusermos como certa a existência de FN e dada pela probabilidade representada pela curva II ($f_2(x)$) cuja expressão é dada por

$$f_2(x) = P_G(x, \langle N_b \rangle + \langle N_s \rangle, \langle N_b \rangle + \langle N_s \rangle) \quad (6.13)$$

enquanto que a probabilidade κ para que a FP dada pela curva I ($f_1(x)$) cuja expressão é dada por

$$f_1(x) = P_G(x, \langle N_b \rangle, \langle N_b \rangle) \quad (6.14)$$

simule a FN é dada pela parte comum as duas curvas I e II.

Numéricamente, podemos calcular a expressão ,

$$\begin{aligned} \kappa &= \frac{1}{\sigma_2 \sqrt{2\pi}} \int_{-\infty}^{\sigma_1 \sigma_2} \exp \left[-\frac{1}{2} \frac{(x - \sigma_2)^2}{\sigma_2} \right] dx + \int_{\sigma_1 \sigma_2}^{+\infty} \exp \left[-\frac{1}{2} \frac{(x - \sigma_1)^2}{\sigma_1} \right] dx \\ &= \frac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\sigma_1 - \sigma_2} \exp \left[-\frac{y^2}{2} \right] dy + \int_{\sigma_2 - \sigma_1}^{+\infty} \exp \left[-\frac{y^2}{2} \right] dy \right\} \end{aligned}$$

e temos

$$\kappa = 1 - \operatorname{erf} \left(\frac{\sigma_2 - \sigma_1}{\sqrt{2}} \right) \quad (6.15)$$

onde $\sigma_1 = \sqrt{N_b}$ e $\sigma_2 = \sqrt{N_b + N_s}$.

Para encontrarmos o ponto de intersecção das duas curvas I e II, achamos a solução da igualdade

$$\frac{x_0 - \sigma_1^2}{\sigma_1} = \frac{x_0 - \sigma_2^2}{\sigma_2} \quad (6.16)$$

ou seja, $x_0 = \sigma_1 \sigma_2$.

Vamos agora discutir o significado da expressão (6.15). Para

$$x \leq x_0 = \sigma_1 \sigma_2 \implies f_1(x) \geq f_2(x) \quad (6.17)$$

temos que a densidade de probabilidade da FP é maior que a densidade de probabilidade de realização da FN. Isto é, para $x \leq x_0$ não temos nenhuma indicação em favor da FN.

A probabilidade para que o número de eventos seja $x < x_0$ é dada por

$$\alpha = \int_{-\infty}^{x_0} f_2(x) dx.$$

Para $x > x_0$ temos evidência para FN. Mas, nessa região $x > x_0$ as flutuações do background podem simular FN. Esse fato nos permite separar a probabilidade de que a FP simula FN em duas partes ($\kappa = \alpha + \beta$). Se $\kappa = 1$, FN nunca será encontrada no experimento. Por outro lado, se $\kappa = 0$ a primeira medida com probabilidade um tem que responder a questão da presença ou ausência de FN.

Isso significa que a área comum das funções da densidade de probabilidade do puro *background* e do *background* mais *signal* é a medida da probabilidade (não negativa, aditiva, e monotônica) do futuro experimento do potencial de não descoberta de FN.

De (6.15) segue que o papel desempenhado pela significância S é

$$S_{12} = \sigma_2 - \sigma_1 = \sqrt{N_b + N_s} - \sqrt{N_b} \quad (6.18)$$

Esta expressão tem sido usada para estimar valores sobre a probabilidade de descoberta dos bósons de Higgs (veja referências [9] e [10]).

S_{12} pode ser expressa em função das significâncias S_1 e S_2 como,

$$S_{12} = \frac{S_1 S_2}{S_1 + S_2} \quad (6.19)$$

No caso $N_b \gg N_s$, como na procura do bóson de Higgs no modo $H \rightarrow \gamma\gamma$, encontra-se $S_{12} \simeq 0.5$ e $S_1 \simeq 0.5S_2$. Para mostrar isto basta lembrar as expressões (6.9) e (6.9) e fazer um pequeno exercício de cálculo.

Isto quer dizer que $S_1 = 5$, de acordo com o usual, para uma convenção de 5σ de nível de confiança quer dizer possível a descoberta de FN. Se $S_{12} = 2.5$ que é a significância real, correspondendo a $\kappa = 1.24\%$.

A tabela damos alguns valores bem definidos de S_1 , S_2 , $\langle N_s \rangle$ e $\langle N_b \rangle$.

$\langle N_s \rangle$	$\langle N_b \rangle$	$\kappa(S_1 = \bar{5})$
5	1	0.1423
10	4	0.0828
15	9	0.0564
20	16	0.0448
25	25	0.0383
30	36	0.0333
35	49	0.0303
40	64	0.0278
45	81	0.0260
50	100	0.0245
55	121	0.0234
60	144	0.0224
65	169	0.0216
70	196	0.0209
75	225	0.0203
80	256	0.0198
85	289	0.0193
90	324	0.0189
95	361	0.0185
100	400	0.0182
150	900	0.0162
500	10^4	0.0136
5000	10^6	0.0125

Para o caso $N_s = k N_b$, $S_{12} = k_{12} S_2$ onde $k = 0.5, 1., 4., 10.$ o valor de k_{12} é dado por $k_{12} = 0.55, 0.56, 0.69, 0.77.$

Para valores de $\langle N_b \rangle$ e $\langle N_s \rangle + \langle N_b \rangle$ não muito grandes a comparação com a distribuição de Poisson se faz necessária. Lembremos que a distribuição de Poisson com a área de definição para inteiros não negativos podemos definir para todo real x ,

$$\bar{P}(x, \langle n \rangle) = \begin{cases} 0 & x \leq 0 \\ P(x, \langle n \rangle) & x > 0 \end{cases} \quad (6.20)$$

Com,

$$\int_{-\infty}^{+\infty} \bar{P}(x, \langle n \rangle) dx = 1. \quad (6.21)$$

A generalização da determinação anterior de κ nesse caso é direta pois κ é nada mais que a área comum das curvas descritas por $\bar{P}(x, \langle N_b \rangle)$ (curva I) e $\bar{P}(x, \langle N_b \rangle + \langle N_s \rangle)$ (curva II) da figura 6.2

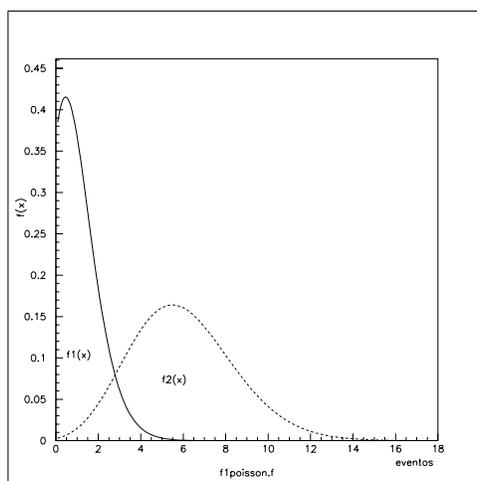


Figura 6.2: $f_1(x), f_2(x)$ representam as densidades de probabilidade, dadas pela distribuição de Poisson $f_{12}(x) \equiv \bar{P}(x, \mu_{12})$ para $\mu_1 = \langle N_b \rangle = 1$ e $\mu_2 = \langle N_b \rangle + \langle N_s \rangle = 6$.

Pode-se então mostrar que

$$\kappa = \alpha + \beta \quad (6.22)$$

$$\begin{aligned} \alpha &= \sum_{n=0}^{n_0} \frac{(\langle N_b \rangle + \langle N_s \rangle)^n}{n!} \exp(-(\langle N_b \rangle + \langle N_s \rangle)) \\ &= \frac{\Gamma(n_0 + 1, \langle N_b \rangle + \langle N_s \rangle)}{\Gamma(n_0 + 1)} \end{aligned} \quad (6.23)$$

$$\begin{aligned} \beta &= \sum_{n=n_0+1}^{\infty} \frac{(\langle N_b \rangle)^n}{n!} \exp(-\langle N_b \rangle) \\ &= 1 - \frac{\Gamma(n_0 + 1, \langle N_b \rangle)}{\Gamma(n_0 + 1)} \end{aligned} \quad (6.24)$$

$$n_0 = \frac{\langle N_s \rangle}{\ln(1 + \langle N_s \rangle / \langle N_b \rangle)} \quad (6.25)$$

6.3 Levando em Consideração os Erros Sistemáticos

Nas seções anteriores foram determinados os erros sistemáticos κ que é a probabilidade de que a FN seja descrita pela FP. Nesta seção vamos examinar a influência dos erros sistemáticos relacionados com a não desconhecida ou conhecida parcialmente, seção de choque do *background* e do sinal, sobre a probabilidade κ evitando confundir a nova física com a velha.

Denominamos de σ_b^0 e σ_s^0 o *background* e o sinal de “Born” respectivamente. Levando em consideração a correção de um “loop” temos

$$\begin{aligned}\sigma_b^0 &\implies \sigma_b^0(1 + \delta_{1b}) \\ \sigma_s^0 &\implies \sigma_s^0(1 + \delta_{1s})\end{aligned}\tag{6.26}$$

onde δ_{1b} e δ_{1s} é da ordem de $\mathcal{O}(0.5)$. Não se conhece a contribuição ou ordem de grandeza da correção de 2 “loops”. Supõe-se então que no presente as seções de choque que podem ser estimadas são

$$\begin{aligned}(\sigma_b^0, \sigma_b^0(1 + \delta_{1b})) \\ (\sigma_s^0, \sigma_s^0(1 + \delta_{1s}))\end{aligned}$$

E os números médios dos eventos de sinal e *background* são ,

$$\begin{aligned}[\langle N_b \rangle^0, \langle N_b \rangle^0(1 + 2\delta_{1b})] \\ [\langle N_s \rangle^0, \langle N_s \rangle^0(1 + 2\delta_{1s})]\end{aligned}\tag{6.27}$$

onde,

$$\begin{aligned}\langle N_b \rangle^0 &= \sigma_b^0 \cdot \mathcal{L} \\ \langle N_s \rangle^0 &= \sigma_s^0 \cdot \mathcal{L}\end{aligned}$$

Para descrever a probabilidade de que a FN seja descrita pela FP temos novamente que comparar com as distribuições de Poisson com e sem FN. Mas, diferentemente da seção 2 compararemos as distribuições de Poisson nas quais os números médios pertencem a um intervalo. Então , a priori só conhecemos os números médios do *background* e do sinal nos intervalos dados pelas expressões (6.27) e não os números exatos de $\langle N_b \rangle$ e $\langle N_s \rangle$.

No pior caso, para determinar a probabilidade da FN ser descrita pela FP, seria quando a FN é descrita por um número de eventos mínimo,

$$\langle N_b^{\min} \rangle = \langle N_s^0 \rangle + \langle N_b^0 \rangle\tag{6.28}$$

Supomos que para o pior caso, teríamos

$$\langle N_b^0 \rangle = \langle N_b^0 \rangle (1 + 2\delta_{1b}) \quad (6.29)$$

já que não conhecemos o valor exato de $\langle N_b^0 \rangle$. Então vamos comparar com a distribuição de Poisson, fazendo

$$\begin{aligned} \langle n \rangle &= \langle N_b^0 \rangle + \langle N_s^0 \rangle \\ &= \langle N_b^0 \rangle (1 + 2\delta_{1b}) + (\langle N_s^0 \rangle - 2\delta_{1b} \langle N_b^0 \rangle) \\ \langle n \rangle &= \langle N_b^0 \rangle (1 + 2\delta_{1b}) \end{aligned}$$

Usando o resultado das seções anteriores, encontramos que para o caso de $\langle N_b^0 \rangle (\langle N_s^0 \rangle) \gg 1$, a significância efetiva é dada por

$$S_{12s} = \sqrt{\langle N_b^0 \rangle + \langle N_s^0 \rangle} - \sqrt{\langle N_b^0 \rangle (1 + 2\delta_{1b})} \quad (6.30)$$

No caso limite $\delta_{1b} \rightarrow 0$ reproduzimos a expressão (6.18). Para valores não muito grandes de $\langle N_b^0 \rangle$ e $\langle N_s^0 \rangle$ usamos os resultados das seções anteriores e as tabelas dos resultados aqui dispostos.

Consideremos um exemplo particular com os seguintes parâmetros, $\delta_{1b} = 0.5$, $\langle N_s \rangle = 100$ e $\langle N_b \rangle = 50$ que é a situação típica da procura de sléptons.

Neste caso achamos que,

$$\begin{aligned} S_1 &= \frac{\langle N_s \rangle}{\sqrt{\langle N_b \rangle}} = 14.1 \\ S_2 &= \frac{\langle N_s \rangle}{\sqrt{\langle N_s \rangle + \langle N_b \rangle}} = 8.2 \\ S_{12} &= \sqrt{\langle N_b \rangle + \langle N_s \rangle} - \sqrt{\langle N_b \rangle} = 5.2 \\ S_{12s} &= \sqrt{\langle N_b \rangle + \langle N_s \rangle} - \sqrt{2 \langle N_b \rangle} = 2.25 \end{aligned}$$

Vemos que é grande a diferença entre o valor adotado pelo CMS $S_2 = 8.2$ que corresponde a $\kappa = 0.24 \times 10^{-15}$ e a significância $S_{12s} = 2.25$ levando em conta os erros sistemáticos devido a seção de choque do *background* desconhecido ou parcialmente desconhecido é 3.6. A comparação direta com a distribuição de Poisson se faz necessária.

$$\langle N_b \rangle (1 + 2\delta_{1b}) = 100$$

$$\langle N_b \rangle (1 + 2\delta_{1b}) + \langle N_{s,\text{eff}} \rangle$$

com

$$\langle N_{s,\text{eff}} \rangle = \langle N_s \rangle - 2\delta_{1b} \langle N_b \rangle = 50$$

o que dá $\kappa = 0.245$.

Tomemos outro exemplo com os seguintes valores médios, $\langle N_s \rangle = 28$, $\langle N_b \rangle = 8$ e $\delta_{1b} = 0.5$, e encontraremos os seguintes resultados para as significâncias,

$$\begin{cases} S_1 & = & 9.9 \\ S_2 & = & 4.7 \\ S_{12} & = & 3.2 \\ S_{12s} & = & 2.0 \end{cases} \quad \kappa = 0.045$$

Esses resultados mostram-nos que é muito importante considerar o erro sistemático relacionado ao desconhecimento das seções de choque do *background* pois essas análises dão um resultado menos otimista para o potencial de descoberta da supersimetria.

6.4 Limites de Exclusão de Física Nova

Suponhamos agora que conhecemos a seção de choque do *background* σ_b e que estamos interessados em obter um limite para a seção de choque do sinal, a qual depende de alguns parâmetros (massas das novas partículas, constantes de acoplamento, etc) e descrever alguma FN, indo além do modelo padrão.

Vamos considerar primeiramente o caso,

$$\langle N_b \rangle = \sigma_b \cdot \mathcal{L} \gg 1 \quad \text{e} \quad \langle N_s \rangle = \sigma_s \cdot \mathcal{L} \gg 1$$

Nesse caso a distribuição de Poisson se aproxima de uma Gaussiana. Como vimos na seção 2, a área comum das duas Gaussianas *com eventos de background e com eventos de sinal mais background*, é a probabilidade de que a FN possa ser descrita pela FP. Por exemplo, se exigirmos que a probabilidade da FP simular a FN seja de 10 % (o que nos dá pela expressão (6.18), $S_{12} > 1.64$). Isto quer dizer que,

$$\sqrt{\langle N_b \rangle + \langle N_s \rangle} - \sqrt{\langle N_b \rangle} \leq 1.64 \quad (6.31)$$

nos dando 90 % para o limite de exclusão sobre $\langle N_s \rangle$.

De um modo geral, quando exigimos que a probabilidade de que a FP simule FN seja $\pm\epsilon$, a expressão

$$\sqrt{\langle N_b \rangle + \langle N_s \rangle} - \sqrt{\langle N_b \rangle} \leq S(\epsilon) \quad (6.32)$$

nos permite obter o limite de exclusão $1 - \epsilon$ sobre a seção de choque de sinal σ_s . Isto é, supomos que $\epsilon = \kappa$. E encontramos,

$$\sigma_s \leq \frac{S(\epsilon)^2}{\mathcal{L}} + 2S(\epsilon) \sqrt{\frac{\sigma_b}{\mathcal{L}}} \quad (6.33)$$

Outros casos, como os de valores de $\langle N_b \rangle$ e $\langle N_s \rangle$ muito altos estamos obrigados a fazer a comparação com a distribuição de Poisson.

Mas os resultados (6.31) e (6.32) não levam em consideração os erros sistemáticos relacionados ao desconhecimento das seções de choque do *background*. Para fazermos isto, temos que usar as expressões ,

$$\sqrt{\langle N_b \rangle + \langle N_s \rangle} - \sqrt{\langle N_b \rangle (1 + 2\delta_{1b})} \leq S(\epsilon) \quad (6.34)$$

e

$$\sigma_s \leq \frac{S(\epsilon)^2}{\mathcal{L}} + 2S(\epsilon) \sqrt{\frac{\sigma_b(1 + 2\delta_{1b})}{\mathcal{L}}} + 2\delta_{1b} \sigma_b \quad (6.35)$$

onde δ_{1b} descreve as incertezas teóricas no cálculo da seção de choque do *background*.

Na tabela 8 encontramos alguns valores significativos como 90 % para o limite de exclusão sobre σ_s com $\mathcal{L} = 10^4 \text{ pb}^{-1}$, $2\delta_{1b} = 0.15$.

6.5 A Probabilidade de Descoberta de Física Nova

Foi mostrado na seção 2 a probabilidade κ de que a FP simule a FN. É importante agora determinar a probabilidade de que a FN venha a ser descoberta em futuros experimentos.

De acordo com as definições correntes de FN, essa corresponde a probabilidade de que o *background* simule o sinal desejado seja menor que 5σ ou 5.7×10^{-7} de probabilidade, desprezando é claro, nesse caso, todo tipo de erro sistemático.

Se exigirmos que a probabilidade da flutuação do *background* para $n > n(\delta)$, $\beta(\delta)$ é menor que δ ,

$$\beta(\delta) = \sum_{n=n_0(\delta)+1}^{\infty} P(\langle N_b \rangle, n) \leq \delta \quad (6.36)$$

A probabilidade $1 - \alpha(\delta)$ para que os eventos de sinal seja maior que $n_0(\delta)$ é igual a,

$$1 - \alpha(\delta) = \sum_{n=n_0(\delta)+1}^{\infty} P(\langle N_b \rangle + \langle N_s \rangle, n) \leq \delta \quad (6.37)$$

É importante ressaltar que δ é um número e $\alpha(\delta)$ é uma função de δ . Usualmente os físicos dizem que descobriu-se um certo fenômeno se a probabilidade de flutuação do *background* é menor que 5σ o que corresponde a $\delta_{dis} = 5.7 \times 10^{-7}$. Então pela equação (6.36) encontra-se $n_0(\delta)$ e estima-se a probabilidade $1 - \alpha(\delta)$ que um experimento terá que satisfazer como critério de descoberta.

Um exemplo seria dado pela procura do bóson de Higgs padrão com massa de $m_H = 110 \text{ GeV}$ no detetor CMS. Para uma luminosidade total de $\mathcal{L} = 3 \cdot 10^4 \text{ pb}^{-1}$ ($2 \times 10^4 \text{ pb}^{-1}$) encontramos $\langle N_b \rangle = 2893$ (1929), $\langle N_s \rangle = 357$ (238) e $S_1 = \langle N_s \rangle / \sqrt{\langle N_b \rangle}$.

Usando as expressões (6.36) e (6.37) para $\delta_{dis} = 5.7 \times 10^{-7}$ (critério de 5σ para descoberta) encontramos $1 - \alpha(\delta_{dis}) = 0.96$ (0.73). Isto quer dizer que o experimento CMS descobrirá os bósons de Higgs com um sinal $\geq 5\sigma$ com a massa de $m_H = 110 \text{ GeV}$ e com uma probabilidade de 96 % (73 %).

6.6 Conclusão

Com essas ferramentas podemos agora fazer vários exercícios com experiências passadas para saber a consistência do método e sobretudo com experiências futuras para avaliação das possibilidades do experimento.

Apêndice A

Tópicos sobre a Teoria de Probabilidades

A.1 Eventos equivalentes

Se duas variáveis aleatórias x e y estão relacionadas ¹ por $y = f(x)$ e Y é um conjunto de condições (eventos) sobre a variável y ² e $X = \{x \mid f(x) \in Y\}$, diz-se que os conjuntos X e Y são probabilisticamente independentes, ou que são eventos equivalentes, no sentido de que suas probabilidades de ocorrência são idênticas,

$$P(Y) = P(X)$$

A.1.1 pdf de eventos equivalentes

Se a **pdf**³ associada à variável x é $f(x)$ e $y = h(x)$ é uma outra variável aleatória, por exemplo, $y = 1 + x = h(x)$ e $f(x) = x/2$ para $0 < x < 2$, a **fd**⁴ $G(t)$ associada à y será

¹Ou, y é uma variável aleatória definida a partir de x .

²Por exemplo, uma relação do tipo $y < a$, onde a é uma constante.

³Função densidade de probabilidade

⁴Função de distribuição acumulada

dada por

$$\begin{aligned} G(t) &= P(y \leq t) = P(1 + x \leq t) = P(x \leq t - 1) \\ &= \int_0^{t-1} f(x) dx = \int_0^{t-1} \frac{x}{2} dx = \frac{(t-1)^2}{4} \end{aligned}$$

de modo que,

$$g(y) = \frac{dG}{dy} = \frac{y-1}{2}$$

é a **pdf** de y .

Uma vez que

$$G(t) = P(x \leq \underbrace{t-1}_{h^{-1}(t)}) = F[\underbrace{h^{-1}(t)}_{x(t)}]$$

implica

$$g(y) = \underbrace{\frac{dF}{dx}}_{f[x(y)]} \frac{dx(y)}{dx} = \frac{y-1}{2}$$

Esse resultado pode ser sistematizado como,

$$\boxed{g(y) = f[x(y)] \left| \frac{dx(y)}{dy} \right|}$$

se $y = h(x)$ for uma função monótona. Pois,

a) f crescente

$$G(t) = P(y \leq t) = P[f(x) \leq t] = P[(x \leq h^{-1}(t))] = F[h^{-1}(t)]$$

↓

$$\underbrace{\frac{dG}{dy}}_{g(y)} = \underbrace{\frac{dF}{dx}}_{f(x)} \frac{dx}{dy} \quad \left(\frac{dx}{dy} > 0 \right)$$

b) f decrescente

$$G(t) = P(y \leq t) = P[(x > h^{-1}(t))] = 1 - P[(x \leq h^{-1}(t))]$$

↓

$$g(y) = -f(x) \frac{dx}{dy} \quad \left(\frac{dx}{dy} < 0 \right)$$

A.1.2 Valor médio de função de uma variável aleatória

A princípio, para a determinação do valor esperado de uma função $y = h(x)$ de uma variável aleatória x de **pdf** $f(x)$, seria necessário primeiro, a determinação da **pdf** $g(y)$ de y e, então calculá-lo por

$$\langle y \rangle = \int_{-\infty}^{\infty} y g(y) dy$$

Entretanto, desde que

$$g(y) = f[x(y)] \frac{dx}{dy}$$

↓

$$\langle y \rangle = \int_{-\infty}^{\infty} y f[x(y)] \left| \frac{dx}{dy} \right| dy = \int_{-\infty}^{\infty} y f[x(y)] \frac{dx}{dy} dy$$

$$\langle h(x) \rangle = \int_{-\infty}^{\infty} h(x) f(x) dx$$

A.2 Variáveis aleatórias multidimensionais

Se a um determinado conjunto (sistema físico) estão associados alguns atributos (grandezas) $\{x_1, x_2 \dots x_n\}$, e as ocorrências de seus valores (medidas) são aleatórias, a n -úpla $X = (x_1, x_2 \dots x_n)$ é denominada **variável aleatória n-dimensional** ou **vetor aleatório n-dimensional**.

A **pdf** das ocorrências conjuntas de cada combinação de valores, associada a uma variável aleatória n -dimensional, é uma função de n variáveis $f(x_1, x_2 \dots x_n)$ tal que

$$\underbrace{\int \int \dots \int}_R f(x_1, x_2 \dots x_n) dx_1 dx_2 \dots dx_n = 1$$

onde R é o domínio de definição das variáveis.

A partir da **pdf** conjunta $f(x_1, x_2 \dots x_n)$, a **pdf** de qualquer uma das variáveis x_k , denominada **pdf marginal** de x_k , é definida por

$$\underbrace{\int \dots \int f(x_1 \dots x_n) dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_n}_{f(x_k)} = 1$$

uma vez que $\int f(x_k) dx_k = 1$

A.2.1 Variáveis independentes

Duas variáveis aleatórias x e y , associadas às **pdf**'s $\rho(x)$ e $\lambda(y)$ são ditas independentes se a **pdf** conjunta $f(x, y)$ for dada por

$$f(x, y) = \rho(x)\lambda(y)$$

ou seja, se a **pdf** conjunta for fatorável.

Se duas variáveis aleatórias x e y são independentes,

$$E(xy) = E(x) E(y)$$

pois,

$$\begin{aligned} E(xy) &= \int \int xy f(x, y) dx dy \\ &= \underbrace{\int x \rho(x) dx}_{E(x)} \underbrace{\int y \lambda(y) dy}_{E(y)} \end{aligned}$$

A.2.2 Funções de várias variáveis aleatórias

Se x e y são variáveis aleatórias com **pdf** conjunta $f(x, y)$ e $u = U(x, y)$, a **pdf** $\rho(u)$ pode ser determinada como uma **pdf** marginal, a partir de uma **pdf** conjunta $g(u, w)$ de u com outra função auxiliar arbitrária ⁵ $w = W(x, y)$, ou seja,

$$\rho(u) = \int_{-\infty}^{\infty} g(u, w) dw$$

Assim, se as funções U e W admitem inversas,

$$\begin{aligned} \int \int f(x, y) dx dy &= \\ \int \int f[x(u, w), y(u, w)] \left| \frac{\partial(x, y)}{\partial(u, w)} \right| du dw \end{aligned}$$

onde $\frac{\partial(x, y)}{\partial(u, w)} = \det \begin{pmatrix} \partial x / \partial u & \partial x / \partial w \\ \partial y / \partial u & \partial y / \partial w \end{pmatrix} = J(u, w)$

⁵Em geral, a mais simples possível.

é o chamado **jacobiano** da transformação

$$(x, y) \rightarrow (u, w)$$

Desse modo,

$$g(u, w) = f[x(u, w), y(u, w)] |J(u, w)|$$

e,

$$\rho(u) = \int_{-\infty}^{\infty} f[x(u, w), y(u, w)] |J(u, w)| dw$$

Exemplos típicos da determinação de **pdf** de funções de duas variáveis x e y com **pdf**s $\rho(x)$ e $\lambda(y)$.

$$1) \quad \begin{cases} u = xy \\ w = x \end{cases} \implies \begin{cases} x = w \\ y = u/w \end{cases}$$

Assim,

$$J = \begin{vmatrix} 0 & 1 \\ 1/w & -u/w^2 \end{vmatrix} = -\frac{1}{w}$$

e,

$$\begin{aligned} g(u, w) &= \rho(w) \lambda(u/w) |1/w| \\ g(u) &= \int_{-\infty}^{\infty} \rho(w) \lambda(u/w) |1/w| dw \\ &= \int_0^{\infty} \rho(w) \lambda(u/w) \frac{1}{w} dw - \int_{-\infty}^0 \rho(w) \lambda(u/w) \frac{1}{w} dw \end{aligned}$$

$$2) \quad \begin{cases} u = y/x \\ v = x \end{cases} \implies \begin{cases} x = v \\ y = uv \end{cases}$$

Assim,

$$J = \begin{vmatrix} 0 & 1 \\ v & u \end{vmatrix} = -v$$

e,

$$\begin{aligned} g(u, v) &= \rho(v) \lambda(uv) |v| \\ g(u) &= \int_0^{\infty} \rho(v) \lambda(uv) v dv - \int_{-\infty}^0 \rho(v) \lambda(uv) v dv \end{aligned}$$

$$3) \quad \begin{cases} u = x + y \\ v = x \end{cases} \implies \begin{cases} x = v \\ y = u - v \end{cases}$$

Assim,

$$J = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

e,

$$\begin{aligned} g(u, v) &= \rho(v) \lambda(u - v) \\ g(u) &= \int_{-\infty}^{\infty} \rho(v) \lambda(u - v) dv \end{aligned}$$

A integral acima é denominada integral de convolução .

Por outro método, se $u = x + y$

$$\begin{aligned} G(t) &= P(u \leq t) = P(x + y \leq t) \\ &= \underbrace{\int \int}_R \rho(x) \lambda(y) dx dy \end{aligned}$$

onde $R = \{(x, y) \mid x + y \leq t\}$

Assim,

$$G(t) = \int_{-\infty}^{\infty} dx \rho(x) \int_{-\infty}^{t-x} dy \lambda(y)$$

e,

$$G'(u) = g(u) = \int_{-\infty}^{\infty} \rho(x) \lambda(u - x) dx$$

A.3 Desigualdades e Teoremas

A.3.1 Desigualdade de Markov

Se x é uma variável aleatória não-negativa, com valor esperado $\mu = E(x)$ e ε é qualquer real positivo,

$$\boxed{P(x \geq \varepsilon) \leq \mu/\varepsilon = E(x)/\varepsilon}$$

Uma vez que

$$E(x) = \int_0^{\infty} x \rho(x) dx = \int_0^{\varepsilon} x \rho(x) dx + \int_{\varepsilon}^{\infty} x \rho(x) dx$$

$$\mu \geq \int_{\varepsilon}^{\infty} x \rho(x) dx \geq \varepsilon \int_{\varepsilon}^{\infty} \rho(x) dx = \varepsilon P(x \geq \varepsilon)$$

A.3.2 Desigualdade de Chebyshev

Para qualquer variável aleatória x , com média μ e variância σ^2 , e qualquer real ε ,

$$P(|x - \mu| \geq \varepsilon) \leq \sigma^2/\varepsilon^2$$

ou,

$$P(|x - \mu| < \varepsilon) \leq 1 - \sigma^2/\varepsilon^2$$

pois, substituindo-se na desigualdade de Markov x por $(x - \mu)^2$,

$$P[(x - \mu)^2 \geq \varepsilon^2] \leq \frac{E[(x - \mu)^2]}{\varepsilon^2} = \sigma^2/\varepsilon^2$$

segue-se a desigualdade, desde que $(x - \mu)^2 \geq \varepsilon^2$ é equivalente ⁶ a $|x - \mu| \geq \varepsilon$.

Essas desigualdades, obtidas a partir de pouquíssimas hipóteses sobre o comportamento de uma variável aleatória ⁷, permitem estabelecer limites para probabilidades associadas a determinados intervalos de definição da variável.

A.3.3 Lei dos grandes números

Uma aplicação imediata da desigualdade de Chebyshev é na prova da chamada **lei dos grandes números** ou do **limite da média das médias**, que estabelece:

A média \bar{x}_N de N variáveis aleatórias independentes x_1, x_2, \dots, x_N com mesma média e variância tende ao valor esperado μ de cada uma das variáveis.

⁶Definem eventos equivalentes.

⁷Apenas a existência da média e da variância.

$$\lim_{N \rightarrow \infty} P[|\bar{x}_N - \mu| < \varepsilon] = 0$$

onde $\bar{x}_N = \frac{1}{N} \sum_1^N x_i$ e $\varepsilon > 0$

Desde que $V(x) = \sigma^2 \implies V(\bar{x}_N) = \sigma^2/N$

segue-se, da desigualdade de Chebyshev,

$$P[|\bar{x}_N - \mu| < \varepsilon] \leq 1 - \frac{\sigma^2}{N\varepsilon^2}$$

ou seja,

$$\lim_{N \rightarrow \infty} \bar{x}_N = \mu$$

A.3.4 Função característica

A Transformada de Fourier (*TF*) de uma **pdf** $\rho(x)$,

$$TF[\rho(x)] = \rho(t) = \int_{-\infty}^{\infty} e^{ixt} \rho(x) dx$$

onde t é um número real, é chamada **função característica**, ou seja,

$$\rho(t) = E(e^{ixt})$$

Expandindo-se a exponencial e^{ixt} em série de Taylor,

$$e^{ixt} = 1 + ixt + \frac{(ixt)^2}{2} + \frac{(ixt)^3}{3!} + \frac{(ixt)^4}{4!} + \dots$$

implica

$$E(e^{ixt}) = \underbrace{E(1)}_1 + it \underbrace{E(x)}_{\langle x \rangle = \mu} + \underbrace{-\frac{t^2}{2} E(x^2)}_{\langle x^2 \rangle = \sigma_x^2 + \mu^2} - \frac{t^3}{3!} \underbrace{E(x^3)}_{\langle x^3 \rangle} + \dots$$

ou seja,

$$\rho(t) = 1 + it\mu - \frac{t^2}{2}(\sigma_x^2 + \mu^2) + \dots$$

A.3.5 Propriedades da função característica

Se x e y são variáveis independentes, $\rho_x(t)$ e $\rho_y(t)$ suas respectivas funções características e α e β são números reais, a função característica de $z = f(x, y)$ satisfará,

$$1) \text{ se } z = x + y, \quad \boxed{\rho_z(t) = \rho_x(t)\rho_y(t)}$$

uma vez que

$$\rho_z(t) = E(e^{izt}) = E(e^{ixt}e^{iyt}) = \underbrace{E(e^{ixt})}_{\rho_x(t)} \underbrace{E(e^{iyt})}_{\rho_y(t)}$$

$$2) \text{ se } y = \alpha x + \beta, \quad \boxed{\rho_y(t) = e^{i\beta t} \rho_x(\alpha t)}$$

pois,

$$\rho_y(t) = E(e^{iyt}) = E[e^{i(\alpha x + \beta)t}] = e^{i\beta t} \underbrace{E(e^{i\alpha x t})}_{\rho_x(\alpha t)}$$

3) se a **pdf** de x é gaussiana,

$$\rho_G(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma_x}\right)^2}$$

e,

$$\rho_G(t) = \int_{-\infty}^{\infty} e^{ixt} \rho_G(x) dx$$

$$\text{Fazendo-se } z = \frac{x - \mu}{\sigma_x} \implies dz = \frac{dx}{\sigma_x}$$

Assim,

$$\rho_G(t) = \frac{e^{i\mu t}}{2\pi} \int_{-\infty}^{\infty} e^{i\sigma_x z t - z^2/2} dz = \frac{e^{i\mu t - \sigma_x^2 t^2/2}}{2\pi} \underbrace{\int_{-\infty}^{\infty} e^{-\frac{(z-i\sigma_x t)^2}{2}} dz}_{\int_{-\infty}^{\infty} e^{-\frac{w^2}{2}} dw = \sqrt{2\pi}}$$

ou seja,

$$\boxed{\rho_G(t) = e^{i\mu t - \sigma_x^2 t^2/2}}$$

\Updownarrow

$$\log \rho_G(t) = i\mu t - \sigma_x^2 t^2/2$$

Para uma distribuição normal padrão, $\mathcal{N}_t(0, 1)$,

$$\boxed{\phi(t) = e^{-t^2/2}}$$

A.3.6 Teorema do limite central

Se $\{x_i\}$ é um conjunto de N variáveis aleatórias independentes associadas à uma mesma pdf $\rho(x)$, não necessariamente normal, com média e variância dados por

$$\begin{cases} E(x_i) = \mu \\ V(x_i) = \sigma^2 \end{cases}$$

A média $\bar{x} = \sum_{i=1}^N x_i/N$ é também uma variável aleatória de mesmo valor esperado $E(\bar{x}) = \mu$ e variância $V(\bar{x}) = \sigma^2/N$.

Nessas circunstâncias, a variável reduzida $\bar{z}_N = \frac{\bar{x} - \mu}{\sigma/\sqrt{N}}$ tende à distribuição normal padrão $\mathcal{N}_{\bar{z}_N}(0, 1)$, no sentido de que

$$\boxed{\lim_{N \rightarrow \infty} P(\bar{z}_N \leq t) = \phi(t)}$$

Desde que $\rho_{\bar{x}}(t) = [\rho(t/N)]^N$,

$$\rho_{\bar{z}_N}(t) = e^{-\frac{i\mu t}{\sigma\sqrt{N}}} \rho_{\bar{x}}(\sqrt{N}t/\sigma) = e^{-\frac{i\mu t}{\sigma\sqrt{N}}} \left[\rho\left(t/\sqrt{N}\sigma\right) \right]^N$$

Tomando-se o logaritmo,

$$\log \rho_{\bar{z}_N}(t) = -\frac{i\mu t}{\sigma\sqrt{N}} + N \log \rho\left(t/\sqrt{N}\sigma\right)$$

e, levando-se em conta que,

$$\begin{aligned} \rho\left(t/\sqrt{N}\sigma\right) &= 1 + \frac{i\mu t}{\sqrt{N}\sigma} - \frac{(\sigma^2 + \mu^2)}{2N\sigma^2}t^2 + \dots \\ \log \rho_{\bar{z}_N}(t) &= -\frac{i\mu t}{\sigma\sqrt{N}} + N \log \left[1 + \frac{i\mu t}{\sqrt{N}\sigma} - \frac{(\sigma^2 + \mu^2)}{2N\sigma^2}t^2 + \dots \right] \\ &= -\frac{i\mu t}{\sigma\sqrt{N}} + N \left[\frac{i\mu t}{\sqrt{N}\sigma} - \frac{(\sigma^2 + \mu^2)}{2N\sigma^2}t^2 + \dots \right] + \\ &\quad - \frac{1}{2} \left(\frac{-\mu^2}{N\sigma^2}t^2 - \frac{i\mu t^3}{\sqrt{N}} \frac{(\sigma^2 + \mu^2)}{N} + \frac{(\sigma^2 + \mu^2)}{4N^2\sigma^4}t^4 + \dots \right) \end{aligned}$$

Assim,

$$\lim_{N \rightarrow \infty} \log \rho_{\bar{z}_N}(t) = -\frac{t^2}{2}$$

↓

$$\mathcal{N}_{\bar{z}_N}(0, 1)$$

A.4 Distribuições

A.4.1 Distribuição gama

Se z é uma variável aleatória distribuída normalmente, segundo $\mathcal{N}(0, 1)$ e,

$$u(z) = \frac{1}{2} z^2$$

a **pdf** associada a u é dada por

$$g(u) = \frac{1}{\sqrt{\pi}} u^{-\frac{1}{2}} e^{-u}$$

Pois, uma vez que a função $u(z)$ não é monótona, sua **pdf** deve ser determinada por

$$\begin{aligned} G(t) &= P(u \leq t) = P(1/2z^2 \leq t) \\ &= P(z^2 \leq 2t) = P(-\sqrt{2t} \leq z \leq \sqrt{2t}) \\ &= F(\sqrt{2t}) - F(-\sqrt{2t}) = \int_{-\sqrt{2t}}^{\sqrt{2t}} \rho_G(z) dz \end{aligned}$$

↓

$$\frac{dG}{du} = g(u) = \frac{1}{\sqrt{2\pi}} 2(2u)^{-1/2} e^{-u}$$

Como $\Gamma(1/2) = \sqrt{\pi}$, $g(u)$ pode ser escrita

$$g(u) = \frac{u^{-1/2}}{\Gamma(1/2)} e^{-u}$$

Essa distribuição é um caso particular da chamada **distribuição gama**,

$$\gamma(u|\alpha) = \frac{u^{\alpha-1}}{\Gamma(\alpha)} e^{-u}$$

onde α é o parâmetro da distribuição denominado **número de graus de liberdade**.

A função característica de uma distribuição gama é dada por

$$\begin{aligned} \gamma(t) &= \frac{1}{\Gamma(\alpha)} \int_{-\infty}^{\infty} e^{-\overbrace{u(1-it)}^v} u^{\alpha-1} du \\ &= \frac{(1-it)^{-\alpha}}{\Gamma(\alpha)} \underbrace{\int_{-\infty}^{\infty} e^{-v} v^{\alpha-1} dv}_{\Gamma(\alpha)} \end{aligned}$$

ou seja,

$$\gamma(t) = (1-it)^{-\alpha}$$

A.4.2 Distribuição de χ^2

Se $\{x_i\}$ é um conjunto de n variáveis independentes distribuídas normalmente com média μ_i e desvio padrão σ_i , a variável

$$z_i = \frac{x_i - \mu_i}{\sigma_i}$$

distribuir-se-á normalmente segundo $\mathcal{N}(0, 1)$ e,

$$\xi = \sum_{i=1}^n \frac{z_i^2}{2} = \sum_{i=1}^n u_i$$

estará associada à função característica

$$g_{\xi}(t) = \prod_{i=1}^n g_{u_i}(t) = g_{u_i}^n(t) = (1-it)^{-n\alpha} = (1-it)^{-n/2} \quad (\alpha = 1/2)$$

cuja **pdf** é dada por

$$g(\xi|n) = \frac{\xi^{n/2-1} e^{-\xi}}{\Gamma(n/2)} = \gamma(\xi|n/2)$$

Assim, a variável $\chi^2 = 2\xi$ ($\xi \geq 0$) terá **pdf** dada por

$$\rho(\chi^2) = g[\xi(\chi^2)] \frac{d\xi}{d\chi^2}$$

ou seja,

$$\rho(\chi^2|n) = \frac{(\chi^2/2)^{n/2-1} e^{-\chi^2/2}}{2\Gamma(n/2)}$$

Essa **pdf** é denominada **distribuição de χ^2** com n graus de liberdade.

A.4.3 Distribuição de Student

Se x, x_1, x_2, \dots, x_n são $(n + 1)$ variáveis independentes e distribuídas normalmente segundo $\mathcal{N}(0, \sigma)$, a variável

$$t = \frac{x}{\sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}}$$

obedecerá a chamada **distribuição de Student**.

Pois, uma vez que a variável

$$\lambda = \sum_{i=1}^n \left(\frac{z_i^2}{2} \right)$$

onde $z_i = x_i/\sigma$, possui **pdf** igual a

$$\rho(\lambda) = \gamma(\lambda|n/2) = \frac{\lambda^{n/2} e^{-\lambda}}{\Gamma(n/2)}$$

a variável

$$u = \sum_{i=1}^n x_i^2 = 2\sigma^2 \sum_{i=1}^n \frac{1}{2} \left(\frac{x_i^2}{\sigma} \right)^2 = 2\sigma^2 \sum_{i=1}^n \frac{z_i^2}{2} = 2\sigma^2 \lambda$$

terá **pdf**

$$\begin{aligned}\rho(u) &= \frac{d\lambda(u)}{du} \gamma[\lambda(u)|n/2] \\ &= \frac{1}{2\sigma} \gamma\left(\frac{u}{2\sigma^2} \middle| \frac{n}{2}\right) = \frac{1}{2\sigma} \left(\frac{u}{2\sigma^2}\right)^{n/2-1} \frac{e^{-u/(2\sigma^2)}}{\Gamma(n/2)}\end{aligned}$$

Além disso, a variável

$$\eta = \left(\frac{\sum_{i=1}^n x_i^2}{n}\right)^{1/2} = \frac{1}{n^{1/2}}(2\sigma^2)^{1/2} \left[\sum_{i=1}^n \frac{1}{2} \left(\frac{x_i}{\sigma}\right)^2\right]^{1/2}$$

ou seja,

$$\eta = \left(\frac{2\sigma^2}{n}\right)^{1/2} \lambda^{1/2} \quad (\eta > 0)$$

terá **pdf** dada por

$$\rho(\eta) = \rho[\lambda(\eta)] \frac{d\lambda}{d\eta}$$

ou seja,

$$\begin{aligned}\rho(\eta) &= \frac{n\eta}{\sigma^2} \gamma\left(\frac{n}{2} \frac{\eta^2}{\sigma^2} \middle| \frac{n}{2}\right) = \frac{n\eta}{\sigma^2} \frac{\left(\frac{n}{2} \frac{\eta^2}{\sigma^2}\right)^{n/2-1}}{\Gamma(n/2)} e^{-\frac{n}{2} \frac{\eta^2}{\sigma^2}} \\ &= \frac{2(n/2)^{n/2} \eta^{n-1} e^{-\frac{n}{2} \frac{\eta^2}{\sigma^2}}}{\sigma^n \Gamma(n/2)}\end{aligned}$$

Assim, a **pdf** conjunta de $t = x/\eta$ e η será dada por

$$\rho(x/\eta, \eta) = \rho[x(\eta)] \rho(\eta) |\eta|$$

ou seja,

$$\rho(x/\eta, \eta) = \frac{\sqrt{\frac{2}{\pi}} \left(\frac{n}{2}\right)^{n/2}}{\sigma^{n+1} \Gamma(n/2)} \eta^{n-1} e^{-\frac{(x^2 + n\eta^2)}{2\sigma^2}} |\eta|$$

Fazendo-se

$$\frac{\sqrt{\frac{2}{\pi}} \left(\frac{n}{2}\right)^{n/2}}{\sigma^{n+1} \Gamma(n/2)} = c_n$$

e

$$\frac{(x^2 + n\eta^2)}{2\sigma^2} = \frac{\eta^2}{2\sigma^2} [(x/\eta)^2 + n] = \eta^2 \frac{(t^2 + n)}{2\sigma^2} = a_n$$

A pdf de t será dada por

$$\rho(t) = c_n \int_0^\infty \eta^n e^{-a_n \eta^2} d\eta$$

Fazendo-se

$$\begin{cases} a_n \eta^2 = v \\ 2a_n \eta d\eta = dv \end{cases} \implies \begin{cases} \eta^n = (v/a_n)^{n/2} \\ d\eta = \frac{dv}{2a_n^{1/2} v^{1/2}} \end{cases}$$

$$\rho(t) = \frac{c_n}{2} a_n^{\left(\frac{n+1}{2}\right)} \underbrace{\int_0^\infty v^{\frac{n-1}{2}} e^{-v} dv}_{\Gamma\left(\frac{n+1}{2}\right)}$$

Desde que

$$\begin{aligned} a_n^{\left(\frac{n+1}{2}\right)} &= \left(\frac{t^2 + n}{2\sigma^2}\right)^{-\frac{n+1}{2}} \\ &= 2^{\left(\frac{n+1}{2}\right)} \sigma^{n+1} n^{-\left(\frac{n+1}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} \end{aligned}$$

e

$$\frac{c_n}{2} = \frac{2^{-1/2} n^{n/2} 2^{-n/2}}{\sqrt{\pi} \sigma^{n+1} \Gamma(n/2)} = \frac{2^{-\left(\frac{n+1}{2}\right)} n^{n/2}}{\sqrt{\pi} \sigma^{n+1} \Gamma(n/2)}$$

a chamada **distribuição de Student** com n graus de liberdade é dada por

$$\rho(t|n) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + t^2/n\right)^{-\left(\frac{n+1}{2}\right)}$$

Apêndice B

Ferramentas para Simulação e Análise de Dados

O desenvolvimento de pacotes de programação (*softwares*) para a simulação e análise de dados é uma prática corrente na Física Experimental de Altas Energias. Estes desenvolvimentos têm sido transformados em subprodutos abertos aos pesquisadores, não só das Altas Energias mas, de qualquer área.

Paralelamente, existem inúmeros pacotes comerciais de programação que são atualizados periodicamente mas, a maioria desses pacotes, por terem sido criados para pequenos computadores pessoais, ainda possuem as limitações e restrições impostas pelos primeiros computadores pessoais, como limites estreitos de armazenamento de dados e gerenciamento de memória.

Desse modo, os pesquisadores da área Física de Altas Energias desenvolveram e ainda desenvolvem pacotes gratuitos mais robustos e estáveis, tanto do ponto de vista da concepção quanto da confiabilidade e longevidade (versões).

Ao contrário dos comerciais, os pacotes desenvolvidos na Física de Altas Energias, desde o início, foram criados em sistemas múltiplos, mais complexos e robustos, ou seja, para várias plataformas e sistemas operacionais de computação.

Um exemplo é o pacote gráfico para análise de dados, denominado **PAW** (*Physics Analysis Workstation*), desenvolvido por uma equipe de pesquisadores do CERN ¹, em 1988, e que vem sendo largamente utilizado por toda a comunidade da Física de Altas Energias.

¹ *European Laboratory for Particle Physics, Geneva, Switzerland*

Por ter sido desenvolvido em cima da linguagem de programação científica FORTRAN, que é a linguagem que tem sido utilizada praticamente desde a criação dos computadores comerciais (1950), o **PAW** teve que suprir e implementar várias características de programação e técnicas de uso de memória para o armazenamento de dados ausentes no FORTRAN.

Mais recentemente, em 19XX, outra equipe do CERN criou e desenvolveu um outro pacote gráfico de análise, denominado **ROOT**, já baseado numa linguagem de programação de concepção distinta do FORTRAN, o **C++**.

Num estágio inicial de formação, é conveniente o aprendizado de ambas as ferramentas.

B.1 PAW

O PAW apoia-se na definição de três elementos básicos: **vetores**, **histogramas** e **n-úplas**.

B.1.1 Vetores

Vetores são arranjos numéricos, equivalentes ao **array** do FORTRAN ou, de maneira mais matemática, que podem ser associados, num dado espaço de dimensão N , a variáveis tensoriais de ordem 1, 2 ou 3. São identificados na memória do programa por um nome.

Exemplo:

```
PAW > v/cr x(5) r 1 4 9 16 25
```

```
PAW > v/pr x
```

Esses são os comandos mais simples para criar (**vector/create**) um vetor e mostrar (**vector/print**) seus valores na tela. Ou seja, o primeiro comando (**v/cr**)² cria um vetor real (**r**) de dimensão 5, denominado **x**, cujas componentes possuem os valores 1, 4, 9, 16 e 25. Enquanto o segundo comando (**v/pr**) apenas mostra, indexadamente, as componentes de **x**.

²Uma característica do PAW é que os comandos podem ser abreviados quando não houver ambiguidades. Assim, por exemplo, o comando **vector/create** pode ser escrito, simplesmente, como **v/cr**. Analogamente, **vector/print** pode ser escrito como **v/pr**.

É claro que esses comandos são utilizados apenas para vetores de baixa dimensão. Em geral, os atributos de um objeto, que podem ser representados numericamente e lidos como vetores, são gerados e armazenados em arquivos no disco do computador. O exemplo seguinte mostra como ler um arquivo ASCII (não formatado) de dados (`vector/read`), criado externamente ³, associá-lo a dois vetores (`x` e `y`) e mostrá-los na tela (`vector/write`) de acordo com a sintaxe FORTRAN.

Exemplo:

```
PAW > v/re x,y vec.dat
```

```
PAW > v/wr x,y ! 5(F6.1,2x,F6.2)
```

Outras operações com vetores são ilustradas pelo seguinte arquivos de comandos ou KUMAC,

```
* Comentarios dentro de uma KUMAC podem ser feitos a partir do simbolo *
* Alertas e instrucoes durante a execucao de uma KUMAC podem ser dados
* pelo comando MESSAGE, como abaixo
```

```
mess '***** Atencao leia as instrucoes dadas na KUMAC *****'
mess ' '
mess '*** Esta KUMAC cria dois vetores, com o pacote SIGMA, que'
mess '*** representam as equacoes parametricas de um circulo.'
```

```
**** Criar 2 vetores cujas componentes sao 1000 valores das
**** funcoes seno e cosseno de 0 a 2*pi
sigma x=sin(array(1000,0#2*pi))
sigma y=cos(array(1000,0#2*pi))
```

```
**** comando para plotar 4 graficos na mesma tela
zon 2 2
```

³Crie um arquivo ASCII denominado `vec.dat` contendo números dispostos como:

1	2
2	4
3	9
4	16
5	25

```

v/dr x
v/dr y
v/dr x
v/dr y ! s
graph 1000 x y

**** comando para dar uma pausa na execucao da KUMAC
wait

**** restitui o estado anterior
zon
graph 1000 x y

**** habilita dados estatisticos no grafico
opt stat
opt fit

sigma d=(array(10000,2#2))
sigma z=sqrt(-2*log(rndm(array(10000,1))))*3*cos(2*pi*rndm(array(10000,2)))
vadd d z w
mess '*** Os vetores "d", "z" e "w" forma criados. '
mess '*** Execute os comandos "v/pl w" e "h/fit 12345 g" '

```

B.1.2 Histogramas

Histogramas são conjuntos de dados numéricos (distribuições) agrupados em intervalos (bins) que contêm informações adicionais sobre; o conteúdo (nome da variável) dos dados, o título do histograma, a frequência associada a cada intervalo, outros parâmetros estatísticos como erros, médias e desvios ou uma função de ajuste associada a distribuição. Desse modo, usualmente, representam a distribuição de um atributo associado a algum objeto. São identificados na memória do programa por um arquivo (unidade) do tipo FORTRAN, ou seja, por um número.

B.1.3 N-úplas

N-úplas são conjuntos de dados mais complexos que permitem a representação dos vários atributos (numéricos ou não) de um objeto (evento). Ou seja, podem ser encaradas como um conjunto de histogramas concatenados que são também identificados na memória do programa por um arquivo (unidade) do tipo FORTRAN, ou seja, por um número.

A interação entre esses elementos básicos do PAW pode ser esquematizada na figura, que mostra alguns comandos que permitem operações de conversão e leitura (em ambos os sentidos, input-output) desses elementos básicos.

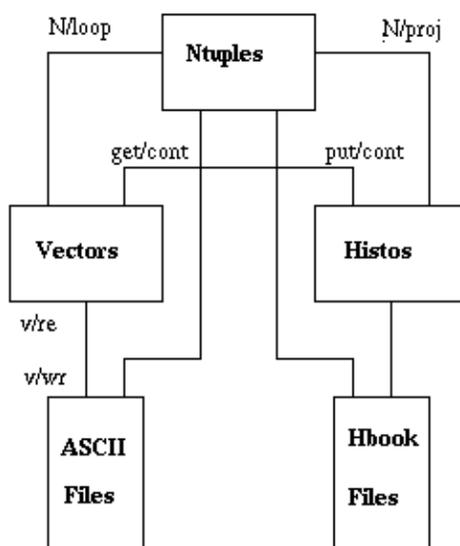


Figura B.1: Elementos básicos do PAW.

Além desses elementos básicos, o PAW possui um interpretador FORTRAN e utiliza vários outros pacotes de programas previamente desenvolvidos no CERN como o HBOOK/HPLOT, que contém rotinas de biblioteca (libraries) em FORTRAN, a CERN-LIB, e o HIGZ (Highlevel Interface to Graphics and Zebra) para implementar os recursos gráficos do programa. E, para incorporar de modo mais efetivo as diversas utilidades gráficas, o PAW contém também um sistema de instruções próprio, denominado KUIP (Kit for a User Interface Package), que pode ser utilizado interativamente ou por meio de arquivos de comandos, denominados KUMAC (Kuip Macros).

Um terceiro bloco de pacotes também previamente existentes utilizados pelo PAW são:

Minuit : programa de minimização utilizado no ajuste de funções .

Comis : interpretador FORTRAN interativo.

SIGMA : manipulador de vetores.

Nesse sentido, além das operações de conversão e leitura (em ambos os sentidos, input-output) dos elementos básicos do pacote, há que se aprender toda uma série de operações ou instruções de manipulação, tanto em seus aspectos gráficos quanto formal, que permitam a realização de uma análise de dados.

B.2 ROOT

O pacote gráfico ROOT⁴ foi criado a partir da linguagem orientada para objetos C++ e, por isso, seus elementos básicos, são rotinas de programação pré-definidas, denominadas **classes**, por meio das quais podem-se criar os diversos objetos que servirão de ponto de partida para uma análise de dados.

Ao contrário do PAW, tanto os elementos básicos para um tratamento estatístico de dados (histogramas e n-túplas) quanto as operações realizadas sobre eles são implementadas como rotinas escritas em C++ que são lidas diretamente por seu próprio interpretador de C++. Nesse sentido não há a necessidade do aprendizado de um novo sistema de instruções, apenas do conhecimento das classes pré-definidas que estão a disposição para uso interativo ou para arquivos de comando.

B.3 Exemplos

Como primeira aplicação simples, utilizemos o PAW e o ROOT para representar graficamente uma distribuição normal.

PAW	ROOT
PAW > fu/pl exp(-x**2/2) -3 3	root [1] TF1 f1("f1","exp(-x**2/2)",-3,3)
	root [1] f1.Draw()

⁴<http://root.cern.ch>

Um exemplo mais complexo, que envolve a geração de uma amostra aleatória, o armazenamento de dados em histogramas e o ajuste de funções é ilustrado pelo programa **dados.f**, em FORTRAN, para o PAW.

```

SUBROUTINE dados(N,M)
dimension x(6)

ER=N/6. *frequencia esperada de cada resultado
CALL HBOOK1(10000,'Distribuicao da Media ',200,3.,4.,0.)
CALL HBOOK1(20000,'Chi-2 ',80,0.,20.,0.)

do k=1,M
  CALL HBOOK1(k,'Brincando com dados ',6,1.,7.,0.)
  XM=0.
  chi_2=0.
  x(1)=0.
  x(2)=0.
  x(3)=0.
  x(4)=0.
  x(5)=0.
  x(6)=0.

  do J=1,N
    IX1=6*RNDM(.01234)+1 *gerar inteiro de 1 a 6
    XM=XM+FLOAT(IX1)
    if (IX1.eq.1) then x(1)=x(1)+1 endif
    if (IX1.eq.2) then x(2)=x(2)+1 endif
    if (IX1.eq.3) then x(3)=x(3)+1 endif
    if (IX1.eq.4) then x(4)=x(4)+1 endif
    if (IX1.eq.5) then x(5)=x(5)+1 endif
    if (IX1.eq.6) then x(6)=x(6)+1 endif
    CALL HFILL(k,FLOAT(IX1),0.,1.)
  end do
  chi_2=(x(1)**2+x(2)**2+x(3)**2+x(4)**2+x(5)**2+x(6)**2)/ER
  chi_2=chi_2-N
  XM=XM/N
  CALL HFILL(10000, XM,0.,1.)
  CALL HFILL(20000,chi_2,0.,1.)
end do

```

end

Este programa, que necessita de dois parâmetros N e M , simula a geração de M amostras de N lançamentos de um dado. Nele, são criados os histogramas 10000 e 20000, que contêm as distribuições da **média** dos resultados e do parâmetro χ^2 relativas às M amostras, e M histogramas (numerados de 1 a M) dos resultados de cada amostra. Pode ser executado, por exemplo, interativamente, do seguinte modo:

```
PAW > call dados.f(100,300)
```

onde os números 100 e 300 são passados como parâmetros N e M , respectivamente.

Uma vez gerado um histograma, no PAW, o mesmo pode ser armazenado num arquivo externo de forma interativa. Por exemplo, para armazenar o histograma (20000) da distribuição de χ^2 , pode-se proceder do seguinte modo:

```
PAW > h/fil 10 dados.his 1024 N
PAW > hrout 20000
PAW > close 10
```

Onde:

- `h/fil` é o comando para criar (parâmetro N), no disco do computador, um arquivo denominado `dados.his` e identificado pelo número 10.
- `hrout` é o comando para copiar para o arquivo `dados.his` o histograma 20000 armazenado na memória.
- `close 10` é simplesmente o comando para fechar o arquivo `dados.his` a novos comandos.

Se quisermos ajustar, interativamente, o histograma (10000) da distribuição da média por uma gaussiana, já incorporada no PAW, pode-se proceder do seguinte modo:

```
PAW > h/fit 10000 G
```

onde o comando `h/fit` necessita dos parâmetros 10000, que identifica o histograma a ser ajustado, e `G`, que identifica a função de ajuste (Func) como uma gaussiana. A função gaussiana as outras funções pré-definidas (exponencial e polinomial) possuem os seguintes parâmetros (`par()`):

```
G : Func=par(1)*exp(-0.5*((x-par(2))/par(3))**2)
E : Func=exp(par(1)+par(2)*x)
Pn: Func=par(1)+par(2)*x+par(3)*x**2+...+par(n+1)*x**n
```

que podem ser visualizados se antes do ajuste utilizarmos o comando,

```
PAW > opt fit
```

No caso do histograma (20000) da distribuição de χ^2 a função de ajuste deve ser escrita num arquivo FORTRAN externo. Por exemplo, o arquivo `chi.f`:

```
real function chi(x)
common/pawpar/par(2)
chi=par(1)*x**(par(2)/2-1)*exp(-x/2)
end
```

Pode se utilizado da seguinte forma:

```
PAW > v/cr par(2) R 10 2
PAW > h/fit 20000 chi.f L 2 par
```

onde os comandos:

- `v/cr` cria o vetor de parâmetros `par` reais (`R`), de dimensão 2, com os valores 10 e 2.
- `h/fit` ajusta pelo método da máxima verossimilhança (`L`) a função `chi.f`, que possui dois parâmetros (2), inicializados pelo vetor `par`.

Alternativamente, todos esses comandos poderiam ser agrupados em um arquivo de comandos `dados.kumac`,

```

macro dados 1=100 2=9

call dados.f([1],[2])
h/fil 10 dados.his 1024 N
hrout 20000
close 10
*
opt fit
h/fit 10000 G
*
v/cr par(2) R 10. 2.
h/fit 20000 chi.f L 2 par

```

e chamados do seguinte modo:

```
PAW > exe dados 100 300
```

onde o comando `exe` chama o arquivo de comandos **dados.kumac**, usualmente chamado simplesmente de KUMAC, que passa os parâmetros 100 e 300 ao arquivo FORTRAN **dados.f**.

Esse modo, de primeiro trabalhar interativamente e, após o estabelecimento de critérios mais definitivos, sintetizá-los numa KUMAC, é procedimento usual no trabalho de análise com o PAW.

Podemos repetir o mesmo exemplo com o ROOT, através do seguinte programa escrito em C++.

```

{ //dados.C
gROOT->Reset();
TFile *hfile=new TFile("dados.root","RECREATE","Brincando com Dados");
TH1F *h1=new TH1F("h1","Media",200,3.,4.);
TH1F *h2=new TH1F("h2","Chi_2",80,0.,20.);
TNtuple *nt=new TNtuple("nt","Dado_ntup","XM:chi_2");
Int_t X1,M,N;
Float_t ER,XM,chi_2,x[6];
M=600; // Numero de amostras
N=100; // Numero de lancamentos

```

```

ER=N/6. // Frequencia esperada de cada resultado
gRandom->SetSeed();
for(Int_t i=1;i<=M;i++)
{
  XM=0.;x[0]=0.;x[1]=0.;x[2]=0.;x[3]=0.;x[4]=0.;x[5]=0.;chi_2=0.;
  for(Int_t i=1;i<=M;i++)
  {
    X1=6*gRandom->Rndm(1); //gerar inteiro de 1 a 6
    X1=X1+1;
    // printf(" X1=%d",X1);
    XM=XM+1;
    if (X1==1) x[0]=x[0]+1;
    if (X1==2) x[1]=x[1]+1;
    if (X1==3) x[2]=x[2]+1;
    if (X1==4) x[3]=x[3]+1;
    if (X1==5) x[4]=x[4]+1;
    if (X1==6) x[5]=x[5]+1;
  }
  XM=XM/N;
  chi_2=(x[0]**2+x[1]**2+x[2]**2+x[3]**2+x[4]**2+x[5]**2)/ER;
  chi_2=chi_2-N;
  h1->Fill(XM);
  h2->Fill(chi_2);
  nt->Fill(XM,chi_2);
}
hfile->Write();
}

```

Este programa, armazenado no arquivo **dados.C**, simula a mesma situação que o programa **dados.f**, com algumas modificações. Ou seja, a geração de M amostras de N lançamentos de um dado. Nele, ainda são criados histogramas, denominados **Media** e **Chi-2**, que contêm as distribuições da **média** dos resultados e do parâmetro χ^2 relativas às M amostras e, como novidade, em vez de M histogramas dos resultados de cada amostra, cria-se uma **n-úpla** que contém os histogramas **Media** e **Chi-2** e um arquivo **dados.root**, no formato ROOT, que contém todos esses objetos, para ser utilizado num trabalho de análise posterior.

Em geral, simulações mais complexas, de processos físicos são realizadas por programas que utilizam Métodos de Monte Carlo e, utilizamos o PAW ou ROOT apenas no trabalho de análise.

Apêndice C

Cinemática Relativística

C.1 Relatividade Restrita

C.1.1 Transformação de Lorentz

Consideremos um sistema de referência inercial S' onde um ponto do espaço-tempo é descrito pelas coordenadas (t', x', y', z') se movendo com velocidade constante $\vec{\beta}_s = \vec{V}_s/c$ em relação à um sistema S onde o mesmo ponto do espaço-tempo possui coordenadas (t, x, y, z) . Escolhendo $z \parallel z' \parallel \vec{\beta}_s$ e assumindo que a origem de ambos os sistemas de coordenadas coincidem em $t = t' = 0$ a transformação de Lorentz e sua inversa são dadas por:

$$\begin{aligned}t &= \gamma_s(t' + \beta_s z') , & x &= x' , & y &= y' , & z &= \gamma_s(z' + \beta_s t') \\t' &= \gamma_s(t - \beta_s z) , & x' &= x , & y' &= y , & z' &= \gamma_s(z - \beta_s t)\end{aligned}\tag{C.1}$$

onde tomamos $c = 1$ e definimos o fator de Lorentz como:

$$\gamma_s = (1 - \beta_s^2)^{-1/2}$$

As equações da Relatividade Restrita e, em particular, da Física de Partículas ficam bastante simplificadas se considerarmos o conceito de quadri-vetores. Para as coordenadas do espaço-tempo, escrevemos

$$x^\mu = (x^0, x^1, x^2, x^3) = (x^0, \vec{x}) = (t, x, y, z) \quad (\text{C.2})$$

Qualquer quadri-vetor A^μ se transforma sob Lorentz como x^μ , ou seja, sob a transformação Eq.(C.1), temos:

$$A^0 = \gamma_s(A'^0 + \beta_s A'^3), \quad A^1 = A'^1, \quad A^2 = A'^2, \quad A^3 = \gamma_s(A'^3 + \beta_s A'^0)$$

Introduzindo o tensor métrico:

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (\text{C.3})$$

temos:

$$A_\mu = g_{\mu\nu} A^\nu$$

Onde A^μ são chamadas de componentes *contravariantes* e A_μ *covariantes*, com:

$$A_0 = A^0, \quad A_1 = -A^1, \quad A_2 = -A^2, \quad A_3 = -A^3,$$

em particular,

$$x_\mu = (x_0, x_1, x_2, x_3) = (x_0, -\vec{x}) = (t, -x, -y, -z)$$

Definimos também o tensor unitário δ_ν^μ , tal que $\delta_\nu^\mu A^\nu = A^\mu$, com $\delta_\mu^\mu = 4$. Além destes dois tensores (*i.e.* $g^{\mu\nu}$ e δ_ν^μ), o tensor totalmente antissimétrico $\epsilon^{\mu\nu\alpha\beta}$ ($\epsilon^{0123} = +1$) também possui as mesmas componentes em qualquer sistema de referência.¹

¹Na realidade, $\epsilon^{\mu\nu\alpha\beta}$ é um pseudo-tensor uma vez que sob rotações ele se comporta como um tensor mas a mudança de sinal de uma ou três coordenadas **não** acarreta na mudança de sinal de suas coordenadas. Sob reflexão do sistema de coordenadas, *i.e.* sob a mudança de sinal de todas as coordenadas: um escalar e um pseudo-vetor ou vetor axial (*e.g.* $\vec{A} \times \vec{B}$) **não mudam** de sinal, enquanto, um pseudo-escalar (*e.g.* $\vec{A} \times \vec{B} \cdot \vec{C}$) e um vetor **mudam**. Ele possui ainda as seguintes propriedades:

$$\epsilon^{\mu\nu\alpha\beta} \epsilon_{\rho\sigma\alpha\beta} = -2(\delta_\rho^\mu \delta_\sigma^\nu - \delta_\sigma^\mu \delta_\rho^\nu), \quad \epsilon^{\mu\nu\alpha\beta} \epsilon_{\rho\nu\alpha\beta} = -6\delta_\rho^\mu \quad \text{e} \quad \epsilon^{\mu\nu\alpha\beta} \epsilon_{\mu\nu\alpha\beta} = -24$$

Podemos escrever o produto escalar entre dois quadrivetores como:

$$\begin{aligned}
 A \cdot B &\equiv \sum_{\mu=0}^3 A^\mu B_\mu = g_{\mu\nu} A^\mu B^\nu \\
 &= A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3 = A^0 B^0 - A^1 B^1 - A^2 B^2 - A^3 B^3 \\
 &= A^0 B^0 - \vec{A} \cdot \vec{B}
 \end{aligned} \tag{C.4}$$

Podemos também definir o *quadrimomento* de uma partícula como:

$$p^\mu = (E, \vec{p}) \tag{C.5}$$

onde,

$$\begin{aligned}
 E &= m\gamma \\
 \vec{p} &= m\gamma\vec{\beta}
 \end{aligned} \tag{C.6}$$

onde m é a massa da partícula com velocidade β :

$$\vec{\beta} = \vec{p}/E, \quad \text{e} \quad \gamma = (1 - \beta^2)^{-1/2} = E/m$$

O quadrado do quadrimomento de uma partícula satisfaz:

$$p^2 = p^\mu p_\mu = E^2 - |\vec{p}|^2 = m^2$$

Vamos escrever o momento de uma partícula em coordenadas esféricas:

$$\vec{p} = (p_x, p_y, p_z) = |\vec{p}| (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$$

e, como iremos tratar na maioria dos casos transformações ao longo do eixo z , vamos definir o *momento longitudinal* e *transversal* como:

$$\begin{aligned}
 p_{\parallel} &= p_z = |\vec{p}| \cos \theta \\
 p_{\perp} &= (p_x^2 + p_y^2)^{1/2} = |\vec{p}| \sin \theta
 \end{aligned} \tag{C.7}$$

Desta forma, a energia e o momento (E', \vec{p}') de uma partícula vistos de um sistema que se move com velocidade $\vec{\beta}_s$ na direção z são dados por:

$$\begin{pmatrix} E' \\ p'_{\parallel} \end{pmatrix} = \begin{pmatrix} \gamma_s & -\gamma_s \beta_s \\ -\gamma_s \beta_s & \gamma_s \end{pmatrix} \begin{pmatrix} E \\ p_{\parallel} \end{pmatrix}, \quad p'_{\perp} = p_{\perp} \tag{C.8}$$

onde p_{\perp} (p_{\parallel}) são as componentes de \vec{p} perpendicular (paralela) à $\vec{\beta}_s$. Esta última expressão é apenas uma maneira mais conveniente de se escrever a Eq.(C.1).

Exercício 3.1 Mostre que o produto escalar de quaisquer dois quadri-
vetores $p_1 \cdot p_2 = E_1 E_2 - \vec{p}_1 \cdot \vec{p}_2$ é invariante sob transformações de Lorentz,
ou seja, independente de sistema de referência.

A transformação de ângulos sob transformações de Lorentz pode ser facilmente ex-
traída da Eq.(C.8). Para o caso do ângulo azimutal ϕ , temos:

$$\tan \phi' = \frac{p'_y}{p'_x} = \frac{p_y}{p_x} = \tan \phi$$

ou seja, o ângulo azimutal ao redor de um eixo é invariante sob transformações de Lorentz
ao longo deste eixo.

Para o ângulo polar, ou seja, o ângulo que o momento da partícula faz com o eixo do
“boost” visto do sistema S' é:

$$\begin{aligned} \tan \theta' &= \frac{p'_\perp}{p'_\parallel} = \frac{p_\perp}{p_\parallel} \\ &= \frac{p_\perp}{\gamma_s(p_\parallel - \beta_s E)} = \frac{1}{\gamma_s} \frac{p_\perp/|\vec{p}|}{(p_\parallel/|\vec{p}| - \beta_s E/|\vec{p}|)} \\ &= \frac{1}{\gamma_s} \frac{\sin \theta}{(\cos \theta - \beta_s/\beta)} \end{aligned} \quad (\text{C.9})$$

ou, equivalentemente:

$$\tan \theta = \frac{1}{\gamma_s} \frac{\sin \theta'}{(\cos \theta' + \beta_s/\beta')}$$

Para introduzirmos o conceito de *rapidez* vamos escrever a transformação Eq.(C.8)
como:

$$\begin{pmatrix} E' \\ p'_\parallel \end{pmatrix} = \begin{pmatrix} \cosh y & -\sinh y \\ -\sinh y & \cosh y \end{pmatrix} \begin{pmatrix} E \\ p_\parallel \end{pmatrix}, \quad p'_\perp = p_\perp \quad (\text{C.10})$$

onde definimos o parametro y chamado de rapidez através de:

$$\beta_s = \tanh y, \quad \gamma_s = \cosh y, \quad \gamma_s \beta_s = \sinh y \quad (\text{C.11})$$

ou, inversamente²:

$$\begin{aligned} y &= \cosh^{-1}(\gamma) = \ln(\gamma + \gamma\beta_s) \\ &= \tanh^{-1}(\beta_s) = \frac{1}{2} \ln \left(\frac{1 + \beta_s}{1 - \beta_s} \right) \end{aligned} \quad (\text{C.12})$$

Supondo a partícula parada em s' , temos $\beta = \beta_s$ e, como:

$$E \pm p_{\parallel} = m\gamma(1 \pm \beta)$$

podemos escrever a rapidez de uma partícula como:

$$y = \frac{1}{2} \ln \left(\frac{E + p_{\parallel}}{E - p_{\parallel}} \right)$$

Enquanto a velocidade varia no intervalo $-1 \leq \beta \leq 1$, a rapidez varia $-\infty \leq y \leq \infty$.

Quando executamos duas transformações de Lorentz paralelas consecutivas de parâmetros β_1 e β_2 o resultado combinado é expresso pelos parâmetros:

$$\begin{aligned} \beta_3 &= \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2} \\ \gamma_3 &= \gamma_1\gamma_2(1 + \beta_1\beta_2) \end{aligned} \quad (\text{C.13})$$

Exercício 3.2 Mostre as relações (C.13)

Desta forma, em termos da rapidez, temos:

$$\begin{aligned} \beta_3 &= \tanh y_3 = \frac{\tanh y_1 + \tanh y_2}{1 + \tanh y_1 \tanh y_2} \\ &\equiv \tanh(y_1 + y_2) \end{aligned} \quad (\text{C.14})$$

ou seja, a rapidez é aditiva sob transformações de Lorentz paralelas *i.e.* : $y_3 = y_1 + y_2$

²Lembre que:

$$\sinh^{-1}(z) = \ln(z + \sqrt{z^2 + 1}), \quad \cosh^{-1}(z) = \ln(z + \sqrt{z^2 - 1}), \quad \text{e} \quad \tanh^{-1}(z) = \frac{1}{2} \ln \left(\frac{1 + z}{1 - z} \right)$$

C.1.2 Sistemas de Referência

Consideremos a colisão de duas partículas de quadrimomento (E_a, \vec{p}_a) e (E_b, \vec{p}_b) . Na descrição destas colisões, dois sistemas de referência são usualmente utilizados:

(i) *Sistema de Centro de Massa (CM)*: É definido como o sistema onde

$$\vec{p}_a = \vec{p}_b$$

e, as quantidade referentes a este sistema serão denotadas a partir de agora por um asterisco (*).

(ii) *Sistema de Laboratório (LAB)*: É o sistema no qual são feitas as medidas. No caso de experimento de alvo fixo este sistema coincide com o sistema do alvo, onde uma das partículas (*e.g.* b) encontra-se em repouso *i.e.* $\vec{p}_b = 0$. Nos experimentos de anéis de colisão, onde feixes de partículas idênticas colidem em direções opostas, este sistema coincide com o CM.

Escolhendo o eixo z como a direção de movimento, os momentos das partículas a e b podem ser escritos, no sistema de CM como:

$$\begin{aligned} p_a^* &= (E_a^*, 0, 0, p_a^*) \\ p_b^* &= (E_b^*, 0, 0, -p_a^*) \end{aligned} \quad (\text{C.15})$$

e, no sistema LAB (alvo fixo, b em repouso) como:

$$\begin{aligned} p_a^{lab} &= (E_a^{lab}, 0, 0, p_a^{lab}) \\ p_b^{lab} &= (m_b, 0, 0, 0) \end{aligned} \quad (\text{C.16})$$

A transformação de Lorentz entre os dois sistemas será dada por:

$$\begin{aligned} E_i^* &= \gamma_{cm}(E_i^{lab} - \beta_{cm}p_i^{lab}) \\ p_i^* &= \gamma_{cm}(p_i^{lab} - \beta_{cm}E_i^{lab}) \end{aligned} \quad (\text{C.17})$$

onde $i = a, b$ e β_{cm} é a velocidade do CM no LAB. Como $\vec{p}_a^* + \vec{p}_b^* = 0$ da Eq.(C.17) temos:

$$\gamma_{cm}[(p_a^{lab} + p_b^{lab}) - \beta_{cm}(E_a^{lab} + E_b^{lab})] = 0$$

e portanto β_{cm} é dada por:

$$\beta_{cm} = \frac{|\vec{p}_a + \vec{p}_b|}{E_a^{lab} + E_b^{lab}} (\vec{p}_b = 0) = \frac{p_a^{lab}}{E_a^{lab} + m_b}$$

No caso de uma colisão de duas partículas de massa m_a e m_b a energia total da colisão ou *massa invariante* é dada por:

$$\begin{aligned} E_T &= (p_a + p_b)^{1/2} = [(E_a + E_b)^2 - (\vec{p}_a + \vec{p}_b)^2]^{1/2} \\ &= [m_a^2 + m_b^2 + 2E_a E_b (1 - \beta_a \beta_b \cos \theta)]^{1/2} \end{aligned} \quad (\text{C.18})$$

onde θ é o ângulo entre as duas partículas. Podemos ver que na definição de E_T aparecem apenas produtos escalares de quadrimomentos o que faz de E_T uma quantidade independente de sistema de referência. Em termos das variáveis de CM e LAB onde a partícula b está em repouso temos:

$$\begin{aligned} E_T &= (E_a^* + E_b^*) \\ &= [(E_a^{lab} + m_b)^2 - (\vec{p}_a^{lab})^2]^{1/2} = (m_a^2 + m_b^2 + 2m_b E_a^{lab})^{1/2} \end{aligned} \quad (\text{C.19})$$

Exercício 3.3 Mostre que $\gamma_{cm} = (E_a^{lab} + m_b)/E_T$

A quantidade E_T pode ser considerada como a energia útil de um acelerador. Para termos idéia de algumas ordens de magnitude consideremos protons ($m_p \simeq 1 \text{ GeV}$) colidindo a uma energia de 1 TeV com outro proton em repouso. Neste caso:

$$E_T \simeq \sqrt{2m_p E_p^{lab}} = \sqrt{2000} \text{ GeV} \simeq 45 \text{ GeV}$$

Por outro lado se considerarmos dois protons em um anél de colisão colidindo com 1 TeV, a energia total será:

$$E_T = E_{pa}^* + E_{pb}^* = 2E_p = 2000 \text{ GeV}$$

Da Eq.(C.17) podemos escrever as variáveis de CM em termos das do LAB:

$$\begin{aligned} E_a^* &= \frac{m_a^2 + m_b E_a^{lab}}{E_T} \\ E_b^* &= \frac{m_b(m_b + E_a^{lab})}{E_T} \\ p_a^* &= \frac{m_b p_a^{lab}}{E_T} \\ p_b^* &= \frac{-m_b p_a^{lab}}{E_T} = -p_a^* \end{aligned} \quad (\text{C.20})$$

Tabela

C.1: Energia e Momento CM versus Momento do Feixe de Protons

$p_a^{lab} (GeV)$	$E_T (GeV)$		$p_a^* (GeV)$	
	ep	pp	ep	pp
1	1.66	2.08	0.57	0.45
10	4.43	4.54	2.12	2.07
100	13.73	13.76	6.83	6.82
500	30.65	30.66	15.3	15.3
1000	43.33	43.34	21.7	21.6
10000	137.0	137.0	68.5	68.5

Exercício 3.4 Confira as Eq.(C.20)

Apresentamos na Tabela (C.1) os valores da energia total (E_T) e do momento (p_a^*) no CM para a colisão de feixes de eletron ($m_a = m_e \simeq 0$) e proton ($m_a = m_p \simeq 0.938$ GeV) colidindo com um proton fixo em função do momento do feixe (p_a^{lab}).

Podemos escrever de maneira compacta as variáveis não invariantes ($E_{a,b}$ e $p_{a,b}$) em termos das invariantes (E_T e $m_{a,b}$) nos diferentes sistemas de referência. A quantidade E_T^2 é usualmente designada pela variável de Mandelstam s , que será introduzida posteriormente. Passaremos portanto a utilizar esta notação para o quadrado da massa invariante.

Vamos definir a função cinemática:

$$\begin{aligned}
 \lambda(x, y, z) &= (x - y - z)^2 - 4yz \\
 &= x^2 + y^2 + z^2 - 2xy - 2yz - 2zx \\
 &= [x - (\sqrt{y} + \sqrt{z})^2] [x - (\sqrt{y} - \sqrt{z})^2]
 \end{aligned} \tag{C.21}$$

que, para alguns valores especiais fica:

$$\begin{aligned}
 \lambda(x, y, y) &= x(x - 4y) \\
 \lambda(x, y, 0) &= (x - y)^2
 \end{aligned} \tag{C.22}$$

Utilizando a definição de $\lambda(x, y, z)$ podemos reescrever Eq.(C.20). Como $\vec{p}_a^* + \vec{p}_b^* = 0$ escrevemos:

$$|\vec{p}_a^*| = |\vec{p}_b^*| = p^*$$

e, conforme Eq.(C.19)

$$\sqrt{s} = E_a^* + E_b^*$$

ou seja:

$$\sqrt{s} = [p^{*2} + m_a^2]^{1/2} + [p^{*2} + m_b^2]^{1/2} = E_a^* + [E_a^{*2} - m_a^2 + m_b^2]^{1/2}$$

Obtemos portanto:

$$\begin{aligned} E_a^* &= \frac{s + m_a^2 - m_b^2}{2\sqrt{s}} \\ E_b^* &= \sqrt{s} - E_a^* = \frac{s - m_a^2 + m_b^2}{2\sqrt{s}} \\ p^* &= \frac{\lambda^{1/2}(s, m_a^2, m_b^2)}{2\sqrt{s}} \end{aligned} \tag{C.23}$$

E, utilizando a Eq.(C.19), temos:

$$\begin{aligned} E_a^{lab} &= \frac{s - m_a^2 - m_b^2}{2m_b} \\ p_a^{lab} &= [(E_a^{lab})^2 - m_a^2]^{1/2} = \frac{\lambda^{1/2}(s, m_a^2, m_b^2)}{2m_b} \end{aligned} \tag{C.24}$$

Exercício 3.5 Obtenha as Eq.(C.23) e Eq.(C.24)

C.2 Decaimento

C.2.1 Decaimento de Dois Corpos

Consideremos uma partícula de massa M e quadrimomento P decaindo em duas outras de quadrimomentos:

$$p_i = (E_i, \vec{p}_i), \quad i = 1, 2$$

No sistema de repouso da partícula de massa M , os quadrimomentos se escrevem com:

$$\begin{aligned} P^* &= (M, 0, 0, 0) \\ p_1^* &= (E_1^*, \vec{p}^*) \\ p_2^* &= (E_2^*, -\vec{p}^*) \end{aligned} \quad (\text{C.25})$$

onde $M = E_1^* + E_2^*$ e $\vec{p}_1^* = -\vec{p}_2^* = \vec{p}^*$. Desta forma:

$$\begin{aligned} P^{*2} &= M^2 = (p_1^* + p_2^*)^2 = m_1^2 + m_2^2 + 2(E_1^*E_2^* + |\vec{p}^*|^2) \\ &= m_1^2 + m_2^2 + 2[E_1^*(M - E_1^*) + (E_1^{*2} - m_1^2)] \end{aligned} \quad (\text{C.26})$$

E podemos escrever as energias e momentos finais como:

$$\begin{aligned} E_1^* &= \frac{M^2 + m_1^2 - m_2^2}{2M} \\ E_2^* &= M - E_1^* = \frac{M^2 - m_1^2 + m_2^2}{2M} \\ |\vec{p}_1^*| = |\vec{p}_2^*| &= (E_1^{*2} - m_1^2)^{1/2} \\ &= \frac{[(M^2 - (m_1 + m_2)^2)(M^2 - (m_1 - m_2)^2)]^{1/2}}{2M} \\ &= \frac{\lambda^{1/2}(M^2, m_1^2, m_2^2)}{2M} \end{aligned} \quad (\text{C.27})$$

C.2.2 Decaimento de Três Corpos

Consideremos uma partícula de massa M e quadrimomento P decaindo agora em três outras de massas m_i e quadrimomentos p_i ; $i = 1, 2, 3$. Vamos primeiramente definir:

$$p_{ij} = p_i + p_j, \quad s_{ij} = p_{ij}^2$$

ou seja:

$$\begin{aligned} s_{12} &= (p_1 + p_2)^2 = (P - p_3)^2 \\ s_{23} &= (p_2 + p_3)^2 = (P - p_1)^2 \\ s_{31} &= (p_3 + p_1)^2 = (P - p_2)^2 \end{aligned} \quad (\text{C.28})$$

desta forma, temos:

$$s_{12} + s_{23} + s_{31} = M^2 + m_1^2 + m_2^2 + m_3^2$$

Exercício 3.6 Mostre a relação acima.

No sistema de repouso da partícula que decai temos $\vec{P}^* = \vec{p}_1^* + \vec{p}_2^* + \vec{p}_3^* = 0$ e podemos obter as energias e momentos em termos de s_{ij} substituindo $P = (M, 0, 0, 0)$ em Eq.(C.28). Por exemplo:

$$s_{12} = (P^* - p_3^*)^2 = M^2 + m_3^2 - 2ME_3^*,$$

de onde obtemos E_3^* e conseqüentemente p_3^* . Da mesma forma obtemos as demais energias e momentos:

$$\begin{aligned} E_1^* &= \frac{M^2 + m_1^2 - s_{23}}{2M} \\ E_2^* &= \frac{M^2 + m_2^2 - s_{31}}{2M} \\ E_3^* &= \frac{M^2 + m_3^2 - s_{12}}{2M} \\ p_1^* &= \frac{\lambda^{1/2}(M^2, m_1^2, s_{23})}{2M} \\ p_2^* &= \frac{\lambda^{1/2}(M^2, m_2^2, s_{31})}{2M} \\ p_3^* &= \frac{\lambda^{1/2}(M^2, m_3^2, s_{12})}{2M} \end{aligned} \quad (\text{C.29})$$

Em geral é mais conveniente considerar um sistema no qual duas das três partículas finais estão em repouso. Consideremos por exemplo o referencial (r) no qual as partículas 2 e 3 encontram-se em repouso *i.e.* $p_2^r + p_3^r = (\sqrt{s_{23}}, 0, 0, 0) = P^r - p_1^r$. Conseguimos as energias e momentos neste referencial substituindo esta relação na Eq.(C.28), *e.g.* :

$$\begin{aligned} s_{23} &= (P^r - p_1^r)^2 = [(\sqrt{s_{23}} + E_1^r, \vec{p}_1^r) - (E_1^r, \vec{p}_1^r)]^2 \\ &= M^2 + m_1^2 - 2[(\sqrt{s_{23}} + E_1^r)E_1^r - (E_1^r)^2 - m_1^2] \\ &= M^2 - m_1^2 - 2\sqrt{s_{23}}E_1^r \end{aligned} \quad (\text{C.30})$$

Assim, as energias e momentos neste referencial são dados por:

$$\begin{aligned}
 E^r &= \frac{M^2 + s_{23} - m_1^2}{2\sqrt{s_{23}}} \\
 E_1^r &= \frac{M^2 - s_{23} - m_1^2}{2\sqrt{s_{23}}} \\
 E_2^r &= \frac{s_{23} + m_2^2 - m_3^2}{2\sqrt{s_{23}}} \\
 E_3^r &= \frac{s_{23} - m_2^2 + m_3^2}{2\sqrt{s_{23}}} \\
 P^r = p_1^r &= \frac{\lambda^{1/2}(M^2, s_{23}, m_1^2)}{2\sqrt{s_{23}}} \\
 p_2^r = p_3^r &= \frac{\lambda^{1/2}(s_{23}, m_2^2, m_3^2)}{2\sqrt{s_{23}}} \tag{C.31}
 \end{aligned}$$

C.3 Espalhamento $2 \rightarrow 2$

Vamos definir as chamadas *Variáveis de Mandelstam* para o processo $p_a + p_b \rightarrow p_1 + p_2$:

$$\begin{aligned}
s &= (p_a + p_b)^2 = (p_1 + p_2)^2 \\
&= (E_a^* + E_b^*)^2 = (E_1^* + E_2^*)^2 \\
&= m_a^2 + m_b^2 + 2m_b E_a^{lab} \\
t &= (p_a - p_1)^2 = (p_b - p_2)^2 \\
&= m_a^2 + m_1^2 - 2E_a E_1 + 2p_a p_1 \cos \theta_{a1} , \quad \text{CM ou LAB} \\
&= m_b^2 + m_2^2 - 2m_b E_2^{lab} \\
u &= (p_a - p_2)^2 = (p_b - p_1)^2 \\
&= m_a^2 + m_2^2 - 2E_a E_2 + 2p_a p_2 \cos \theta_{a2} , \quad \text{CM ou LAB} \\
&= m_b^2 + m_1^2 - 2m_b E_1^{lab}
\end{aligned} \tag{C.32}$$

satisfazem,

$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2 \tag{C.33}$$

Exercício 3.7 Mostre a relação (C.33)

A energia e momento das partículas no CM em termos das variáveis de Mandelstam são:

$$\begin{aligned}
E_a^* &= \frac{s + m_a^2 - m_b^2}{2\sqrt{s}} \\
E_b^* &= \frac{s + m_b^2 - m_a^2}{2\sqrt{s}} \\
E_1^* &= \frac{s + m_1^2 - m_2^2}{2\sqrt{s}} \\
E_2^* &= \frac{s + m_2^2 - m_1^2}{2\sqrt{s}} \\
p_a^* = p_b^* &= \frac{\lambda^{1/2}(s, m_a^2, m_b^2)}{2\sqrt{s}} \\
p_1^* = p_2^* &= \frac{\lambda^{1/2}(s, m_1^2, m_2^2)}{2\sqrt{s}}
\end{aligned} \tag{C.34}$$

Introduzindo estes resultados na relação entre t e $\cos \theta_{a1}^*$, Eq.(C.32), temos:

$$\begin{aligned}\cos \theta_{a1}^* &= \frac{t - m_a^2 - m_1^2 + 2E_a^* E_1^*}{2p_a^* p_b^*} \\ &= \frac{2s(t - m_a^2 - m_1^2) + (s + m_a^2 - m_b^2)(s + m_1^2 - m_2^2)}{\lambda^{1/2}(s, m_a^2, m_b^2)\lambda^{1/2}(s, m_1^2, m_2^2)}\end{aligned}\quad (\text{C.35})$$

a relação para θ_{a2}^* pode ser obtida lembrando que $\theta_{a2}^* = \pi - \theta_{a1}^*$

Agora, a energia e momento das partículas no LAB, em termos das variáveis de Mandelstam se escrevem:

$$\begin{aligned}E_a^{lab} &= \frac{s - m_a^2 - m_b^2}{2m_b} \\ E_b^{lab} &= m_b \\ E_1^{lab} &= \frac{-u + m_b^2 + m_1^2}{2m_b} \\ E_2^{lab} &= \frac{-t + m_b^2 + m_2^2}{2m_b} \\ p_a^{lab} &= \frac{\lambda^{1/2}(s, m_a^2, m_b^2)}{2m_b} \\ p_b^{lab} &= 0 \\ p_1^{lab} &= \frac{\lambda^{1/2}(u, m_b^2, m_1^2)}{2m_b} \\ p_2^{lab} &= \frac{\lambda^{1/2}(t, m_b^2, m_2^2)}{2m_b}\end{aligned}\quad (\text{C.36})$$

E, as relações para os ângulos são:

$$\begin{aligned}\cos \theta_{a1}^{lab} &= \frac{2m_b^2(t - m_a^2 - m_1^2) + (s - m_a^2 - m_b^2)(-u + m_b^2 + m_1^2)}{\lambda^{1/2}(s, m_a^2, m_b^2)\lambda^{1/2}(u, m_b^2, m_1^2)} \\ \cos \theta_{a2}^{lab} &= \frac{2m_b^2(u - m_a^2 - m_2^2) + (s - m_a^2 - m_b^2)(-t + m_b^2 + m_2^2)}{\lambda^{1/2}(s, m_a^2, m_b^2)\lambda^{1/2}(t, m_b^2, m_2^2)}\end{aligned}\quad (\text{C.37})$$

Exercício 3.8 Verifique Eq.(C.34), Eq.(C.35), Eq.(C.36) e Eq.(C.37)

C.3.1 Região Física em s , t e u

Vamos determinar a região fisicamente aceitável no plano st para a reação $a+b \rightarrow 1+2$. Vamos assumir que $m_a = m_1 = m$ e $m_b = m_2 = M$. As energias e momentos são dados por:

$$\begin{aligned} E_a^* &= E_1^* = \frac{s + m^2 - M^2}{2\sqrt{s}} \\ E_b^* &= E_2^* = \frac{s + M^2 - m^2}{2\sqrt{s}} \\ p_a^* &= p_b^* = p_1^* = p_2^* \equiv p^* = \frac{\lambda^{1/2}(s, M^2, m^2)}{2\sqrt{s}} \end{aligned} \quad (\text{C.38})$$

O ângulo θ_{a1}^* é obtido da Eq.(C.35):

$$\cos \theta_{a1}^* = 1 + \frac{2st}{\lambda(s, M^2, m^2)} \quad (\text{C.39})$$

e, a variável t :

$$\begin{aligned} t &= -\frac{\lambda(s, M^2, m^2)}{2s}(1 - \cos \theta_{a1}^*) \\ &= -2(p^*)^2(1 - \cos \theta_{a1}^*) \\ &= -4(p^*)^2 \sin(\theta_{a1}^*/2) \end{aligned} \quad (\text{C.40})$$

Podemos obter a relação para u usando:

$$s + t + u = 2M^2 + 2m^2$$

obtendo:

$$u = \frac{(M^2 - m^2)^2}{s} - \frac{\lambda(s, M^2, m^2)}{2s}(1 + \cos \theta_{a1}^*) \quad (\text{C.41})$$

Exercício 3.9 Mostre as Eq.(C.39), Eq.(C.40) e Eq.(C.41)

As fronteiras da região física podem ser obtidas da relação :

$$-1 \leq \cos \theta_{a1}^* \leq 1$$

O limite superior de t é obtido quando $\cos \theta_{a1}^* = 1$

$$\begin{aligned} t &= 0 \\ u &= 2(M^2 + m^2) - s \end{aligned} \quad (\text{C.42})$$

e o limite inferior de t é obtido quando $\cos \theta_{a1}^* = -1$

$$\begin{aligned} t &= \frac{-\lambda(s, M^2, m^2)}{s} \\ u &= \frac{(M^2 - m^2)^2}{s} \end{aligned} \quad (\text{C.43})$$

Desta forma os limites são dados pela reta Eq.(C.42) e pela hipérbole Eq.(C.43) que possuem assíntotas:

$$\begin{aligned} s &= 0 \\ u &= 0 \quad \text{ou} \quad t = -s + 2M^2 + 2m^2 \end{aligned} \quad (\text{C.44})$$

As curvas Eq.(C.42) e Eq.(C.43) se intersectam em $s = (M^2 \pm m^2)^2$. O sinal positivo corresponde ao limiar da reação .

Os valores máximos e mínimos de t podem ser obtidos da Eq.(C.35) impondo $\cos \theta_{a1}^* = \pm 1$ e usando Eq.(C.34) para escrever as energias e momentos em termos de s :

$$\begin{aligned} t_{\pm} &= m_a^2 + m_1^2 - 2E_a^* E_1^* \pm 2p_a^* p_1^* \\ &= m_a^2 + m_1^2 - \frac{1}{2s} [(s + m_a^2 - m_b^2)(s + m_1^2 - m_2^2) \\ &\quad \pm \lambda^{1/2}(s, m_a^2, m_b^2) \lambda^{1/2}(s, m_1^2, m_2^2)] \end{aligned} \quad (\text{C.45})$$

onde t_- (t_+) é o maior (menor) valor de t . Note que t_+ é sempre negativo e $t_- > t_+$

Exercício 3.10 Verifique a Eq.(C.45)

C.4 Seção de Choque e Largura de Decaimento

A Seção de Choque para o processo $a + b \rightarrow 1 + \dots + n$ é dada por:

$$\begin{aligned} d\sigma &= \frac{1}{|\vec{\beta}_a - \vec{\beta}_b|} \frac{1}{(2E_a)} \frac{1}{(2E_b)} |\mathcal{M}|^2 \frac{d^3 p_1}{2E_1 (2\pi)^3} \dots \frac{d^3 p_n}{2E_n (2\pi)^3} \\ &\quad \cdot (2\pi)^4 \delta^4 [p_a + p_b - (p_1 + \dots + p_n)] S \end{aligned} \quad (\text{C.46})$$

ou seja,

$$d\sigma = \frac{1}{|\vec{\beta}_a - \vec{\beta}_b|} \frac{1}{(2E_a)} \frac{1}{(2E_b)} \frac{1}{(2\pi)^{3n-4}} \frac{S}{|\mathcal{M}|^2} \prod_{i=1}^n \frac{d^3 p_i}{2E_i} \delta^4 \left(p_a + p_b - \sum_{i=1}^n p_i \right) \quad (\text{C.47})$$

onde S é o fator estatístico que leva em conta o fato de haver m partículas idêntica no estado final, *i.e.* :

$$S = \prod_i \frac{1}{m_i!}$$

Podemos ver que:

$$\frac{1}{|\vec{\beta}_a - \vec{\beta}_b|} = \frac{E_a E_b}{[(p_a \cdot p_b)^2 - m_a^2 m_b^2]^{1/2}} = \frac{2E_a E_b}{\lambda^{1/2}(s, m_a^2, m_b^2)} \quad (\text{C.48})$$

Exercício 3.11 Mostre as duas igualdades da Eq.(C.48). Lembre que $\cos \theta_{ab} = \pi$

e definindo o fator de fluxo por:

$$\mathcal{F} = 2(2\pi)^{3n-4} \lambda^{1/2}(s, m_a^2, m_b^2) \quad (\text{C.49})$$

e o espaço de fase diferencial por:

$$d\mathcal{R}_n = \prod_{i=1}^n \frac{d^3 p_i}{2E_i} \delta^4 \left(p_a + p_b - \sum_{i=1}^n p_i \right) \quad (\text{C.50})$$

podemos escrever³:

$$d\sigma = \frac{S}{\mathcal{F}} |\mathcal{M}|^2 d\mathcal{R}_n \quad (\text{C.51})$$

³É importante notar que esta fórmula também é válida para fermions quando adotamos a normalização do spinor de Dirac igual a $2m$, ou seja:

$$\sum_{\pm s} u(p, s) \bar{u}(p, s) = \not{p} + m \quad \text{e} \quad \sum_{\pm s} v(p, s) \bar{v}(p, s) = \not{p} - m$$

A quantidade \mathcal{M} é chamada de *Amplitude Invariante*. Ela contém toda a informação sobre a dinâmica do processo que está sendo estudado sendo calculada para um modelo específico através das chamadas *Regras de Feynman*. Podemos ver das Eq.(C.49) e (C.50) que $[\mathcal{F}] = E^2$ e $[\mathcal{R}_n] = E^{2n-4}$. Uma vez que $[\sigma] = E^{-2}$, podemos ver que $[\mathcal{M}] = E^{2-n}$, ou seja, em um processo $2 \rightarrow 2$ a amplitude invariante é adimensional.

Quando a integral em (C.51) é feita sobre todo o espaço de fase de dimensão $3n - 4$ obtemos a *seção de choque total* da reação ; se a integração for restrita a um sub-espaço do espaço de fase temos a *seção de choque diferencial*. Chamamos de *reação exclusiva* aquela na qual a energia e momento de todas as partículas é medido. Numa *reação inclusiva* apenas a energia e momento de algumas das partículas finais são medidos.

A Largura de Decaimento $p \rightarrow p_1 + \dots + p_n$ é dada por:

$$\begin{aligned} d\Gamma &= \frac{1}{2M} |\mathcal{M}|^2 \prod_{i=1}^n \frac{d^3 p_i}{2E_i (2\pi)^3} (2\pi)^4 \delta^4 \left(p - \sum_{i=1}^n p_i \right) S \\ &= \frac{1}{2M} \frac{S}{(2\pi)^{3n-4}} |\mathcal{M}|^2 d\mathcal{R}_n \end{aligned} \quad (\text{C.52})$$

Quando integramos sobre todo o espaço de fase obtemos a largura de decaimento que é o inverso da vida média da partícula. Podemos ver que o mesmo espaço de fase diferencial Eq.(C.50) aparece aqui ocorrendo apenas a troca $p_a + p_b \rightarrow p$ onde $p_{a,b}$ são os quadrimomentos iniciais da colisão e p é o quadrimomento da partícula que decai.

C.5 Espaço de Fase

Em uma reação $2 \rightarrow n$ (*i.e.* $a + b \rightarrow 1 + 2 + \dots + n$) devemos impor a conservação de energia e momento:

$$\begin{aligned} E_a + E_b &= \sum_{i=1}^n E_i \\ \vec{p}_a + \vec{p}_b &= \sum_{i=1}^n \vec{p}_i \end{aligned} \quad (\text{C.53})$$

com,

$$E_j^2 = |\vec{p}_j|^2 + m_j^2, \quad j = a, b, 1, \dots, n$$

Chamamos de *espaço dos momentos* o espaço de $3n$ dimensões dos momentos \vec{p}_i . As condições (C.53) restringem o conjunto de \vec{p}_i possíveis e definem neste espaço um sub-espaço de dimensão $3n - 4$ chamado de *espaço de fase*.

Numa reação $2 \rightarrow n$ existem $3n - 4$ variáveis do estado final ($3n$ componentes de momento menos 4 vínculos (C.53)). Porém, neste caso, o eixo do feixe inicial define uma direção no espaço em relação à qual o processo possui simetria levando a uma variável trivial ϕ . Assim temos $3n - 5$ variáveis essenciais no estado final. Se levarmos em conta o estado inicial, existe mais uma variável essencial que é o quadrado da energia total s levando portanto a um total de $3n - 4$ variáveis essenciais (no processo $2 \rightarrow 2$ estas duas variáveis são s e t).

Vamos nos deter agora ao espaço de fase diferencial. Podemos ver que o elemento de integração $d^3p/2E$ é invariante de Lorentz. Supondo um “boost” na direção de p_z , temos:

$$dp_z = \gamma(dp'_z + \beta dE') = \frac{E}{E'} dp'_z \quad (\text{C.54})$$

Exercício 3.12 Verifique a Eq.(C.54). Lembre-se que $E'^2 = p'^2 + m^2$ e $E = \gamma(E' + \beta p'_z)$

Desta forma,

$$\frac{d^3p}{2E} = \frac{dp_x dp_y dp_z}{2E} = \frac{d^3p'}{2E'}$$

Podemos também escrever o elemento de espaço de fase em forma integral como:

$$\frac{d^3p}{2E} = \int d^4p \delta(p^2 - m^2) \Theta(p^0) \quad (\text{C.55})$$

onde a integral é feita sobre todos os valores das componentes p^μ e $\Theta(p^0) = 0(1)$ para $p^0 < (>)0$.

Exercício 3.13 Escreva $p^2 = (p^0)^2 - \vec{p}^2$ e $E^2 = \vec{p}^2 + m^2$ e use a propriedade da função delta

$$\delta[f(x)] = \frac{1}{|f'(x_0)|} \delta(x - x_0), \quad f(x_0) = 0$$

para mostrar a Eq.(C.55)

Desta forma \mathcal{R}_n Eq.(C.50) pode ser escrito como:

$$\mathcal{R}_n = \int \prod_{i=1}^n d^4 p_i \delta(p_i^2 - m_i^2) \Theta(p_i^0) \delta^4 \left(p_a + p_b - \sum_{i=1}^n p_i \right) \quad (\text{C.56})$$

Em geral omite-se a função $\Theta(p_i^0)$ que apenas denota o fato da energia ser uma quantidade positiva.

C.5.1 Espaço de Fase de Uma Partícula

O espaço de fase de uma partícula é trivial. Utilizando a Eq.(C.50) e Eq.(C.55) o espaço de fase para a reação $p_a + p_b \rightarrow p$, fica:

$$d\mathcal{R}_1 = \int d^4 p \delta(p^2 - m^2) \delta^4(p_a + p_b - p)$$

ou, integrando em $d^4 p$ usando a função δ^4 :

$$\mathcal{R}_1 = \delta[(p_a + p_b)^2 - m^2] = \delta(s - m^2) \quad (\text{C.57})$$

C.5.2 Espaço de Fase de Duas Partículas

Vamos agora considerar a integral do espaço de fase de duas partículas:

$$\mathcal{R}_2 = \int d^4 p_1 d^4 p_2 \delta(p_1^2 - m_1^2) \delta(p_2^2 - m_2^2) \delta^4(p - p_1 - p_2) \quad (\text{C.58})$$

No caso do espalhamento $2 \rightarrow 2$ temos $p = p_a + p_b$ e no caso do decaimento de uma partícula, p é o seu momento.

Integrando em $d^4 p_2$ usando a δ^4 obtendo:

$$\mathcal{R}_2 = \int d^4 p_1 \delta(p_1^2 - m_1^2) \delta[(p - p_1)^2 - m_2^2]$$

Lembrando que \mathcal{R}_n é invariante de Lorentz podemos escolher o referencial onde $p = (\sqrt{s}, 0, 0, 0)$ ou seja o CM das partículas iniciais ou o sistema de repouso da partícula que decai ($\sqrt{s} = M$):

$$\mathcal{R}_2 = \int d^4 p_1 \delta(p_1^2 - m_1^2) \delta(s - 2\sqrt{s}E_1 + m_1^2 - m_2^2)$$

Usando a Eq.(C.55) temos:

$$\mathcal{R}_2 = \int \frac{d^3 p_1}{2E_1} \delta(s - 2\sqrt{s}E_1^* + m_1^2 - m_2^2)$$

Note que não mencionamos a função $\Theta(p_0) = \Theta(\sqrt{s} - m_1 - m_2)$ que aparece na Eq.(C.55). Ela neste caso apenas garante que \mathcal{R}_2 seja nulo abaixo do limiar de produção das partículas 1 e 2 em repouso *i.e.* $\sqrt{s} < m_1 + m_2$. Como

$$d^3 p = p^2 dp d\Omega = p E dE d\Omega$$

temos:

$$\mathcal{R}_2 = \int \frac{p_1^* dE_1^* d\Omega_1^*}{2} \delta(s - 2\sqrt{s}E_1^* + m_1^2 - m_2^2)$$

onde, $d\Omega_1^* = \sin\theta d\theta_1^* d\phi_1^*$ descreve a orientação do momento \vec{p}_1^* no sistema CM (repouso) de p . A função δ fixa o valor de E_1^* e conseqüentemente o modulo do momento p_1^* (ver Eq.(C.34)), *i.e.* :

$$p_1^* = p_2^* = [(E_1^*)^2 - (m_1^*)^2]^{1/2} = \frac{\lambda^{1/2}(s, m_1^2, m_2^2)}{2\sqrt{s}}$$

Desta forma, integrando a função δ , lembrando que $\delta(ax) = (1/a)\delta(x)$ temos:

$$\mathcal{R}_2 = \frac{p_1^*}{4\sqrt{s}} \int d\Omega_1^* = \frac{\lambda^{1/2}(s, m_1^2, m_2^2)}{8s} \int d\Omega_1^* \quad (\text{C.59})$$

Em geral a Amplitude Invariante depende do ângulo θ^* . No entanto, se \mathcal{M} independe de θ^* , chegamos ao resultado final com $\int d\Omega_1^* = 4\pi$:

$$\mathcal{R}_2 = \frac{\pi p_1^*}{\sqrt{s}} = \frac{\pi \lambda^{1/2}(s, m_1^2, m_2^2)}{2s}$$

Utilizando as Eq.(C.51,C.59), a seção de choque de um processo $a + b \rightarrow 1 + 2$ fica:

$$\begin{aligned} d\sigma &= \frac{S}{2(2\pi)^2 \lambda^{1/2}(s, m_a^2, m_b^2)} |\mathcal{M}|^2 d\mathcal{R}_2 \\ &= \frac{S}{2(2\pi)^2 \lambda^{1/2}(s, m_a^2, m_b^2)} |\mathcal{M}|^2 \frac{\lambda^{1/2}(s, m_1^2, m_2^2)}{8s} d\Omega_1^* \end{aligned} \quad (\text{C.60})$$

Portanto:

$$\begin{aligned}\frac{d\sigma}{d\Omega_1^*} &= \frac{S}{64\pi^2 s} \frac{\lambda^{1/2}(s, m_1^2, m_2^2)}{\lambda^{1/2}(s, m_a^2, m_b^2)} |\mathcal{M}|^2 \\ &= \frac{S}{64\pi^2 s} \frac{p_1^*}{p_a^*} |\mathcal{M}|^2\end{aligned}\quad (\text{C.61})$$

onde $p_{a(1)}^*$ são os momentos iniciais (finais) no CM, dados pela Eq.(C.34).

Em geral é mais interessante termos uma expressão invariante para a seção de choque diferencial. Isto pode ser feito usando a definição de t Eq.(C.32):

$$dt = 2p_a^* p_1^* d(\cos \theta_{a1}^*) = \frac{1}{\pi} p_a^* p_1^* d\Omega_1^*$$

Desta forma:

$$\frac{d\sigma}{dt} = \frac{d\sigma}{d\Omega_1^*} \frac{d\Omega_1^*}{dt} = \frac{\pi}{p_a^* p_1^*} \frac{d\sigma}{d\Omega_1^*}$$

ou seja:

$$\begin{aligned}\frac{d\sigma}{dt} &= \frac{S}{64\pi s} \frac{1}{(p_a^*)^2} |\mathcal{M}|^2 \\ &= \frac{S}{16\pi} \frac{1}{\lambda(s, m_a^2, m_b^2)} |\mathcal{M}|^2\end{aligned}\quad (\text{C.62})$$

Devemos lembrar que a integral em dt é feita levando em conta os limites Eq.(C.45), ou seja

$$\sigma = \int_{t_+}^{t_-} \frac{d\sigma}{dt}$$

Exercício 3.14 Suponha que $|\mathcal{M}|^2 = -t/s$. Calcule a seção de choque total neste caso utilizando Eq.(C.61). Compare o resultado obtido quando se utiliza a Eq.(C.62). Considere $m_a = m_1 = 0$ e $m_b = m_2 = m$

Podemos também computar a largura de decaimento Eq.(C.52) de uma partícula de massa M , em repouso, em duas outras *i.e.* $P \rightarrow p_1 + p_2$:

$$\begin{aligned}d\Gamma &= \frac{1}{2M} \frac{S}{(2\pi)^2} |\mathcal{M}|^2 d\mathcal{R}_2 \\ &= \frac{1}{2M} \frac{S}{(2\pi)^2} |\mathcal{M}|^2 \frac{\lambda^{1/2}(M^2, m_1^2, m_2^2)}{8M^2} d\Omega_1^* \\ &= \frac{S}{64\pi^2 M^3} \lambda^{1/2}(M^2, m_1^2, m_2^2) |\mathcal{M}|^2 d\Omega_1^* \\ &= \frac{S}{32\pi^2 M^2} p_1^* |\mathcal{M}|^2 d\Omega_1^*\end{aligned}\quad (\text{C.63})$$

C.5.3 Espaço de Fase de Três Partículas

O espaço de fase para três partículas é:

$$\mathcal{R}_3 = \int \prod_{i=1}^3 \frac{d^3 p_i}{2E_i} \delta^3(\vec{p} - \vec{p}_1 - \vec{p}_2 - \vec{p}_3) \delta(E - E_1 - E_2 - E_3)$$

Vamos escolher o referencial onde $\vec{p} = 0$ e integrar sobre \vec{p}_2 usando a função δ^3 :

$$\mathcal{R}_3 = \int \frac{d^3 p_1 d^3 p_3}{8E_1 E_2 E_3} \delta(\sqrt{s} - E_1 - E_2 - E_3)$$

sendo E_2 dado por:

$$E_2^2 = |\vec{p}_1 + \vec{p}_3|^2 + m_2^2$$

Escrevendo o elemento de integração como:

$$d^3 p_1 d^3 p_3 = (p_1^2 dp_1 d\Omega_1)(p_3^2 dp_3 d\Omega_3) = (p_1 E_1 dE_1 d\Omega_1)[p_3 E_3 dE_3 (d \cos \theta_{13} d\phi_3)]$$

onde Ω_3 descreve a orientação de \vec{p}_3 em relação a \vec{p}_1 e Ω_1 a orientação de \vec{p}_1 em relação a algum eixo.

Podemos assim usar a função δ para integrar em $d \cos \theta_{13}$ usando:

$$\frac{dE_2}{d \cos \theta_{13}} = \frac{p_1 p_3}{E_2}$$

obtendo,

$$\mathcal{R}_3 = \frac{1}{8} \int dE_1 dE_3 d\Omega_1 d\phi_3$$

Podemos também escrever \mathcal{R}_3 em função de s_{12} e s_{23} , definidos anteriormente na Eq.(C.28) ($s = M^2$), usando o Jacobiano:

$$\frac{\partial(E_1, E_3)}{\partial(s_{12}, s_{23})} = \frac{1}{4s}$$

obtendo:

$$\mathcal{R}_3 = \frac{1}{32s} \int ds_{12} ds_{23} d\Omega_1 d\phi_3 \quad (\text{C.64})$$

Devemos notar que no caso da colisão de duas partículas (*i.e.* $p = p_a + p_b$) Ω_1 descreve a orientação de \vec{p}_1 no CM e ϕ_3 a rotação da configuração de todos os momentos em torno de um eixo. Para o caso do decaimento como não existe um eixo preferencial (imagine a partícula em repouso) podemos integrar em $d\Omega_1$ e $d\phi_3$ obtendo:

$$\mathcal{R}_3 = \pi^2 \int dE_1 dE_3 = \frac{\pi^2}{4s} \int ds_1 ds_2$$

Desta forma a largura de decaimento Eq.(C.52) fica:

$$\begin{aligned} \Gamma &= \frac{S}{(2\pi)^3} \frac{1}{8M} \int |\mathcal{M}|^2 dE_1 dE_3 \\ &= \frac{S}{(2\pi)^3} \frac{1}{32M^3} \int |\mathcal{M}|^2 ds_{12} ds_{23} \end{aligned} \quad (\text{C.65})$$

Podemos notar que a distribuição de espaço de fase:

$$\frac{d\mathcal{R}_3}{ds_{12} ds_{23}} = \frac{\pi^2}{4s}$$

é constante para s fixo. Desta forma se os dados de um experimento de decaimento forem “plotados” no plano $s_{12} \times s_{23}$ a densidade de pontos será proporcional ao modulo ao quadrado da amplitude invariante. Ou, dito de outra maneira, a não uniformidade no “plot” fornece informação sobre a dinâmica do processo. Por exemplo, no caso do decaimento $D \rightarrow K\pi\pi$, o aparecimento de bandas quando $m_{(K\pi)} = m_{K^*}$, reflete o aparecimento do canal de decaimento $D \rightarrow K^*\pi \rightarrow K\pi\pi$. Esta distribuição é chamada de *Dalitz Plot*.

A região física do Dalitz Plot pode ser determinada considerando s_{12} para o caso em que \vec{p}_1 é paralelo ou antiparalelo a \vec{p}_1 , ou seja:

$$\begin{aligned} s_{12}^\pm &= m_1^2 + m_2^2 + 2(E_1^r E_2^r \pm |\vec{p}_1^r| |\vec{p}_2^r|) \\ &= m_1^2 + m_2^2 + \frac{1}{2s_{23}} [(M^2 - s_{23} - m_1^2)(s_{23} + m_2^2 - m_3^2) \\ &\quad \pm \lambda^{1/2}(M^2, s_{23}, m_1^2) \lambda^{1/2}(s_{23}, m_2^2, m_3^2)] \end{aligned} \quad (\text{C.66})$$

No caso em que apenas m_2 é diferente de zero a Eq.(C.66) impõe os limites:

$$\frac{M^2 m_2^2}{s_{23}} \leq s_{12} \leq M^2 + m_2^2 - s_{23}$$

e, no caso em que todas as massas são nulas, temos:

$$0 \leq s_{12} \leq M^2 - s_{23}$$

As igualdades determinam as fronteiras do Dalitz Plot. No primeiro caso temos:

$$\begin{aligned} s_{23} &= \frac{M^2 m_2^2}{s_{12}} \\ s_{23} &= M^2 + m_2^2 - s_{12} \end{aligned} \tag{C.67}$$

e para o caso de todas as massas nulas:

$$\begin{aligned} s_{12} &= 0 \\ s_{23} &= 0 \\ s_{23} &= M^2 - s_{12} \end{aligned} \tag{C.68}$$

Bibliografia

- [1] P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences* (McGraw–Hill Book Company, N.Y. USA, 1969);
- [2] L. Lyons, *Statistics for Nuclear and Particle Physics* (Cambridge University Press, N.Y. USA, 1986).
- [3] G. Cowan, *Statistical Data Analysis* (Clarendon Press — Oxford, 1998).
- [4] F. James, *Monte Carlo Phase Space*, CERN/Yellow report 68-15 (1968).
- [5] E. Byckling and K. Kajantie, *Particle Kinematics* (John Wiley & Sons, 1973).
- [6] R. D. Cousins *et al.*, *Nucl. Instr. Meth.* **A320** (1992) 330.
- [7] O. Helene, *Nucl. Instr. Meth.* **A212** (1983) 319; *idem*, *Nucl. Instr. Meth.* **A228** (1984) 120; G. Zech, *Nucl. Instr. Meth.* **A277** (1989) 608.
- [8] A. C. Melissinos, *Experiments in Modern Physics* (Academic Press Inc. N.Y. USA, 1966);
Ver Capítulo 10: Elementos da Teoria Estatística.
- [9] S. I. Bityukov e N. V. Krasnikov, *New Physics Discovery Potential in Future Experiments*, hep-ph/9811025 (1998).
Ver também referências neste artigo.
- [10] N. Brown, *Z. Phys.* **C49** (1991) 657
- [11] L.D. Landau e E.M. Lifshitz, *The Classical Theory of Fields* (Pergamon Press, 1975)
- [12] J.D. Bjorken e S.D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, 1964)
- [13] Particle Data Group, *Review of Particle Properties*,

Índice Remissivo

- Ajustes
 - Curvas, 29
 - Método
 - da Máxima Verossimilhança, 32
 - dos Mínimos Quadrados, 34
 - dos Momentos, 31
- Combinatória
 - Arranjo, 22
 - Combinação, 23
 - Permutação, 22
- Confidence Level (CL), 56
- Correlação
 - Coefficiente, 8
 - Eficiência, 16
- Covariância, 8
- Desvio Padrão , 8
- Distribuição
 - Binomial, 21, 23
 - Média, 23
 - Variância, 23
 - Breit-Wigner, 26
 - Cauchy, 26
 - Cumulativa, 6
 - de Probabilidades, 4
 - Gaussiana, 9, 25
 - Média, 10
 - Variância, 10
 - Lorentziana, 26
 - Normal, *veja* Gaussiana, 25
 - Original, 4
 - Poisson, 24
 - Média, 24
 - Variância, 24
- Erro
 - Sistemático, 71
- Erro na Média, 15
- Erros
 - Aleatórios, 2
 - Estatísticos, 2
 - Medidas Diretas, 14
 - Medidas Indiretas, 11
 - Propagação, 11
 - Sistemáticos, 2
- Espaço Amostral, 2
- Estatística
 - Bayesiana, 3
 - Clássica, 3
- Incerteza na Média, 15
- Incertezas, *veja* Erros, 2
- Média, 5
- Método
 - da Máxima Verossimilhança, 32
 - dos Mínimos Quadrados, 34
 - dos Momentos, 31
 - Monte Carlo, *veja* Monte Carlo, 37
- Matriz
 - Covariância, 8
 - Erro, 8
 - Mediana, 6

- Moda, 6
- Monte Carlo, 37
 - Aplicações, 41
 - Espaço de Fase, 42
 - Momentum transversal, 41
 - Erro Estatístico, 47
 - Geradores Aleatórios, 50
 - Integração, 40
 - Precisão, 39
 - Utilidade, 39
 - Vantagem e Desvantagem, 54
- Nível de Confiança, 56
 - Método Integral, 57
- Números Aleatórios, 39
- Parâmetros, 5
 - Estimação, 14
- Probabilidade, 2
 - Condiciona, 2
 - Função Densidade, 4
- Propagação de Erros, 11
 - Adição, 13
 - Divisão, 13
 - Multiplicação, 13
 - Subtração, 13
- Significância, 65
- Teste de χ^2 , 36
- Valor Esperado, 7
- Valor Médio, 7
- Variável Aleatória, 2
- Variância, 8
- Verossimilhança, 4