

PLANEJAMENTO E ANÁLISE ESTATÍSTICA  
DE EXPERIMENTOS EM FÍSICA



Otaviano A.M. Helene

Tese apresentada ao Instituto  
de Física da Universidade de  
São Paulo para o concurso de  
Livre-Docência junto ao Depar-  
tamento de Física Experimen-  
tal.

SBI-IFUSP



305M810T1441

SÃO PAULO

1988

Defesa 26/08/88

## AGRADECIMENTOS

Esta tese não é -- como eu imagino que não devam ser os outros trabalhos de Livre-Docência -- um trabalho estanque. Não é estanque por seu conteúdo que, mesmo que não extrapolasse a física, não estaria restrito a um único detalhe desta. Não é também estanque no tempo pois muitas das idéias exploradas aqui têm origem remota e desconhecida, e nem mesmo posso saber se sou eu o autor delas, e não se esgotarão, espero, neste trabalho ou em uma ou outra publicação que possa dele sair. Não é estanque também no espaço, pois muitas das questões discutidas têm origem e fim fora do Departamento de Física Experimental. Assim, como poderia ser ela estanque na autoria?

Todo o leitor crítico de manuscritos, como o professor Iuda (Gldman), cujos dedos seguiram letra a letra as  $10^5$  letras deste trabalho, discutindo inúmeros pontos e fazendo várias sugestões indispensáveis, perceberia a participação intelectual do Vito (Vanin) em muitas das passagens, mesmo tendo discutido com ele apenas um ou outro aspecto deste trabalho.

Se não fossem os dedos do Max (Cohenca), do Paulo (Pascholati) e do Alfredo (Bonini), não adiantaria nada

tentar apertar qualquer tecla ou botão da sala experimental e de computação do Acelerador Linear. Nada funcionaria e este trabalho não poderia sequer ser iniciado.

A participação do Philippe (Gouffon) é ainda maior por ter ajudado na parte de computação da mesma forma que participou de inúmeras discussões sobre questões de estatística em física experimental.

E de novo o Paulo, que sugeriu a realização de um experimento desafiante, a do duplo gama, que aceitei imediatamente e, como uma espécie de atravessador da ciência, passei-o ao Eduardo (Mucciolo). Que o mediu!

Este trabalho teria encaixado há alguns anos se o Bruno (Escoubes) e a Salomé (de Unamuno), do grupo de altas energias do Centre de Recherches Nucléaires de Estrasburgo e com atividades experimentais no CERN, não tivessem aceito com grande interesse a proposta de realizar inúmeros experimentos com poucos eventos para testar, na prática, diferentes visões filosóficas da estatística.

Parte dele também não seria jamais feito sem a oportunidade de conhecer em detalhe o experimento do duplo decaimento beta da colaboração Bordeaux-Zaragoza e de colaborar na interpretação estatística dos resultados. Isso foi possível graças ao convite do professor Pierre Menrath

de Bordeaux e a acolhida atenciosa por parte de seu grupo, especialmente de Philippe Hubert.

Mas é claro que nada disso teria sido feito se os dedos da Lourdes (Morais) não tivessem digitado aquelas  $10^5$  letras examinada, uma a uma, pelo Iuda.

E houve um dedo a atrapalhar. Um dedo duro incapaz de fazer uma revisão mínima das bobagens que podem sair de uma listagem de computador. A escrita deste trabalho foi interrompida por duas semanas, ocupadas em reuniões, assembléias e redação de documentos em defesa desta Universidade e de 1100 colegas, vários deles co-autores deste trabalho, apesar disso acusados de improdutivos.

Enfim, muitos dedos contribuíram para a realização deste trabalho. Vários colegas, ligados ou não à infraestrutura do Laboratório do Acelerador Linear, estudantes e bolsistas não foram citados nominalmente aqui por injusta economia de espaço. Se há algo proveitoso, todos eles são responsáveis por isso. Mas todo e qualquer erro só pode ter sido produzido pelo autor.

ÍNDICE

	Página
Resumo.....	v
Abstract.....	vii
1. INTRODUÇÃO.....	01
1.1 - POUCOS EVENTOS.....	02
1.2 - PLANEJAMENTO DE EXPERIMENTOS.....	06
1.3 - ORGANIZAÇÃO DO TRABALHO.....	09
2. INFERÊNCIA EM FÍSICA EXPERIMENTAL E TEORIA ESTATÍSTICA.....	13
2.1 - ESTIMATIVA DE PONTO.....	14
a) Consistência.....	15
b) Não tendenciosidade.....	17
c) Mínima variância.....	18
d) Eficiência.....	20
e) Suficiência.....	21
f) Conclusão.....	23

2.2 - ESTIMATIVA DE INTERVALO.....	24
a) Abordagem clássica: intervalos de confiança.....	25
b) Abordagens bayesianas e fiducialista.....	28
2.3 - INFORMAÇÃO EM FÍSICA EXPERIMENTAL.....	29
3. EXPERIMENTOS COM POUCOS EVENTOS: ESTIMATIVA DE INTERVALOS.....	32
3.1 - TRABALHOS RECENTES.....	33
3.2 - INTERVALOS DE CONFIANÇA.....	37
a) Aproximação gaussiana.....	40
b) Clássico.....	41
c) Bayesianos.....	43
c1) Função a priori constante.....	45
c2) Função a priori divergente.....	46
c3) Funções a priori $a^{-1/2}$ .....	47
3.3 - DIAGNÓSTICO.....	48
a) Aproximação gaussiana.....	51
b) Clássico.....	53
c) Bayesiano uniforme mínimo.....	55
d) Bayesiano uniforme central.....	58

e) Bayesianos divergentes: central e mínimo.....	58
f) Bayesiano a priori $a^{-1/2}$ central.....	61
g) Bayesiano a priori $a^{-1/2}$ mínimo.....	63
3.4 - CONCLUSÃO.....	63
4. PLANEJAMENTO E ANÁLISE DE EXPERIMENTOS.....	71
4.1 - USO DE VÍNCULOS E COVARIÂNCIAS NA ANÁLISE DE DADOS.....	72
a) Uso de vínculos.....	73
b) Uso de covariâncias.....	80
4.2 - PLANEJAMENTO DE UM EXPERIMENTO.....	84
4.3 - ALTERAÇÃO DE CALIBRAÇÃO EM ESPECTROS MULTICANAIS.....	91
a) Alteração de calibração.....	92
b) Alteração de calibração de espec- tros: o duplo decaimento beta do $^{76}\text{Ge}$ .....	98
4.4 - QUANTIDADE DE INFORMAÇÃO EM DIFE- RENTES ARRANJOS EXPERIMENTAIS.....	103

5. CONCLUSÃO.....	116
APÊNDICE A.....	122
A.1 - INTERVALOS DE CONFIANÇA.....	122
a) Distribuição de Poisson.....	122
b) Razão entre pequenos números.....	123
b1) Abordagem clássica.....	123
b2) Abordagem Bayesiana.....	126
A.2 - TESTE EXPERIMENTAL.....	127
APÊNDICE B.....	132
B.1 - ANÁLISE DO ESPECTRO BIDIMENSIONAL.....	134
B.2 - ANÁLISE DA PROJEÇÃO NOS DETECTORES	
HPGe.....	141
APÊNDICE C.....	145
REFERÊNCIAS.....	150



R E S U M O

Este trabalho discute alguns aspectos estatísticos relacionados à análise e planejamento de experimentos em física. Os casos analisados correspondem a experimentos com poucos eventos bem como a experimentos com muitos eventos. O primeiro desses casos ocorre quando o número de eventos  $n$  é tal que  $n = \sqrt{n}$ , e/ou quando o número de eventos é comparável à flutuação estatística do ruído de fundo. São considerados experimentos com muitos eventos aqueles em que o desvio padrão é comparável com eventuais erros sistemáticos.

É discutida a questão da definição de intervalos de confiança no caso de eventos que obedecem a distribuições de Poisson, sendo que entre as abordagens estatísticas estudadas a bayesiana com função densidade de probabilidade *a priori* uniforme é aquela que se mostrou em melhor acordo com o que se espera de um resultado em física experimental.

No caso de experimentos com muitos eventos é mostrado que a explicitação de vínculos físicos entre os parâmetros ajustados pode contribuir para reduzir as consequências de eventuais erros sistemáticos. Ainda neste caso é mostrada a importância da preservação das matrizes de covariância em sua totalidade tanto para que o conteúdo probabilístico dos intervalos sejam conservados como para atualização dos parâmetros ajustados quando algum ou alguns deles forem medidos em outro experimento. Esta situação é

ilustrada com resultado recente de espectroscopia gama do  $^{133}\text{Ba}$ .

A questão do estabelecimento de janelas em espectros multicanais com o fim de discriminar eventos de origens diferentes é estudada com base na quantidade de informação contida em dados experimentais. Neste caso mostra-se que a largura dessas janelas depende não apenas da forma do sinal correspondente aos eventos discriminados mas também da relação entre sua intensidade e a intensidade do ruído de fundo.

O conceito de quantidade de informação é usado também na comparação de diferentes arranjos experimentais para o estudo do duplo decaimento beta do  $^{76}\text{Ge}$  com alimentação do nível  $2^+$  do  $^{76}\text{Se}$ . Neste caso a evidência do decaimento é dada pela medida simultânea de duas grandezas diferentes. São comparadas duas maneiras de quantificar a informação nesse experimento sendo que ambas mostraram o mesmo comportamento. Discute-se ainda a questão da mudança de calibração de um espectro multicanal sem o comprometimento da flutuação estatística e, portanto, sem impossibilitar a realização de testes estatísticos.

A B S T R A C T

This work focusses some statistical aspects of the design and analysis of experiments in physics. Experiments with small and large numbers of events are analyzed. Small number of events occurs whenever  $n \cong \sqrt{n}$  and/or the number of events are close to the background noise. Experiment with large number of events is the one with the standard deviation comparable to sistematic errors.

This work discusses the definition of confidence intervals for events that follow Poisson distributions; among all statistical approaches studied the Bayesian one with uniform prior distribution is in better agreement to the experimental physics expectative.

For experiments with large number of events it is shown that the use of constraints among adjusted parameters is useful to reduce sistematic errors. It is also shown that the use of the whole covariance matrix is important in order to preserve the probabilistic contents of intervals as well as to calculate new values for the adjusted parameters when some of them are measured in a new and independent experiment. These points are stressed with a recent experiment in gamma-ray spectroscopy of  $^{133}\text{Ba}$ .

The choice of windows in multichannel spectra to discriminate different events is discussed considering the amount of information contained in the experimental data. It is shown that the width of the windows depends on both the peak shape and the signal to background ratio.

The amount of information is also used to compare different experimental designs for the study of the double beta decay of  $^{76}\text{Ge}$  to the  $2^+$  level of  $^{76}\text{Se}$ . In this case the decay is evidenced by the simultaneous measurement of two different quantities.

Finally, we discuss the calibration change of multichannel spectra without changing the statistical fluctuations and therefore allowing the accomplishment of statistical tests.

## 1. INTRODUÇÃO

Existe hoje uma quantidade significativa de experimentos fundamentais para o desenvolvimento da física que são caracterizados pela pequena quantidade de dados (ou eventos), seja em termos absolutos, seja em relação à flutuação estatística de um ruído de fundo que tende a dissimulá-los. E não há muito o que fazer para alterar essa situação. Os eventos nesses experimentos extremos vão surgindo lentamente e não podemos aguardar que os anos se passem até concluirmos algo sobre o fenômeno estudado.

Como exemplos dessa situação podemos citar o decaimento do próton, o duplo decaimento beta com ou sem neutrino, a procura de monopolos magnéticos, os neutrinos solares, modos de decaimento de partículas elementares entre outros. Além dessas há outras situações nas quais o número de eventos é inevitavelmente pequeno, como a observação recente de supernova SN1987A: apenas 19 eventos de neutrino foram observados de forma não dúbia. Qualquer conclusão que se pretenda, sobre mecanismos de explosão de supernovas,

meia vida e massa dos neutrinos, ela dependerá da análise desses únicos 19 eventos.

Em situações experimentais usuais é sempre importante o adequado planejamento do experimento bem como o correto tratamento estatístico dos dados. Não fazê-los implica em nova realização do experimento ou no prolongamento da sua duração, coisas incômodas mas não impossíveis. No entanto no caso de experimentos extremos essas soluções não são sempre possíveis. Um investimento inicial no planejamento bem como um tratamento correto dos dados -- e ambos estão inter-relacionados -- são fundamentais.

No presente trabalho pretendemos apresentar duas contribuições distintas ao tratamento e à interpretação estatística de dados como eles ocorrem em física, ambas adequadas à situações experimentais extremas.

### 1.1 - POUCOS EVENTOS

Há em estatística algumas abordagens diferentes para a estimativa de intervalos de confiança calculados a partir da observação de uma quantidade de eventos que obedece a uma determinada distribuição de probabilidade. Essas abordagens apresentam diferentes fundamentos e

diferentes hipóteses e como regra geral satisfazem diferentes expectativas quanto aos resultados. Embora em muitos casos os resultados fornecidos por essas diferentes abordagens sejam iguais, há casos em que isso não ocorre. No presente trabalho discutiremos os intervalos de confiança calculados com base em algumas dessas possíveis abordagens. Nesta Introdução apresentaremos de forma geral a questão básica que pretendemos discutir, por meio de um exemplo.

Vamos supor que determinada hipótese ou modelo que se pretende testar preveja para certa razão de ramificação do decaimento de determinada partícula o valor  $R_{N/D} = 0,050$ . É feito então um experimento onde se observa  $N = 1$  evento e  $D = 3$  eventos para os dois modos de decaimento. A estimativa é então de uma razão de ramificação  $R = 1/3$ . Para se testar a hipótese é entretanto importante indicar os intervalos de confiança associados a esse resultado. Caso seja seguida a abordagem clássica, uma das diferentes abordagens possíveis para o tratamento do problema, seriam indicados os seguintes intervalos para  $R$  (JA80):

$$\begin{aligned} (0,044 < R < 1,62): P &= 68\% \\ (0,013 < R < 3,02): P &= 90\%. \end{aligned} \tag{1.1}$$

As probabilidades indicadas à direita indicam os níveis de confiança associados aos intervalos. Olhando apenas para os intervalos parece claro que o resultado experimental é compatível com a hipótese  $R_{N/D} = 0,050$ .

No entanto se os intervalos indicados são os centrais e determinados pela abordagem bayesiana com função densidade de probabilidade *a priori* uniforme (o capítulo 2 esclarece as definições desses termos) os intervalos calculados para  $N = 1$  e  $D = 3$  são (HE83a):

$$\begin{aligned} (0,172 < R < 1,10) : P &= 68\% \\ (0,091 < R < 1,92) : P &= 90\% \\ (0,056 < R < 2,56) : P &= 95\% \end{aligned} \tag{1.2}$$

O valor  $R_{N/D} = 0,050$  está fora dos intervalos com níveis de confiança "razoáveis". Alguém que olhasse para esses intervalos tenderia a acreditar pouco na compatibilidade da hipótese com o resultado experimental.

Qual dos dois resultados está "correto"? Os indicados em 1.1 ou em 1.2?

Em um caso extremo como esse é necessário uma análise estatística mais elaborada do que a mera indicação de intervalos. No entanto seria bastante conveniente se pudessemos determinar intervalos (realmente dignos) de



confiança pois a simples indicação deles permitiria uma análise rápida e confiável do resultado.

Embora a situação analisada aqui possa parecer muito artificiosa, há situações reais bastante próximas disso. Um exemplo seria a medida de um único eventual monopolo magnético (CA82) -- que por não ter sido reproduzido posteriormente o evento passou a ser interpretado como consequência de um erro sistemático ou um possível mal funcionamento do equipamento. Os experimentos de decaimento do próton apresentam usualmente, mesmo como fundo, pouquíssimos eventos a partir dos quais é necessário fazer-se inferências necessárias inclusive para o planejamento de novos arranjos ou do prosseguimento dos experimentos.

A explosão da supernova SN1987A observada em fevereiro de 1987 apresenta outra situação real na qual 8 eventos (neutrinos) foram observados no detector da colaboração IMB (BI87) e 11 eventos no detector Kamiokande II (HI87). Como esses eventos não são monoenergéticos, o número de eventos em dada faixa de energia é ainda mais reduzido. E é a partir deles que serão feitas as inferências possíveis.

Neste trabalho discutiremos os intervalos de confiança calculadas a partir de diferentes abordagens

estatísticas comparando o comportamento de cada um deles com o que se espera de um intervalo em física experimental.

Como exemplo ainda de outros experimentos onde não apenas o número de eventos é pequeno mas também comparável à flutuação estatística do fundo podemos citar o estudo do duplo decaimento beta sem neutrino pela colaboração Bordeaux - Zaragoza (M087) ou a duplo decaimento gama do primeiro nível  $0^+$  do  $^{90}\text{Zr}$  estudado no Laboratório do Acelerador Linear do Instituto de Física da USP (MU87b). Esses experimentos são discutidos neste trabalho com algum detalhe.

## 1.2 - PLANEJAMENTO DE EXPERIMENTOS

É conveniente em qualquer experimento que se realize um planejamento anterior cuidadoso. No entanto não fazê-lo, em uma situação trivial, o preço a pagar é uma menor precisão final do resultado, a necessidade da repetição do experimento ou o prolongamento de sua duração. Em uma situação extrema o custo de um mau planejamento pode ser muito maior: a não obtenção de qualquer resultado, a perda de um arranjo experimental caro ou a necessidade de prolongar o experimento por um tempo proibitivamente longo.

Faz parte do planejamento a escolha dos equipamentos bem como definição da estratégia experimental (qual o melhor caminho para se medir a grandeza desejada). Mas para isso é necessário ter em vista claramente o que se quer determinar e como cada fator responsável por um eventual erro influencia o resultado final.

A primeira dessas exigências -- saber o que se quer determinar -- pode não ser tão óbvia como parece. Um exemplo trivial é o uso da mediana e da média como estimativas de uma grandeza cujos dados experimentais obedecem a uma distribuição gaussiana. Ambas são estimativas da grandeza medida. No entanto a média é uma estimativa mais precisa do que a mediana: usando-se a média consegue-se a mesma precisão que se conseguiria com o uso da mediana em um experimento com cerca de 50% a mais de dados.

Outro exemplo menos trivial desenvolvido recentemente (OL86) refere-se a determinação da mistura multipolar E2/M1 para transições eletromagnéticas medida em experimentos de correlação angular. Neste caso houve a necessidade de se estudar entre os diferentes estimadores possíveis qual deles estava em melhor acordo com o que se pretendia medir.

Eadie et al. (EA71) ilustram a necessidade e a importância do planejamento com um exemplo semelhante a

seguinte situação. Deseja-se medir a massa de dois corpos,  $m_a$  e  $m_b$ , usando-se duas vezes uma balança cuja precisão corresponde a uma variância  $\sigma^2$ , independente da massa. Se os dois corpos são pesados isoladamente ambos os resultados apresentarão variâncias  $\sigma^2$ . No entanto se em uma medida os dois corpos forem pesados juntos,  $M_1 = m_a + m_b$ , com variância  $\sigma^2$ , e em outra for medida a diferença de suas massas,  $M_2 = m_a - m_b$ , também com variância  $\sigma^2$ , então

$$m_a = \frac{M_1 + M_2}{2} ; \quad \sigma_a^2 = \sigma^2/2$$

(1.3)

$$m_b = \frac{M_1 - M_2}{2} , \quad \sigma_b^2 = \sigma^2/2$$

onde  $\sigma_a^2$  e  $\sigma_b^2$  indicam as variâncias dos resultados experimentais para  $m_a$  e  $m_b$  respectivamente. Ou seja, com o mesmo esforço e o mesmo aparato experimental consegue-se, como consequência da estratégia adotada, uma precisão final no resultado desejado duas vezes melhor que aquela que seria obtida caso os dois corpos fossem pesados isoladamente.

Este exemplo mostra a importância do planejamento de um experimento. No entanto é necessário dizer que o planejamento só é eficiente se sabemos claramente o que se quer medir.

### 1.3 - ORGANIZAÇÃO DO TRABALHO

A discussão ou resolução dos problemas apresentados neste trabalho dependem de uma interpretação e de um tratamento estatístico rigorosos e adequados às expectativas dos físicos quanto aos resultados de um experimento. Assim no capítulo 2 apresentaremos algumas das idéias e dos princípios básicos da estatística, comparando-os com aquelas expectativas.

É interessante observarmos que a física experimental utiliza menos da teoria estatística na resolução de seus problemas de inferência do que outras ciências, como por exemplo as ciências sociais e da vida (economia, biologia, psicologia entre outras). O próprio desenvolvimento da estatística está ligada a essas ciências na medida em que elas apresentam problemas de inferência ao mesmo tempo que podem criticar, com fundamento em situações concretas, as soluções propostas (veja por exemplo BE82). Esse uso restrito da estatística por parte da física experimental não tem origem na sua pouca adaptação aos problemas de inferência da física mas talvez em um mau hábito originado na pequena necessidade de realização de inferências rigorosas e precisas em física. Como exemplo

desse mau hábito podemos citar o frequente esquecimento das covariâncias na análise de dados experimentais, cuja importância é absolutamente relevante como mostrado por Mannhart (MA81), se se quer preservar o conteúdo probabilístico dos intervalos de confiança indicados.

No entanto o uso rigoroso e criterioso da estatística permite não apenas a preservação dos conteúdos probabilísticos dos intervalos indicados, o que é fundamental, como também o planejamento e a análise de experimentos com o objetivo de maximizar a quantidade de informação por eles gerada.

O conteúdo do capítulo 2 é então justificado não como a apresentação de um resumo da teoria estatística mas apenas como uma comparação entre as nossas expectativas na realização e análise de experimentos e as idéias fundamentais da teoria estatística.

No capítulo 3 discutiremos o problema da indicação de intervalos de confiança no caso da distribuição de Poisson. Neste caso, como em vários outros, é impossível a indicação, a partir de um resultado experimental, de intervalos de confiança com conteúdo probabilístico (nível de confiança) exato. Assim, diferentes abordagens para o problema têm sido apontadas (veja por exemplo os capítulos 20 e 21 da referência KE79). A questão é saber então qual

delas está em melhor acordo com o que se espera como resultado de um experimento em física e, havendo discrepâncias, identificá-las e analisar quão grave elas são.

O conteúdo do capítulo 3 é, de certa forma, continuação de um trabalho anterior (ES87) no qual comparamos intervalos determinados por diferentes abordagens com situações experimentais concretas. No Apêndice A mostramos um resumo daquele trabalho, indicando as principais conclusões.

No capítulo 4 apresentamos a análise e o planejamento de alguns experimentos ou procedimentos experimentais correspondentes a situações extremas. Uma das situações discutidas no capítulo 4 refere-se ao uso de vínculo e covariância entre parâmetros ajustados. O experimento analisado é a medida de transições eletromagnéticas do  $^{133}\text{Ba}$  com o uso de detectores semicondutores (de resoluções da ordem de keV) com precisão menor ou da ordem de 1 eV, recentemente concluída (MU87a). O uso de vínculos permitiu a diminuição dos efeitos dos erros sistemáticos não apenas por ser a energia de cada transição obtida por combinação linear da energia de várias transições como também pela facilidade com que certos artifícios

instrumentais podem ser usados. O uso de covariâncias preservou o conteúdo probabilístico dos intervalos.

Ainda no capítulo 4 discutimos vários aspectos relacionados à quantidade de informação em um experimento. Uma dessas situações refere-se a escolha de "janelas" ideais em espectros multicanais. Procedimentos como esse permitiram a observação clara do duplo decaimento gama no  $^{90}\text{Zr}$  (MU87b). Outra situação analisada nesse capítulo é a comparação entre três arranjos experimentais diferentes para a medida do duplo decaimento beta sem neutrino do  $^{76}\text{Ge}$ : um dos arranjos corresponde a um experimento recentemente concluído pela colaboração Bordeaux - Zaragoza (MO87); os outros dois são possíveis programas de continuação desse experimento.

Discutimos ainda no capítulo 4 o problema da mudança de calibração de um espectro multicanal sem comprometimento da flutuação estatística (de Poisson) da quantidade de eventos em cada canal e, portanto, sem inviabilizar a aplicação de testes estatísticos ao espectro modificado. No Apêndice B ilustramos a alteração de calibração, comparando dois procedimentos diferentes para a sua realização, a alteração determinística e a alteração estatística.



## 2. INFERÊNCIA EM FÍSICA EXPERIMENTAL E TEORIA ESTATÍSTICA

Neste capítulo discutiremos as idéias e os resultados básicos da teoria estatística fazendo uma comparação com o que se espera, em física, de um experimento.

Muitas vezes os físicos experimentais tendem a usar a estatística como um conhecimento acabado e pronto, que ensina como tratar resultados experimentais e que não sofre alterações há muito tempo. Isso certamente não é verdade. Idéias tão simples como o uso da média, no lugar da mediana, como a melhor estimativa do valor de uma grandeza que dá origem a dados distribuídos segundo uma gaussiana, só foram estabelecidas de forma rigorosa e definitiva a partir de trabalhos que datam da década de 20 (KE79, FI50). A opção pelo desvio padrão como sendo a melhor forma de se quantificar a precisão de um experimento, em substituição ao "erro provável", até então usado, é também da mesma época (FI50). Na edição de 1979 do livro "The Advanced Theory of Statistics" de Kendall e Stuart (KE79), cerca de um terço das quase 1500 referências são da década de 70 e muitas

delas não se referem apenas a temas acadêmicos, de interesse exclusivo dos estatísticos, mas tratam de problemas concretos sob o ponto de vista do físico experimental.

No capítulo seguinte nós iremos discutir um problema ao menos parcialmente aberto: qual é a abordagem -- bayesiana ou clássica -- mais adequada à física experimental. Neste capítulo faremos uma revisão dos conceitos e teorias estatísticas mais relevantes à física experimental comparando o que os estatísticos e os físicos esperam de um processo de inferência. Como não pretendemos fazer um resumo da teoria estatística, abriremos mão de todo e qualquer rigorosismo que possa prejudicar nosso objetivo.

## 2.1 - ESTIMATIVA DE PONTO

A menos de casos excepcionais, nos quais a totalidade dos dados brutos devem ser apresentados para que o leitor possa fazer, de forma completa, seu próprio julgamento, é usual reduzir-se os dados de um experimento a apenas dois: um valor que corresponde a melhor estimativa da grandeza medida e um intervalo onde deve estar, com certa probabilidade, o seu valor verdadeiro. Nesta seção discutiremos a questão da estimativa da grandeza, fazendo

uma comparação entre o que se espera em física experimental e as propriedades que devem ter um "bom" estimador em estatística. Na seção seguinte faremos o mesmo com a estimativa de intervalos.

### a) Consistência

A expectativa básica em física é que quanto mais esforço for colocado na medida de uma grandeza, obtendo-se mais e mais dados, mais nos aproximamos do conhecimento do seu valor verdadeiro. Essa expectativa está associada à propriedade de consistência de um estimador. Se  $t_n$  é um estimador calculado a partir de  $n$  dados correspondentes a medida de uma grandeza  $x_0$  então, quaisquer que sejam  $\eta$  e  $\epsilon$  positivos, existe  $N$  tal que

$$P(|t_n - x_0| < \epsilon) > 1 - \eta, \quad n > N, \quad (2.1)$$

onde  $P(|t_n - x_0| < \epsilon)$  é a probabilidade de que a diferença absoluta entre  $t_n$  e  $x_0$  seja menor que  $\epsilon$ . (Em certo sentido a condição 2.1 pode ser comparada à condição de inexistência de erros sistemáticos na análise dos dados.)

Por exemplo na medida de uma grandeza cujos dados experimentais obedecem a uma distribuição gaussiana,

$$f(x_i) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x_i - x_0)^2}{2\sigma^2}}, \quad (2.2)$$

onde  $x_0$  é o valor verdadeiro da grandeza e  $\sigma^2$  a variância, a média de  $n$  dados obtidos,

$$\bar{x} = \frac{1}{n} \sum x_i \quad (2.3)$$

é uma estimativa consistente de  $x_0$ . (Essa propriedade de  $\bar{x}$  não será provada aqui, pois não é objetivo deste trabalho fazer um resumo didático da teoria estatística. A referência (KE79) apresenta os fundamentos da teoria estatística e as demonstrações necessárias à compreensão deles.) No entanto o valor

$$\frac{1}{n-1} \sum x_i = \frac{n}{n-1} \bar{x} \quad (2.4)$$

também é uma estimativa consistente de  $x_0$ . Por que não o usamos?

### b) Não tendenciosidade

Outra propriedade importante do resultado de um experimento é que, qualquer que seja o número de dados obtidos, digamos  $n$ , a repetição múltipla do mesmo experimento com o mesmo número de dados leva a um conjunto de resultados que, se analisados globalmente, tendem ao valor verdadeiro da grandeza medida qualquer que seja  $n$ . Ou, dito de outra forma, o resultado de um experimento não pode ter uma tendência de subestimação ou superestimação da grandeza medida.

Em teoria estatística essa é a propriedade de não tendenciosidade. Se  $t_n$  é o estimador de uma grandeza  $x_0$  a partir de  $n$  dados, então

$$E(t_n) = x_0, \quad (2.5)$$

onde  $E(t_n)$  indica o valor esperado de  $t_n$  calculado a partir de sua função densidade de probabilidade. Isso explica -- do ponto de vista da estatística -- porque não usamos a estimativa da equação 2.4:

$$E\left(\frac{n}{n-1} \bar{x}\right) = \frac{n}{n-1} E(\bar{x}) = \frac{n}{n-1} x_0. \quad (2.6)$$

Se  $n$  é finito a equação 2.4 corresponde a uma superestimação, em média, da grandeza medida.

Em física experimental poder-se-ia dizer que a estimativa 2.4 corresponde à introdução de um erro sistemático na análise dos dados.

### c) Mínima variância

A raiz quadrada da variância de um estimador é frequentemente chamado em física -- especialmente no caso de distribuições normais -- do "erro" da estimativa. É claro portanto que dado um conjunto de dados procure-se aquela combinação deles que corresponda a uma estimativa boa (sem erros sistemáticos provenientes da análise de dados, ou seja, consistente e não tendenciosa) da grandeza medida com menor "erro". No caso da distribuição gaussiana a média e a mediana são boas estimativas. No entanto, a mediana tem uma variância maior que a média, cuja relação é

$$\frac{V(\text{mediana})}{V(\text{média})} = \frac{\pi}{2} \quad (2.7)$$

quando o número de dados é suficientemente grande. É portanto inadequado o uso da mediana como o estimador da grandeza medida: é necessário maior esforço para conseguir a mesma qualidade no resultado final que se conseguiria com o uso da média.

Há dois resultados muito úteis para a física experimental sobre a menor variância possível de um estimador, que vamos mostrar de forma simplificada.

Seja  $F(x|x_0)$  a função densidade de probabilidade (ou, no caso de resultados discretos, a probabilidade) de  $x$  em um experimento cujo valor verdadeiro é  $x_0$ . A função verossimilhança de um conjunto de dados  $x_1, x_2, \dots, x_n$  é

$$L(x_1, x_2, \dots, x_n | x_0) = f(x_1 | x_0) f(x_2 | x_0) \dots f(x_n | x_0). \quad (2.8)$$

(A barra vertical representa a função probabilidade das variáveis à sua esquerda, dados os parâmetros à direita.) É possível mostrar que um estimador não tendencioso  $t$  de  $x_0$  apresenta uma variância  $V(t)$  tal que

$$V(t) \geq \frac{-1}{E(\partial^2 \ln L / \partial x_0^2)} \quad (2.9)$$

onde  $E$  representa o valor esperado calculado a partir da função densidade de probabilidade dos dados.

O segundo resultado é que a eq. 2.9 se torna uma igualdade quando

$$\frac{\partial \ln L}{\partial x_0} = A(x_0) (t - x_0) \quad (2.10)$$

sendo  $t$  neste caso um estimador de mínima variância. As demonstrações desses dois resultados podem ser vistas em Ke79 e EA71.

A quantidade  $I$  dada por

$$I = - E \left( \frac{\partial^2 \ln L}{\partial x_0^2} \right) \quad (2.11)$$

é algumas vezes chamada de quantidade de informação contida nos dados. Quando o estimador é tal que a equação 2.9 é uma igualdade  $I$  é o inverso da sua variância.

#### d) Eficiência

Nem sempre existe um estimador de mínima variância. Uma definição em estatística que corresponde,



grosso modo, ao custo que temos que pagar pela inexistência de tal estimador é a eficiência (KE79). A eficiência de um estimador pode ser definida como a razão entre sua variância e a variância de um estimador de mínima variância quando o número de dados tende a infinito:

$$\xi = \lim_{n \rightarrow \infty} \frac{V(t_1)}{V(t_2)} . \quad (2.12)$$

Nessa expressão  $t_2$  é um estimador de mínima variância e  $V$  indica as variâncias dos estimadores. Por exemplo, a mediana é um estimador de eficiência  $2/\pi$ .

#### e) Suficiência

Um estimador é em geral uma função dos dados experimentais. Ele contém portanto informações contidas no conjunto dos dados. Uma tal função é chamada suficiente se ela contém todas as informações contidas originalmente nos dados. Ou seja, conhecendo-a se torna desnecessário conhecer os dados individualmente. Um caso familiar ao físico experimental é o da média em experimentos cujos dados são distribuídos normalmente e a variância conhecida. Neste caso a média contém todas as informações contidas nos dados. Após

seu cálculo os dados podem ser simplesmente jogados fora sem que com isso perca-se qualquer informação.

É interessante conectarmos a idéia de suficiência com a definição de informação da equação 2.11. Numa distribuição normal a variância dos dados é  $\sigma^2$  e a da média  $\sigma^2/n$ , onde  $n$  é a quantidade de dados usados no cálculo da média. A informação contida em cada dado é  $1/\sigma^2$  e na média é  $n/\sigma^2$ , pois tanto a média como os dados são estimativas para as quais 2.9 é uma igualdade. Assim, vemos que a quantidade de informação da média é igual a soma das quantidades de informação em cada um dos  $n$  dados.

Formalmente a suficiência pode ser estabelecida nos seguintes termos. Se  $f(x_1, x_2, \dots, x_n | x_0)$  é a função densidade de probabilidade dos dados e  $t = t(x_1, x_2, \dots, x_n)$  é uma função dos dados,  $t$  é dito suficiente se

$$f(x_1, x_2, \dots, x_n | x_0) = g(t | x_0) h(x_1, x_2, \dots, x_n | t) \quad (2.13)$$

No caso de uma distribuição gaussiana temos

$$\frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\sum \frac{(x_i - x_0)^2}{2\sigma^2}} = \frac{e^{-\frac{(\bar{x} - x_0)^2 n}{2\sigma^2}}}{(\sqrt{2\pi}\sigma)^n} e^{-\sum \frac{(x_i - \bar{x})^2}{2\sigma^2}} \quad (2.14)$$

onde o termo à esquerda é a função densidade de probabilidade dos dados e a segunda exponencial à direita depende apenas dos dados e da média e não do valor verdadeiro da grandeza, que só aparece na primeira exponencial, juntamente com a média. Como é a função densidade de probabilidade que nos permite fazer inferências sobre a grandeza medida, tendo conhecimento da média o conhecimento dos dados é desnecessário no que diz respeito a informação que podemos ter sobre  $x_0$ .

É a propriedade de suficiência que fundamenta o que chamamos usualmente em física experimental de redução de dados.

#### f) Conclusão

Um problema básico da estatística que não abordaremos aqui é o de procurar estimadores e/ou métodos de estimação que estejam de acordo com as propriedades acima ou mostrar que isso pode ser impossível em alguns casos. Os métodos da máxima verossimilhança (ou da máxima probabilidade como é talvez mais conhecido em física experimental (HE80)), dos mínimos quadrados ou o bayesiano são respostas àquela procura. No entanto não abordaremos

aqui esse ponto por não ser ele fundamental para o objetivo do presente trabalho.

Os exemplos dados nesta seção corresponderam a medidas de grandezas cujos dados eram distribuídos normalmente. No entanto, tudo o que foi dito pode ser estendido para outras distribuições, inclusive para aquelas que geram eventos discretos.

## 2.2 - ESTIMATIVA DE INTERVALO

Quando apresentamos o resultado de um experimento indicamos a melhor estimativa da grandeza medida bem como um intervalo que contém o valor verdadeiro da grandeza com certa probabilidade. A indicação do intervalo é tão importante quanto a indicação da grandeza pois é ele que permite comparar experimentos diferentes, bem como incorporar um resultado experimental a teorias ou modelos.

O usual em física experimental é indicar-se um intervalo com conteúdo probabilístico de 68,3%, que corresponde a um desvio padrão no caso de distribuições gaussianas. Quando necessário são indicados intervalos com conteúdos probabilísticos diferentes que são, nesses casos, explicitados.

Ao indicar um intervalo com certo conteúdo probabilístico espera-se duas coisas: (i) a cada vez que se indica um intervalo há um chance igual ao conteúdo probabilístico de que o valor verdadeiro da grandeza esteja de fato no intervalo e (ii) a longo termo, ou seja, ao repetir-se múltiplas vezes a indicação de intervalos, a fração de vezes que os valores verdadeiros das grandezas estão de fato no intervalo tende ao conteúdo probabilístico indicado. Apesar dessas duas expectativas serem extremamente importantes, nem sempre é possível satisfazê-las. E isso ocorre de forma mais marcante em experimentos com poucos eventos, usualmente ligados a estudos de fenômenos cujo conhecimento é fundamental para os próximos passos no desenvolvimento de teorias e modelos da física atual. Um dos objetivos deste trabalho é procurar entre as várias opções indicadas pela teoria estatística aquela que é mais adequada à física, ou seja, que está em melhor acordo com as expectativas (i) e (ii) acima.

#### **a) Abordagem clássica: intervalos de confiança**

A maneira talvez mais sólida de determinação de intervalos é baseada unicamente na teoria da probabilidade e independente de qualquer argumento que possa ser

caracterizado como doutrinário. Por exemplo, a média calculada a partir de  $n$  pontos experimentais que obedecem a uma distribuição gaussiana centrada em  $x_0$  e com variância  $\sigma^2$  obedece a função densidade de probabilidade, também gaussiana,

$$f(\bar{x}) = \frac{e^{-\frac{n(\bar{x} - x_0)^2}{2\sigma^2}}}{\sqrt{2\pi} \sigma/\sqrt{n}} \quad (2.15)$$

Neste caso para qualquer  $\alpha < 1$ , é possível determinar  $\lambda$  tal que a probabilidade que  $\bar{x} \in (x_0 \pm \lambda\sigma/\sqrt{n})$  seja igual a  $1-\alpha$ ,

$$\int_{x_0 - \lambda\sigma/\sqrt{n}}^{x_0 + \lambda\sigma/\sqrt{n}} f(\bar{x}) d\bar{x} = 1-\alpha \quad (2.16)$$

desde que conheçamos  $\sigma$  mas sem que precisemos conhecer  $x_0$ . A condição acima é equivalente à condição  $x_0 \in (\bar{x} \pm \lambda\sigma/\sqrt{n})$ . Portanto ao indicarmos em um experimento o intervalo  $(\bar{x} - \lambda\sigma/\sqrt{n}, \bar{x} + \lambda\sigma/\sqrt{n})$  haverá uma chance igual à  $(1-\alpha)$  100% que o valor  $x_0$  esteja contido no intervalo indicado. Da mesma forma, ao fazermos múltiplas vezes a indicação de tal intervalo em uma fração  $1-\alpha$  delas o valor verdadeiro da

grandeza estará contida no intervalo. Assim podemos indicar, no caso de distribuições normais com variância conhecida, intervalos com conteúdos probabilísticos de acordo com as expectativas do físico experimental.

No entanto embora isso seja possível para a distribuição normal, bem como para muitas outras, não o é em geral. Por exemplo, no caso de medidas de grandezas que obedecem a distribuições binomiais ou de Poisson não é possível determinar-se intervalos com conteúdos probabilísticos (no sentido (i) e (ii) citados) exatos. Nesses casos só seria possível a indicação de intervalos com conteúdos probabilísticos maiores ou iguais a  $1-\alpha$ . Ou seja, poderíamos dizer que há uma chance maior ou igual à  $1-\alpha$  de que o valor verdadeiro esteja contido em certo intervalo; ou, repetindo-se a longo termo o procedimento, o valor verdadeiro da grandeza estará no intervalo indicado uma fração de vezes maior ou igual a  $1-\alpha$ . As referências KE79 e EA71 discutem a técnica de construção desses intervalos.

Em certas aplicações, especialmente quando o preço a pagar pelo fato do valor verdadeiro da grandeza não estar no intervalo indicado for muito alto, como pode ocorrer em decisões que envolvam saúde humana por exemplo, talvez não seja incômodo a superestimação de seus conteúdos probabilísticos. No entanto em outras aplicações, como em

física experimental, certamente isso é incômodo. Para driblar ao menos parcialmente esse problema alguns autores (ST50, PR85a) têm proposto esquemas artificiosos. No entanto esses esquemas parecem inadequados à física pois levam a intervalos diferentes para o mesmo resultado experimental (ST50) ou a resolução de apenas parte do problema colocado pela física experimental (PR85a).

#### **b) Abordagens bayesiana e fiducialista**

Outras abordagens do problema de determinação de intervalos fundamentam-se na construção de uma "função densidade de probabilidade do grau de confiança que temos nos diversos valores possíveis para a grandeza desconhecida" após a medida (KE79). Em certas circunstâncias (KE79) e sob o ponto de vista do físico experimental, ambas as abordagens baseadas na existência dessa "função de confiança", a bayesiana e a fiducialista, são equivalentes. Nos dois casos obteríamos uma distribuição para o valor verdadeiro da grandeza medida que depende do resultado experimental observado.

Essas abordagens têm sido usadas explicita ou implicitamente por físicos experimentais (veja por exemplo RA47 e AN53, entre outros). A abordagem bayesiana



particularmente é bastante confortável uma vez que informações a priori sobre a grandeza medida -- vínculos ou resultados anteriores -- são facilmente incorporadas na análise dos dados.

No capítulo seguinte faremos uma revisão das abordagens mais comumente usadas por físicos bem como um diagnóstico da consequência dessas abordagens.

### 2.3 - INFORMAÇÃO EM FÍSICA EXPERIMENTAL

A definição 2.11 de quantidade de informação é bastante útil em física experimental, especialmente por permitir a comparação entre diferentes arranjos e procedimentos experimentais. No capítulo 4 discutiremos alguns experimentos extremos onde a identificação do fenômeno em estudo é feita a partir da medida de uma ou mais grandezas usando explicitamente o conceito de informação. É interessante portanto atentarmos para algumas propriedades da quantidade de informação no caso da existência de um bom estimador.

A quantidade de informação é, no caso de um estimador de mínima variância, dada pelo inverso de sua variância,

$$I = \frac{1}{\sigma^2} . \quad (2.17)$$

No caso da realização de dois experimentos independentes para a medida de uma mesma grandeza cujos resultados (não covariantes) apresentam variâncias  $\sigma_1^2$  e  $\sigma_2^2$ , a quantidade de informação total é

$$I = I_1 + I_2 = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} , \quad (2.18)$$

dada a propriedade de aditividade da informação. Esse resultado está de acordo com a expectativa em física experimental onde, por propagação de erros, a variância  $\sigma_t^2$  total é dada por

$$\frac{1}{\sigma_t^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \quad (2.19)$$

Essa propriedade de aditividade -- bem como a propriedade de aditividade das quantidades de informação contidas nos dados para compor a quantidade de informação contida na média no caso de distribuições gaussianas, discutida na seção 2.1.e -- torna a quantidade de informação

uma grandeza bastante útil no planejamento de experimentos: a quantidade de informação cresce (de forma aproximadamente linear) com o esforço dispendido na realização de um experimento. Assim dois arranjos experimentais diferentes podem ser comparados entre si por meio da avaliação da quantidade de informação contida em cada um deles em experimentos de igual duração. Na seção 4.4 discutiremos a quantidade de informação contida nos dados de um experimento correspondente a identificação de um fenômeno que é caracterizado pela medida de duas grandezas diferentes.

### 3. EXPERIMENTOS COM POUCOS EVENTOS: ESTIMATIVA DE INTERVALOS

Quando a partir de um resultado experimental calcula-se um intervalo de confiança  $(a_-, a_+)$  com determinado conteúdo probabilístico  $1-\alpha$ , espera-se que (i) haja  $(1-\alpha)$  100% de chance que o valor verdadeiro da grandeza medida esteja contido no intervalo e (ii) a longo termo -- repetindo-se inúmeras vezes a indicação de intervalos com o mesmo conteúdo probabilístico -- uma fração  $1-\alpha$  das vezes os valores verdadeiros das grandezas medidas estarão contidos nos intervalos indicados. Como discutido no capítulo anterior essas duas expectativas são satisfeitas no caso de medidas cujos dados obedecem a distribuições gaussianas, assim como em muitas outras situações. No entanto em alguns casos -- entre eles situações nas quais os dados experimentais obedecem a distribuições discretas, como a de Poisson por exemplo, discutida neste capítulo -- isso não ocorre, sendo a discrepância em relação às expectativas acima tão mais marcante quanto menor o número de dados observados. Neste capítulo é feita uma revisão das soluções propostas para essas situações e as implicações de cada uma

delas. É sugerido, a partir da análise dessas várias soluções, um conjunto de intervalos -- e um procedimento estatístico -- que está mais próximo das expectativas (i) e (ii) acima para o caso da distribuição de Poisson.

### 3.1 - TRABALHOS RECENTES

Em 1980 James e Roos (JA80) publicaram um artigo onde, segundo o resumo inicial, apresentaram as expressões "corretas" para o erro na razão entre número de eventos que obedecem à distribuições de Poisson. Esse trabalho foi motivado pela necessidade de análise de resultados de experimentos com poucos eventos, especialmente quanto aos intervalos que deveriam ser indicados. Segundo esses autores as expressões "corretas" são aquelas baseadas na abordagem clássica.

Certamente essa publicação de 1980 não foi a primeira, entre os físicos, a se preocupar com a indicação "correta" de intervalos em experimentos com poucos eventos. (Em 1951, por exemplo, Regener (RE51) havia publicado um trabalho no qual sugeria a indicação de intervalos clássicos no caso de experimentos com raios cósmicos.) No entanto, talvez seja aquela publicação que retomou, em anos recentes,

a questão de cálculos de intervalos em casos extremos em física experimental.

Em trabalho concluído em 1982 (HE82, HE83c) foi discutido um problema de certa forma parecido com aquele de James e Roos: a questão da definição de um limite superior para a área de um sinal (um pico gaussiano em um espectro multicanal) na presença de um ruído de fundo (no caso contagens de fundo na região do pico). A diferença em relação ao problema de James e Roos é que naquele trabalho foi estudado a diferença entre duas variáveis que obedecem a distribuições de Poisson, enquanto estes autores discutiram a razão entre elas. A abordagem utilizada foi a bayesiana com uma função densidade de probabilidade a priori uniforme. Não era objetivo daquele trabalho contribuir para qualquer discussão sobre questões fundamentais da estatística. Se foi adotada aquela abordagem estatística era simplesmente porque ela permitia resolver o problema em questão quaisquer que fossem as quantidades de eventos observados. Além disso a definição de um limite superior para a área de um pico parece mais adequado do que testar a hipótese de inexistência do pico, como sugerido por alguns autores (PA71, DO77). O mesmo problema foi estudado por Little (LI82) na mesma época e adotando a mesma abordagem estatística, embora restringindo a situações em que os

números de eventos observados eram suficientemente altos para permitir aproximações gaussianas.

Nessa situação cabia estender a abordagem ao problema da razão entre poucos eventos, o que foi feito a seguir (HE83a, HE84). Nessa publicação eram apresentados intervalos de confiança não só para a razão entre números pequenos, mas também para a diferença (anteriormente (HE82, HE83c) havia-se apresentado apenas limites superiores para a diferença) e para o parâmetro de uma distribuição de Poisson.

Em 1985 Prosper (PR85a), motivado pela necessidade de análise de resultados experimentais com poucos eventos, publicou um trabalho onde discute a definição de limites para a diferença entre valores que obedecem a distribuições de Poisson dentro da abordagem clássica. De fato esse problema não tem solução dentro da abordagem clássica. O que é feito nesse trabalho é uma aproximação que tem como inconveniente a indicação de um valor negativo como a estimativa de uma grandeza essencialmente não negativa. Em artigo posterior esse mesmo autor (PR85b) retoma o assunto adotando dessa vez uma abordagem bayesiana e estendendo o tratamento para a razão entre poucos eventos. O objetivo principal deste último trabalho parece ter sido o de comparar criticamente as abordagens bayesiana e clássica.

A necessidade de análise de experimentos com poucos eventos continuou motivando a publicação de intervalos de confiança. Por exemplo Gehrels (GE86) apresentou recentemente tabelas de limites superiores e inferiores para a distribuição de Poisson e para a razão entre números que obedecem a distribuições de Poisson. De certa forma esse trabalho é uma extensão daquele de James e Roos de 1980.

No entanto o problema básico continuou em aberto: qual a abordagem estatística mais adequada para a física experimental? Ou, com um pouco de ironia, qual a abordagem estatística que devemos usar até que se resolva a disputa doutrinária entre clássicos, fiducialistas e bayesianos?

Uma possibilidade é procurar entre as diferentes abordagens, qual está em melhor acordo com o que se espera de um resultado em física experimental. Com esse objetivo em 1986 foi realizado, em colaboração com Escoubes e Unamuno (ES87), uma grande quantidade de experimentos artificialmente com poucos eventos e cujos resultados exatos poderiam ser facilmente obtidos. Esses experimentos foram analisados a partir das abordagens clássica e bayesiana e referiam a determinação de parâmetros de distribuições de Poisson e da razão entre pequenos números que obedeciam a distribuições de Poisson (a diferença entre pequenos números



não foi analisada uma vez que não há solução clássica exata para o problema). A conclusão é que a abordagem bayesiana está mais próxima do que se espera de um resultado experimental em física. No Apêndice A apresentamos um resumo desse trabalho.

No presente capítulo pretendemos dar continuidade a essa comparação das diferentes abordagens estatísticas com as expectativas, (i) e (ii), em física experimental.

### 3.2 - INTERVALOS DE CONFIANÇA

A probabilidade de observarmos  $n$  eventos correspondentes a uma distribuição de Poisson com parâmetro  $a$  é

$$P_a(n) = \frac{e^{-a} a^n}{n!} \quad (3.1)$$

O problema inverso é: observados  $n$  eventos, o que podemos afirmar sobre  $a$ ? Neste caso a estimativa boa (no sentido discutido no capítulo 2) para  $a$  é  $t = n$ . No entanto a determinação de um intervalo  $(a_{-}^{(\alpha)}(n), a_{+}^{(\alpha)}(n))$  que satisfaça aos requisitos (i) e (ii) explicitados

anteriormente neste capítulo, onde  $1-\alpha$  é o conteúdo probabilístico associado ao intervalo, é menos clara. Diferentes abordagens estatísticas, e dentro delas diferentes métodos, levam a diferentes intervalos para mesmos valores de  $n$  e  $\alpha$ . Qual deles está de acordo com os requisitos (i) e (ii)? De fato, nenhum. O que podemos tentar é descobrir qual procedimento está em menor desacordo com aqueles requisitos.

Repetindo-se inúmeras vezes um experimento com o mesmo valor de  $a$ , a fração de vezes que o valor verdadeiro estará de fato dentro do intervalo indicado,  $\beta(a)$ , é

$$\beta(a) = \sum_{n|a \in (a_{-}^{(\alpha)}(n), a_{+}^{(\alpha)}(n))} P_a(n) \quad (3.2)$$

onde a soma é feita sobre os valores de  $n$  tais que  $a$  esteja contido no intervalo  $(a_{-}^{(\alpha)}(n), a_{+}^{(\alpha)}(n))$  (cf. fig. 3.1). Se  $\beta(a) = 1-\alpha$  para qualquer valor de  $a$  (e de  $\alpha$ ), a expectativa (i), correspondente a probabilidade que o valor verdadeiro esteja contido no intervalo seja  $(1-\alpha)$  100%, é satisfeita. Consequentemente a expectativa (ii) também é satisfeita.

Caso  $\beta(a)$  não seja sempre igual a  $1-\alpha$  mas varie pouco em torno dele, a expectativa (i) não é satisfeita. No

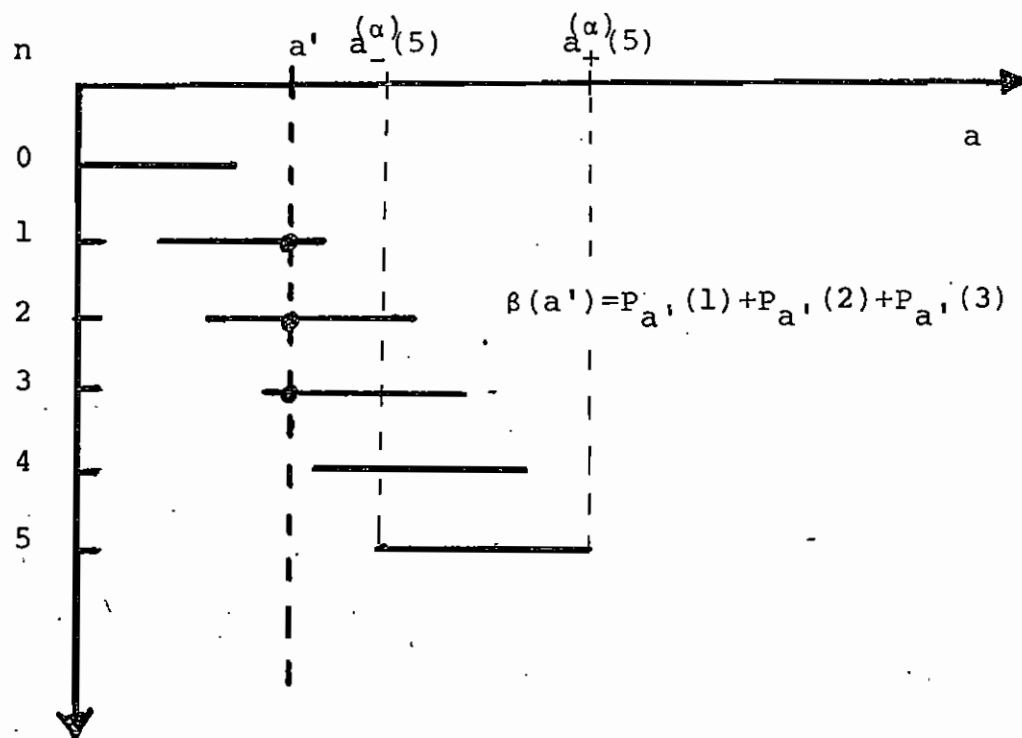


Figura 3.1 - Cada valor (verdadeiro e desconhecido) do parâmetro  $a$  pode dar origem a diferentes resultados experimentais  $n$ . A partir do conhecimento de  $n$  determina-se intervalos que podem ou não conter o valor  $a$ . A fração de vezes que  $a$  está contido no intervalo indicado é igual à soma das probabilidades que aquele valor do parâmetro dê origem a valores experimentais cujos intervalos o contém.

entanto essa variação de  $\beta(a)$  em torno de  $1-\alpha$  pode ser tal que a longo termo  $\langle \beta(a) \rangle = 1-\alpha$ , onde  $\langle \rangle$  indica uma média sobre medidas de diferentes valores de  $a$  -- correspondendo a diferentes arranjos experimentais bem como a diferentes experimentos realizados a longo termo. Assim a expectativa (ii) pode ser satisfeita sem que (i) o seja.

Neste capítulo analisaremos  $\beta(a)$  para intervalos calculados a partir de diferentes abordagens estatísticas. Os intervalos analisados são os que seguem.

#### a) Aproximação gaussiana

Quando o número de eventos é grande o intervalo de confiança pode ser calculado pela aproximação gaussiana (KE79),

$$\psi = \frac{\partial \ln L / \partial a}{E(\partial^2 \ln L / \partial a^2)} \quad (3.3)$$

onde  $L = L(n|a)$  é a função verossimilhança e  $E$  indica o valor esperado ponderado pela própria função  $L$ .  $\psi$  é aproximadamente normal com média nula e variância

$1/E(\partial^2 \ln L / \partial a^2)$ . Nessa aproximação os intervalos são dados por

$$\begin{aligned} a_-(n) &= n - \sqrt{n} \\ a_+(n) &= n + \sqrt{n}, \quad P = 0,683 \end{aligned} \quad (3.4)$$

e

$$a_-(n) = \begin{cases} 0 & n < 4 \\ n - 1,96 \sqrt{n} & n > 3 \end{cases}$$

$$a_+(n) = n + 1,96 \sqrt{n} \quad (3.5)$$

Essa aproximação só é boa quando  $n \gg 1$ . No entanto esses intervalos foram incluídos entre aqueles estudados uma vez que a aproximação gaussiana é bastante popular em física experimental.

#### b) Clássico

No esquema clássico os limites  $a_-^{(\alpha)}(n)$  e  $a_+^{(\alpha)}(n)$  são calculados por, (veja p. ex. GE86)

$$\frac{\alpha}{2} = \frac{n}{\sum_{m=0}^n \frac{e^{-a_+(n)} a_+(n)^m}{m!}} \quad (3.6a)$$

$$\frac{\alpha}{2} = \sum_{m=n}^{\infty} \frac{e^{-a_{-}(n)} a_{-}(n)^m}{m!} \quad (3.6b)$$

Ou seja, o valor  $a_{-}^{(\alpha)}(n)$  é o valor de  $a$  que daria origem a um número de eventos maior ou igual ao observado com probabilidade  $(\alpha/2)100\%$ ; o valor  $a_{+}^{(\alpha)}(n)$  é aquele que daria origem a  $n$  eventos ou menos com a mesma probabilidade.

No esquema clássico  $\beta(a)$  calculado por meio da equação 3.2 é sempre maior ou igual a  $1-\alpha$ . Essa maneira de construir intervalos de confiança pode ser entendido nos seguintes termos: na impossibilidade de determinar intervalos que contenham o valor verdadeiro da grandeza com determinado nível de confiança, é melhor superestimá-los do que subestimá-los. Esse procedimento é bastante razoável e recomendável quando o fato da grandeza não estar no intervalo indicado implica em riscos inaceitáveis. Neste caso pode ser melhor superestimar a margem desses riscos do que subestimá-los. No entanto esse não é o caso em física experimental. Os riscos de superestimação do conteúdo probabilístico são, em física, tão graves quanto os riscos de subestimação: procedendo-se assim de forma sistemática

pode-se tender a validar teorias que não são corretas ou a invalidar teorias corretas.

### c) Bayesianos

O esquema bayesiano corresponde a definir uma "função densidade de probabilidade do grau de confiança que temos em cada valor de  $a$  como sendo o valor verdadeiro da grandeza medida" posteriormente à observação. Obviamente não há sentido em falar-se em "função densidade de probabilidade" de uma grandeza que tem um valor exato. Assim a grande dificuldade da abordagem bayesiana é o seu conteúdo doutrinário ou filosófico (Kendall e Stuart (KE79), por exemplo, criticam a abordagem bayesiana pela perda de objetividade que ela implica). No entanto dada a sua facilidade de manipulação -- quando superado o problema filosófico -- e a possibilidade de incorporar-se conhecimentos prévios sobre a grandeza (baseados em vínculos físicos ou em resultados anteriores), ela é bastante confortável.

A abordagem bayesiana fundamenta-se no teorema de Bayes (KE79) que para uma distribuição de Poisson pode ser escrito na forma

$$f_p(a) = N \frac{e^{-a} a^n}{n!} f_a(a). \quad (3.7)$$

Nessa expressão  $f_p(a)$  e  $f_a(a)$  são as "funções densidades de probabilidades no grau de confiança que temos nos possíveis valores de  $a$ " após o experimento e antes do experimento respectivamente. Frequentemente  $f_a(a)$  é chamada de função densidade de probabilidade *a priori*. Na equação 3.7  $n$  é o número de eventos observados e  $N$  uma constante de normalização.

Sendo  $f_p(a)$  uma função contínua de  $a$ ,  $a > 0$ , é possível construir-se uma infinidade de intervalos de confiança  $(a_-, a_+)$  com determinado conteúdo probabilístico  $1-\alpha$ , dado por

$$\int_{a_-}^{a_+} f_p(a) da = 1-\alpha \quad (3.8)$$

Entre os possíveis intervalos consideraremos o intervalo central, onde  $a_-$  e  $a_+$  obedecem à relação

$$\int_0^{a_-} f_p(a) da = \int_{a_+}^{\infty} f_p(a) da = \frac{\alpha}{2} \quad (3.9)$$

e o intervalo mínimo com  $a_-$  e  $a_+$  obedecendo a



$$f_p(a_-) = f_p(a_+) \quad (3.10)$$

se a função é unimodal e a moda não é nula, ou  $a_- = 0$  se a moda é nula.

No esquema bayesiano quando não se tem qualquer informação a priori da grandeza medida, há várias indicações possíveis para  $f_a(a)$ . As consideradas aqui são descritas em seguida.

#### c1) Função a priori constante

A função a priori  $f_a(a)$  quantifica o conhecimento -- ou a falta de conhecimento -- que se tem da grandeza antes do experimento. Quando nada se sabe sobre a grandeza medida, é comum adotar-se em física experimental uma função densidade de probabilidade a priori uniforme (veja por exemplo VI84, HE84, HE82, RA47) em  $(0, \infty)$ . Neste caso a função densidade de probabilidade a posteriori é dada por

$$f_p(a) = \frac{e^{-a} a^n}{n!} \quad (3.11)$$

onde, diferentemente da equação (3.1),  $a$  é uma variável e  $n$  um parâmetro conhecido.

### c2) Função a priori divergente

Outra forma para quantificar a função densidade de probabilidade a priori correspondente a ignorância total no caso de uma grandeza definida no intervalo  $(0, \infty)$  (JE61, JA68) é

$$f_a(a) = 1/a \quad (3.12)$$

e portanto,

$$f_p(a) = \frac{e^{-a} a^{n-1}}{(n-1)!} \quad (3.13)$$

Esta forma leva a intervalos equivalentes aos intervalos fiduciais (KE79) que, portanto, estão implicitamente incluídos na presente análise.

Os defensores do uso da expressão 3.12 para quantificar a total ignorância da grandeza antes da realização do experimento usam argumentos bastante razoáveis, baseados principalmente em invariância da função densidade de probabilidade com mudança de unidade do

parâmetro quando este não é adimensional (veja especialmente JA68 e MU80, que discutem explicitamente essa questão e advogam em favor do uso da função divergente como a melhor forma de quantificar a ignorância total a priori).

Na impossibilidade de tratar-se o caso  $n = 0$ , nos diagnósticos feitos a seguir adotamos  $a_-(0) = a_+(0) = 0$  para qualquer nível de probabilidade tanto para o intervalo mínimo como central.

### c3) Funções a priori $a^{-1/2}$

Diversas outras formas para a função densidade de probabilidade a priori tem sido estudadas e propostas. Entre essas não podemos deixar de citar aqui o trabalho de Primakoff, de 1953 (AN53), onde os autores discutem como quantificar o conhecimento -- ou a falta de conhecimento -- anterior ao experimento em diversas situações reais.

Lewis (LE82) discute com algum humor uma situação difícil: um "chefe", nas mas palavras, manda-o medir a meia vida de um único núcleo e entregar-lhe um resultado da estimativa bem como do erro a ela associado; depois de algum tempo o núcleo decai. O que responder ao chefe? Em trabalho posterior (LE85) Lewis discute uma situação igualmente

constrangedora: o que fazer se o único núcleo não decair dentro do prazo que ele tem para dar a resposta? Em ambos os trabalhos ele mostra a dependência do resultado apresentado com a função escolhida para quantificar a ignorância total sobre a grandeza medida antes do experimento.

Além desses, vários outros autores têm discutido a questão da quantificação do conhecimento **a priori**. Como não é nosso objetivo aqui fazermos um estudo das diferentes formas dessa quantificação ou de suas bases teóricas, não vamos apresentar um levantamento completo desses trabalhos nem estender o presente estudo a todas as proposições feitas para a forma possível de  $f_a(a)$ .

Como essas outras proposições para a função **a priori** são em certo sentido intermediárias às distribuições uniforme e divergente, vamos incluir neste diagnóstico a função  $1/\sqrt{a}$ . Como a função divergente  $1/a$  esta também privilegia valores pequenos de  $a$  sendo entretanto integrável e permitindo a determinação de limites mesmo para  $n = 0$ .

### 3.3 - DIAGNÓSTICO

As tabelas 3.1 e 3.2 mostram os intervalos de confiança para  $(1 - \alpha)100\%$  igual a 68,3% e 95%

n	A	B	C	D	E	F	G	H
0	0.00	0.00	0.17	1.84	0.00	0.00	0.02	0.00
1	0.00	0.17	0.71	3.30	0.17	0.00	0.42	0.05
2	0.58	3.41	0.71	4.64	0.86	0.26	1.03	0.54
3	1.27	4.73	1.37	5.92	1.37	0.64	1.72	1.20
4	2.00	6.00	2.09	7.17	2.09	1.55	2.46	1.92
5	2.76	7.24	2.84	8.39	2.84	2.29	3.23	2.67
6	3.55	8.45	3.62	9.59	3.62	3.06	4.02	3.45
7	4.35	9.65	4.42	10.8	4.42	3.85	4.82	4.25
8	5.17	10.8	5.23	12.0	5.23	4.65	5.64	5.06
9	6.00	12.0	6.06	13.1	6.06	5.47	6.47	5.83
10	6.84	13.2	6.90	14.3	6.89	6.30	7.31	6.72
11	7.68	14.3	7.73	15.5	7.74	7.14	8.16	7.54
12	8.54	15.5	8.60	16.6	8.59	7.99	9.01	8.41
13	9.39	16.6	9.44	17.7	9.44	8.84	9.87	9.27
14	10.3	17.7	10.3	18.8	10.3	9.72	10.7	10.1
15	11.1	18.9	11.2	20.0	11.2	10.6	11.6	11.0
16	12.0	20.0	12.1	21.1	12.1	11.5	12.5	11.9
17	12.9	21.1	13.0	22.2	13.0	12.4	13.4	12.8
18	13.8	22.2	13.8	23.3	13.8	13.2	14.2	13.6
19	14.6	23.4	14.7	24.4	14.7	14.1	15.1	14.5
20	15.5	24.5	15.6	25.5	15.6	15.0	16.0	15.4

Tabela 3.1 - Limites inferiores e superiores para intervalos de confiança  $1-\alpha = 0,683$ .

(A) Aproximação gaussiana; (B) Abordagem clássica; abordagens bayesianas com função a priori uniforme e intervalos (C) centrais e (D) mínimos; abordagens bayesianas com função a priori  $1/a$  e intervalos (E) centrais e (F) mínimos; abordagens bayesianas com função a priori  $1/\sqrt{a}$  (G) centrais e (H) mínimos.

n	A	B	C	D	E	F	G	H
0	0.00	0.00	3.69	0.00	0.00	0.00	4.9E-4	0.00
1	0.00	2.96	5.57	4.2E-2	3.69	0.00	0.11	1.6E-3
2	0.00	4.77	7.22	0.30	5.57	4.2E-2	0.42	0.15
3	0.00	6.39	8.77	0.71	7.22	0.30	0.84	0.49
4	8.0E-2	7.92	10.2	1.21	8.77	0.71	1.35	0.95
5	0.62	9.38	11.7	1.76	10.2	1.21	1.91	1.48
6	1.20	10.8	13.1	2.35	11.7	1.76	2.50	2.05
7	1.81	12.2	14.4	2.97	13.1	2.35	3.13	2.66
8	2.46	13.5	15.8	3.62	14.4	2.97	3.78	3.28
9	3.12	14.9	17.1	4.29	15.8	3.62	4.45	3.95
10	3.80	16.2	18.4	4.98	17.1	4.29	5.14	4.63
11	4.50	17.5	19.7	5.67	18.4	4.98	5.84	5.33
12	5.21	18.8	21.0	6.40	19.7	5.67	6.56	6.04
13	5.93	20.1	22.3	7.12	21.0	6.40	7.29	6.75
14	6.67	21.3	23.5	7.86	22.3	7.12	8.02	7.49
15	7.41	22.6	24.8	8.60	23.5	7.86	8.77	8.23
16	8.16	23.8	26.0	9.36	24.8	8.60	9.52	8.98
17	8.92	25.1	27.2	10.1	26.0	9.36	10.3	9.73
18	9.68	26.3	28.4	10.9	27.2	10.1	11.1	10.5
19	10.5	27.5	29.7	11.6	28.4	10.9	11.8	11.3
20	11.2	28.8	30.9	12.4	29.6	11.6	12.6	12.1
								29.5

Tabela 3.2 - Limites inferiores e superiores para intervalos de confiança  $1-\alpha = 0,95$ .

(A) Aproximação gaussiana; (B) Abordagem clássica; abordagens bayesianas com função a priori  $1/a$  e intervalos (E) centrais e (F) mínimos; abordagens bayesianas com função a priori  $1/\sqrt{a}$  (G) centrais e (H) mínimos.

respectivamente, para os oito intervalos apresentados na seção 3.2 e para  $n = 0$  à  $n = 20$ . A função  $\beta(a)$ , foi calculada para cada um dos oito diferentes intervalos e para os dois valores de  $1-\alpha$ . As figuras 3.2 a 3.9 mostram graficamente os resultados, comentados a seguir com base no que se espera em física experimental.

#### a) Aproximação gaussiana

A figura 3.2 mostra  $\beta(a)$  para o intervalo estimado por  $\sqrt{n}$ . Como pode se ver nesta figura experimentos com valores pequenos de  $a$ ,  $a < 3$ , implicam em uma grande tendência de subestimação do conteúdo probabilístico dos intervalos indicados. Como diversos experimentos importantes em física corresponde a situações de valores pequenos para os parâmetros, essa tendência de subestimação do conteúdo probabilístico do intervalo é grave. Certamente intervalos com esse defeito são inadequados tanto para a incorporação de resultados experimentais a teorias ou modelos bem como para a comparação entre resultados experimentais,

Para valores  $a > 3$  embora haja flutuações de  $\beta(a)$  em torno do valor  $1-\alpha$  desejado estas não são extremamente graves. Isso se deve ao fato que  $\beta(a)$  é a fração média de vezes que  $a \in (a_-, a_+)$  para um determinado valor de  $a$  ou

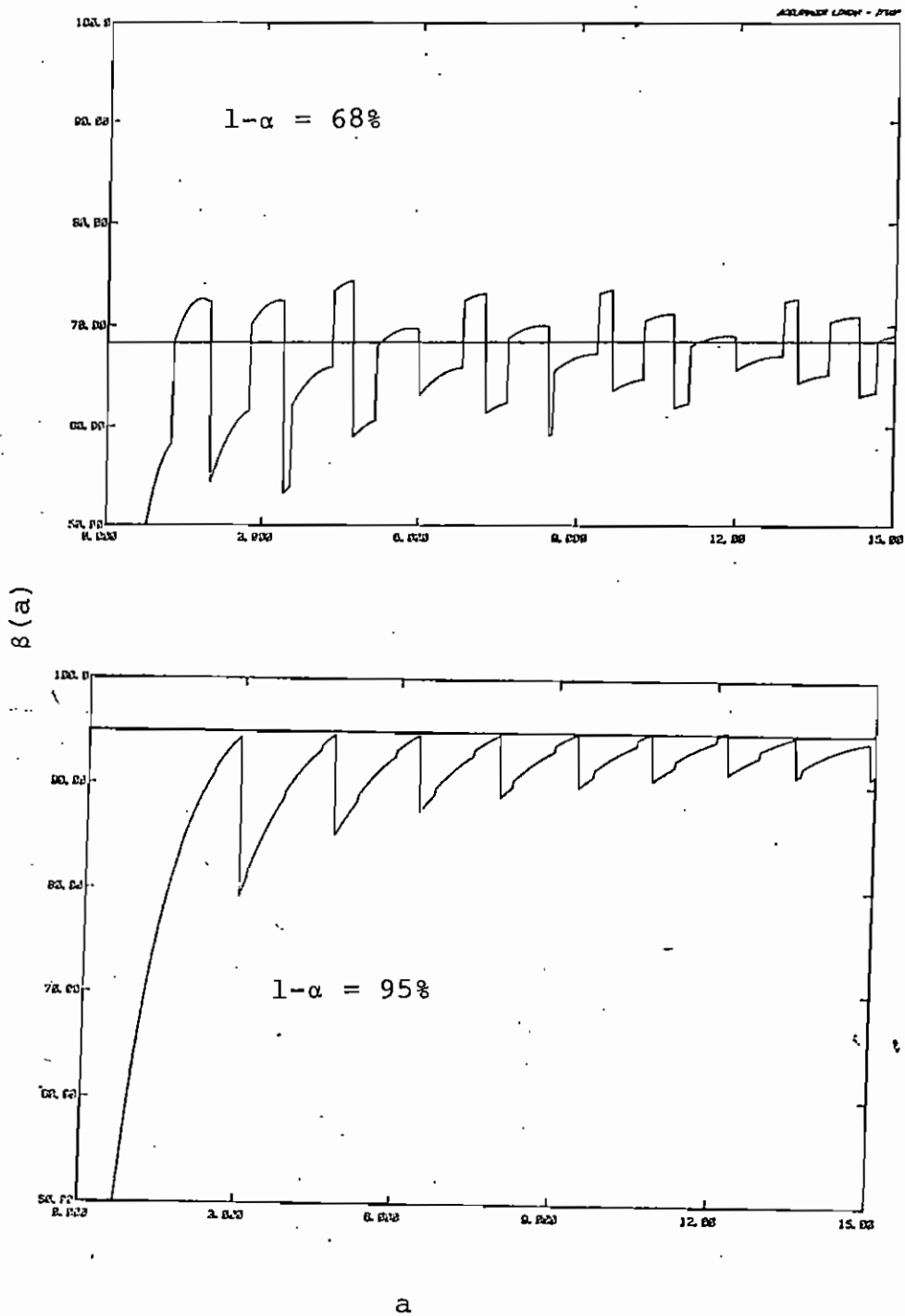


Figura 3.2 -  $\beta(a)$  em função de  $a$  para intervalos de confiança determinados pela aproximação gaussiana.



seja, a fração média de vezes que o valor verdadeiro da grandeza está no intervalo indicado quando se repete infinitas vezes o experimento com aquele valor de  $a$ . Isso de fato não ocorre em física experimental. Mesmo repetindo-se inúmeras vezes experimentos com o objetivo de medir uma mesma grandeza, o valor do parâmetro verdadeiro não permanece o mesmo: há alterações no arranjo experimental (como volume e eficiência de detectores por exemplo), duração da medida e frequência dos eventos medidos. No entanto este comentário não atenua a gravidade da subestimação do conteúdo probabilístico para pequenos valores de  $a$  uma vez que muitos experimentos na física atual correspondem a medidas naquela situação.

#### b) Clássico

Os intervalos clássicos, como comentado anteriormente, superestimam sistematicamente o conteúdo probabilístico. Isso ocorre particularmente para valores pequenos de  $a$ , como mostra a figura 3.3, tanto para  $1-\alpha = 0,683$  como para  $1-\alpha = 0,95$ . Do ponto de vista da física experimental essa superestimação é tão grave quanto à subestimação, por impossibilitar intercomparações de

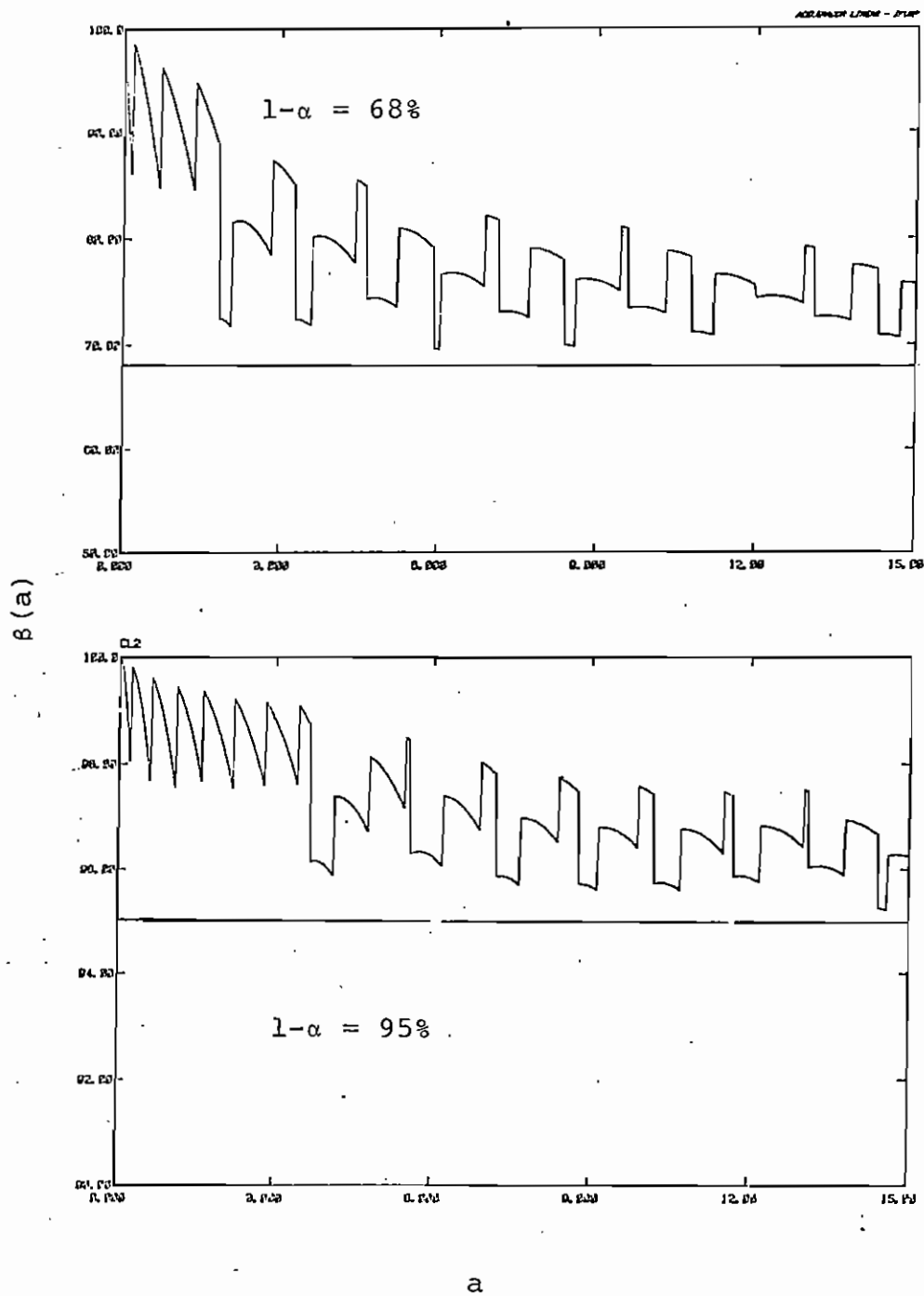


Figura 3.3 -  $\beta(a)$  em função de  $a$  para intervalos determinados pela abordagem clássica. Neste caso  $\beta(a) > 1-\alpha$  para qualquer valor de  $\alpha$  e  $a$ .

experimentos bem como a incorporação de seus resultados a teorias e modelos.

A superestimação sistemática que ocorre para valores mais altos de  $a$  não é extremamente grave, embora seja bastante inconveniente em procedimentos de longo termo. Nestes casos, mesmo que haja uma variação de  $a$  por variações das condições experimentais, como anteriormente comentado, ela não pode compensar a superestimação do conteúdo probabilístico.

#### c) Bayesiano uniforme mínimo

A fig. 3.4 mostra  $\beta(a)$  para a abordagem bayesiana, com função densidade de probabilidade a priori constante e intervalos mínimos. Neste caso há uma superestimação do conteúdo probabilístico para pequenos valores de  $a$ . No entanto a superestimação é menos grave do que no caso clássico por ser menos intensa e de menos alcance em  $a$ .

Quanto à flutuação que ocorre para valores mais altos de  $a$ , ela pode ser compensada por variações experimentais como discutido anteriormente.

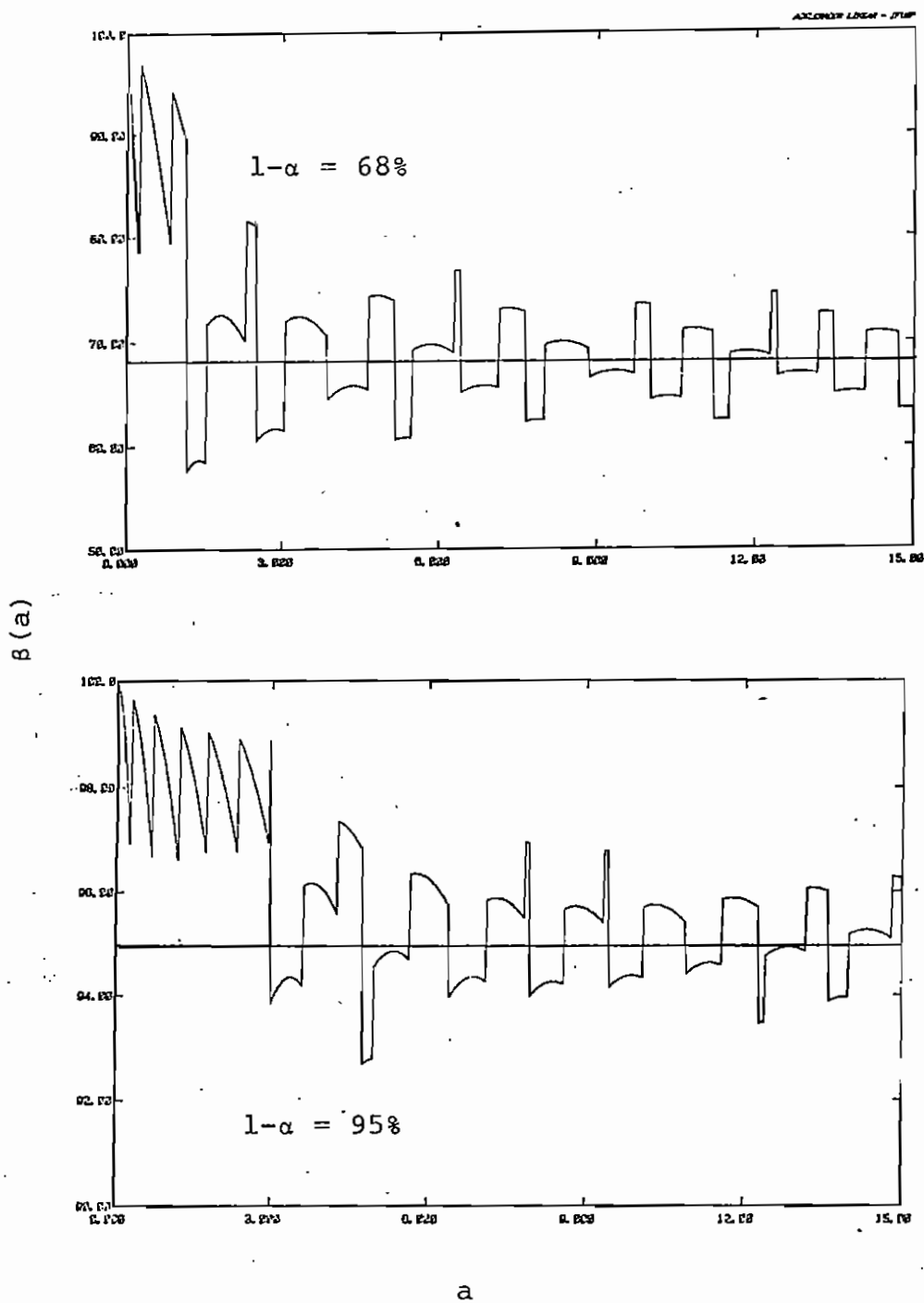
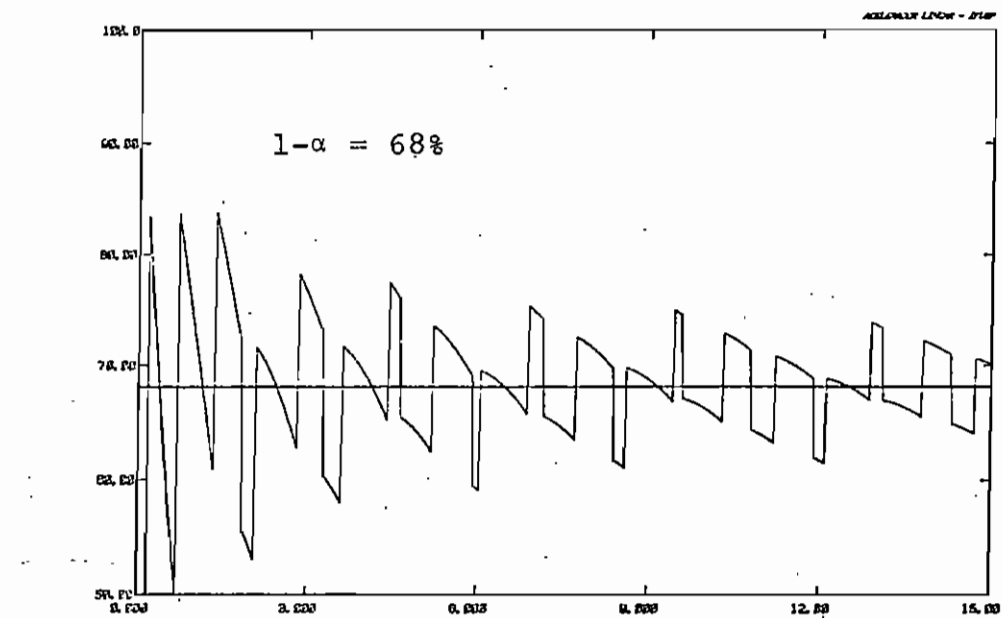
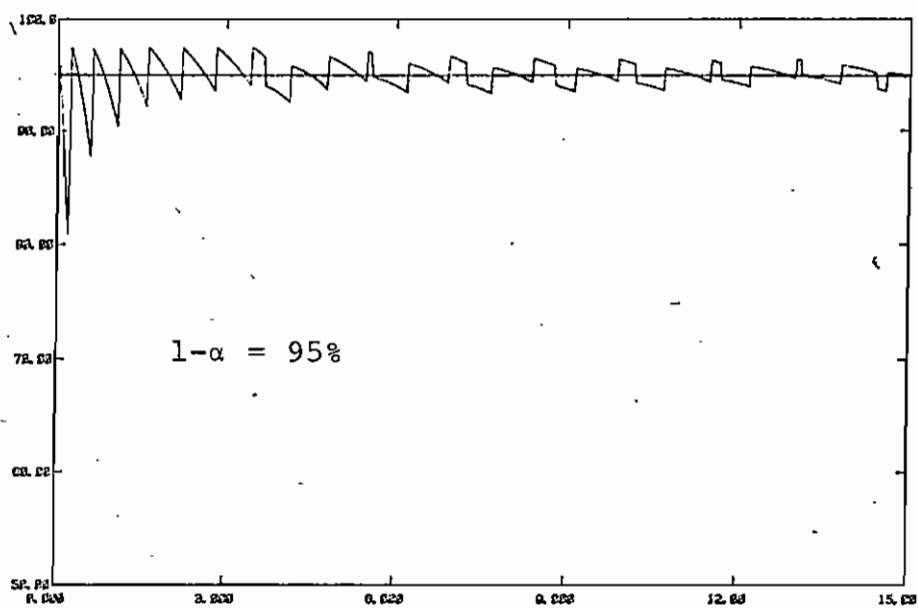


Figura 3.4 - Estas figuras mostram  $\beta(a)$  para a abordagem bayesiana com função densidade de probabilidade a priori uniforme, correspondente a intervalos mínimos.

 $\beta(a)$ 

a

Figura 3.5 - O mesmo que a figura 3.4, correspondente neste caso aos intervalos centrais.

**d) Bayesiano uniforme central**

Este tipo de intervalo, para o qual  $\beta(a)$  é ilustrado na fig. 3.5, mostra o inconveniente de subestimação para pequenos valores de  $a$ . Como anteriormente, as flutuações em torno do valor  $1-\alpha$  para valores mais altos de  $a$  não são graves, sendo inclusive menores do que no caso dos intervalos do item c acima.

**e) Bayesianos divergentes: central e mínimo**

As figuras 3.6 e 3.7 mostram  $\beta(a)$  para os intervalos bayesianos construídos com o uso de uma função densidade de probabilidade a priori  $1/a$ .

A primeira dessas figuras corresponde a intervalos centrais e a segunda a intervalos mínimos. Apesar da fundamentação sólida que existe em defesa do uso da densidade a priori  $1/a$  no caso da distribuição de Poisson (JE61, JA68), não parece haver, na prática, qualquer vantagem desta sobre a densidade uniforme. Ao contrário, estes intervalos tem uma tendência a subestimar, de maneira sistemática e acentuada, o conteúdo probabilístico para valores de  $a$  não muito pequenos. Apenas para quantificar essa tendência podemos supor alguém ou algum laboratório que

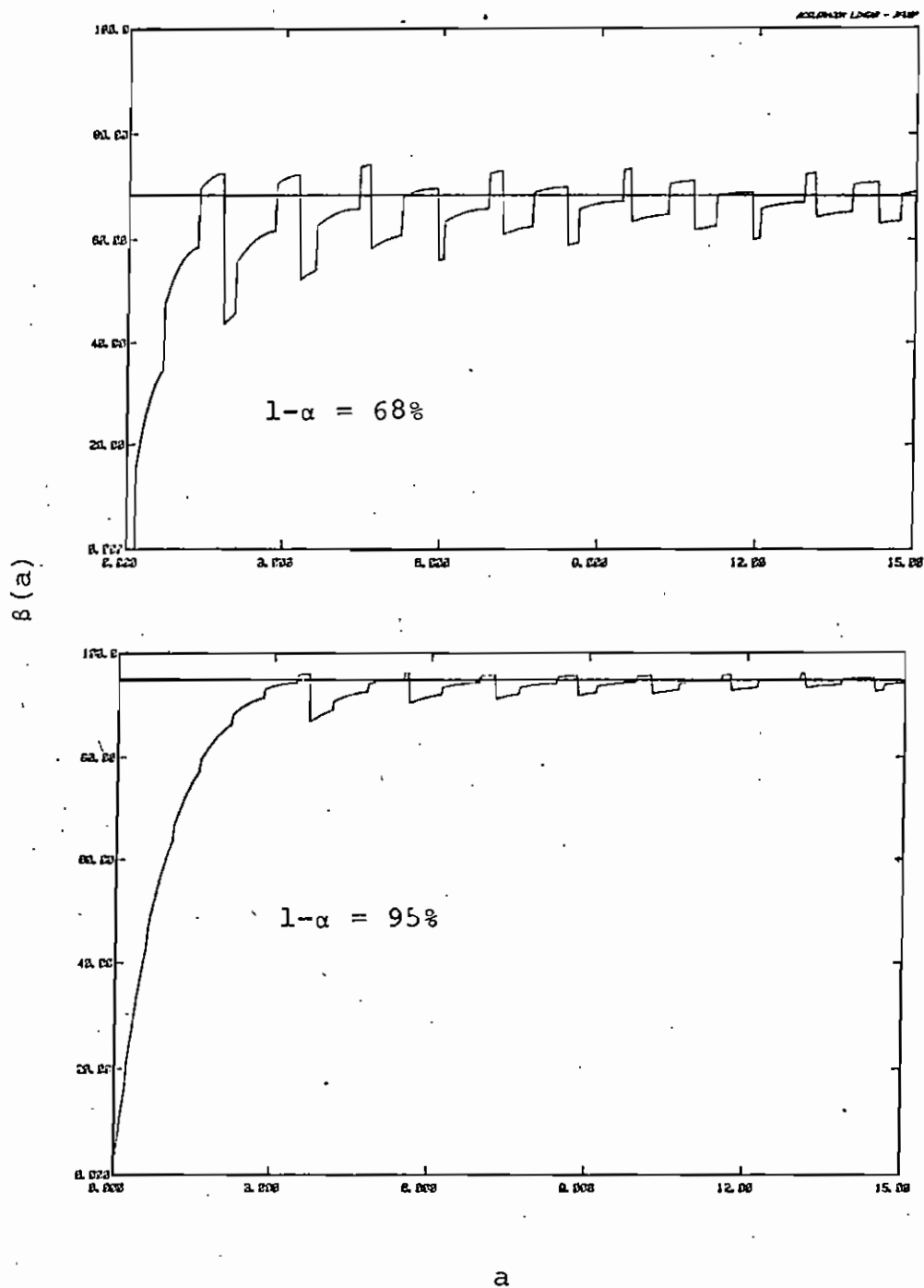


Figura 3.6 -  $\beta(a)$  para a abordagem bayesiana, intervalos centrais, com função densidade de probabilidade a priori  $1/a$ .

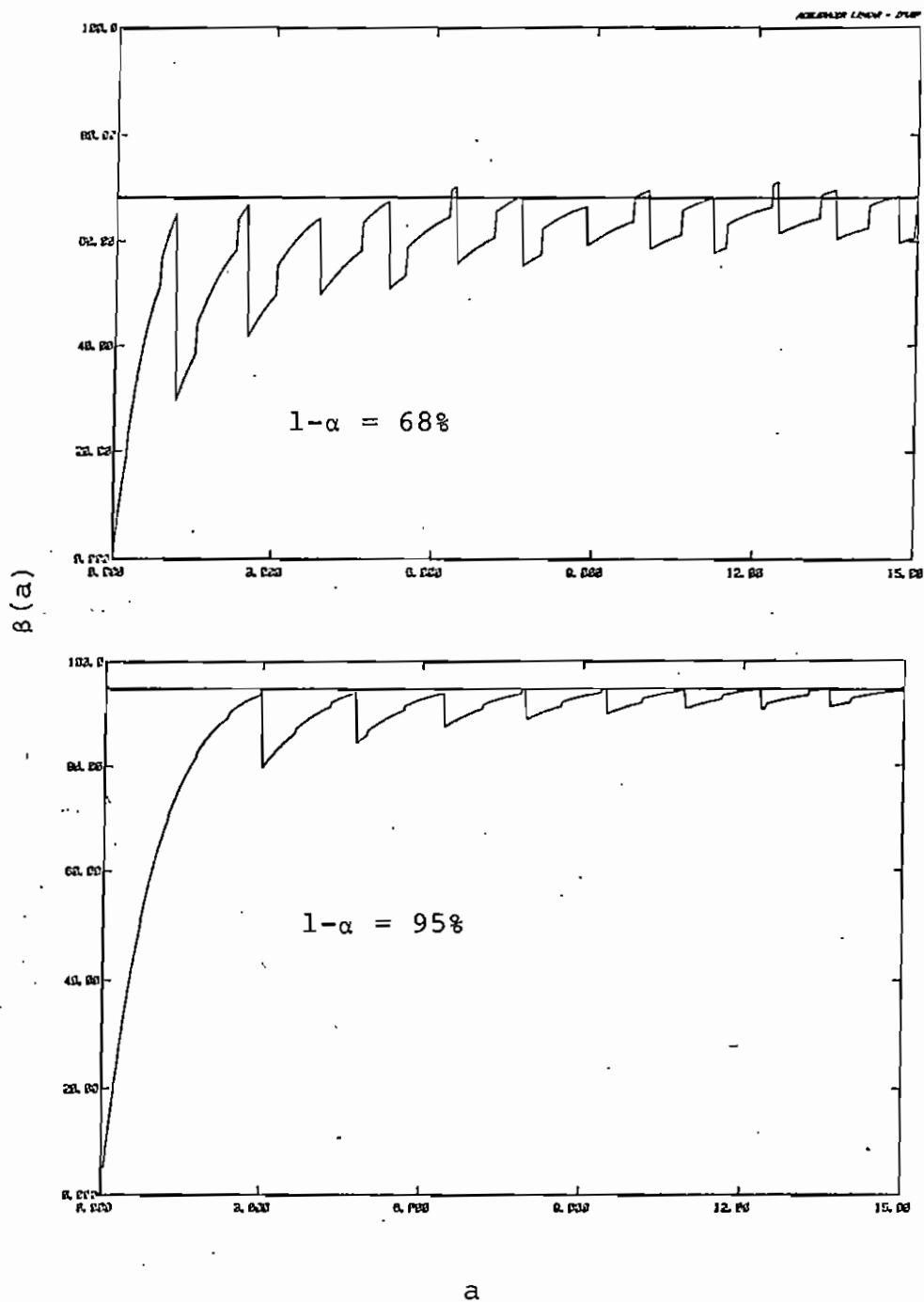


Figura 3.7 -  $\beta(a)$  para a abordagem bayesiana, intervalos mínimos, com função densidade de probabilidade a priori  $1/a$ .



se dedique, a longo termo, a medidas de grandezas que apresentam valores de  $a$  entre 0 e 3 e use intervalos divergentes com "conteúdo probabilístico" de 68,3%; de fato essa pessoa ou esse laboratório estaria em média indicando intervalos de conteúdo probabilístico entre 40% e 45%, uma subestimação nada desprezível.

Para valores mais altos de  $a$  o comportamento desses intervalos não é muito diferente dos outros indicados, embora a tendência à subestimação se prolongue até valores de  $a$  significativamente maiores do que 3, especialmente no caso de intervalos mínimos.

f) Bayesiano a priori  $a^{-1/2}$  central

Embora seja dificilmente perceptível pelos gráficos da figura 3.8 há, neste caso, uma subestimação do conteúdo probabilístico para  $a < 0,02$  no caso  $1-\alpha = 0,683$  e para  $a < 5 \cdot 10^{-4}$  para  $1-\alpha = 0,95$ . Essas subestimações ocorrem pelo fato de que esses valores,  $0,02$  e  $5 \cdot 10^{-4}$ , são os limites inferiores dos intervalos para  $n = 0$  para  $1-\alpha = 0,683$  e  $1-\alpha = 0,95$  respectivamente. Para valores mais altos de  $a$  há uma flutuação de  $\beta(a)$  em torno dos valores  $1-\alpha$  desejados.

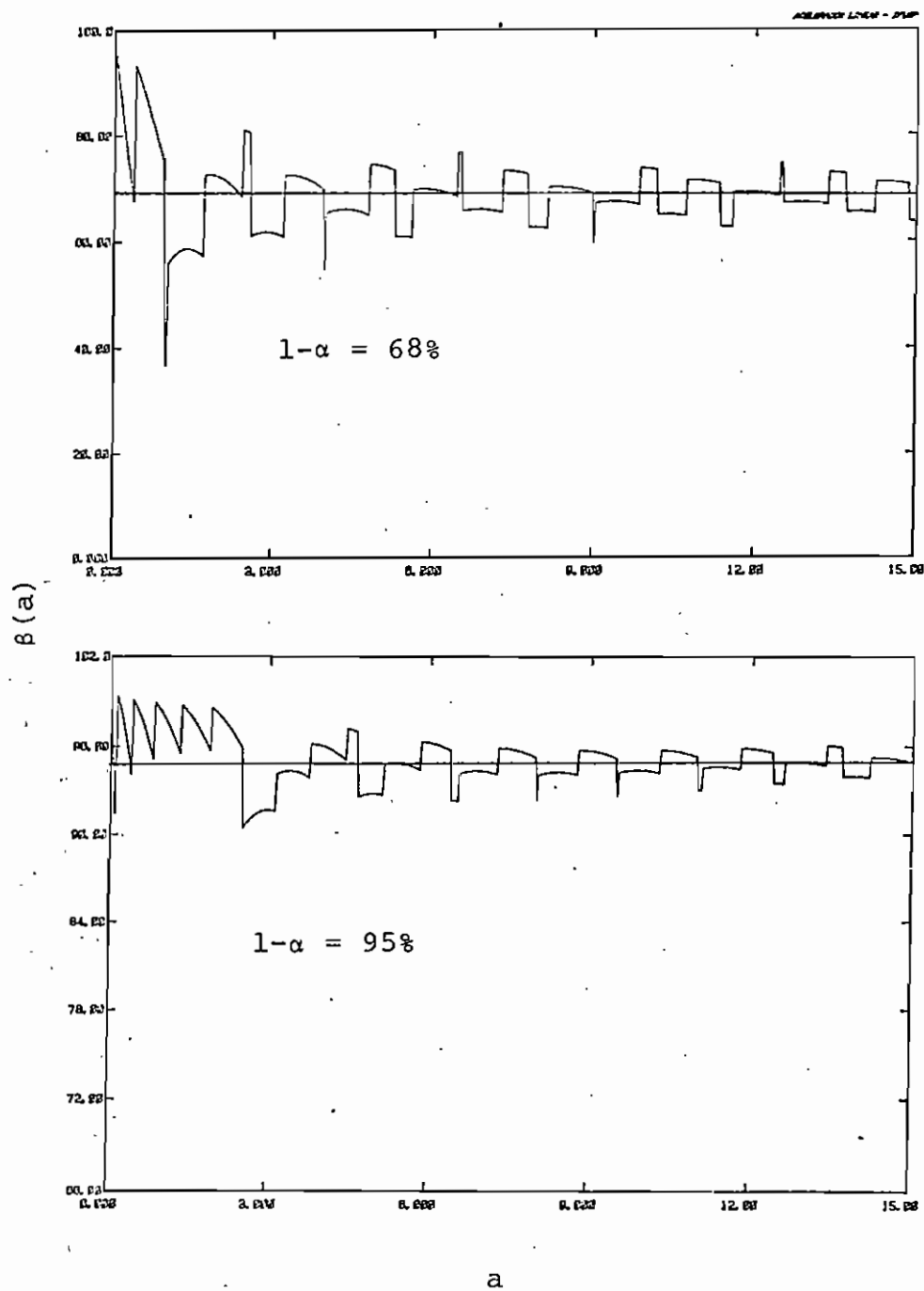


Figura 3.8 -  $\beta(a)$  para intervalos centrais determinado pela abordagem bayesiana com função densidade de probabilidade a priori  $1/\sqrt{a}$ .

g) Bayesiano a priori  $a^{-1/2}$  mínimo

A função densidade de probabilidade a priori  $a^{-1/2}$  dá origem a intervalos mínimos para os quais  $\beta(a)$ , mostrados na figura 3.9, tem um comportamento bastante parecido com o comportamento apresentado no caso de intervalos mínimos com função densidade de probabilidade a priori constante. No entanto a flutuação de  $\beta(a)$  em torno dos valores  $1-\alpha$  desejados para  $a > 2$ , é maior do que naquele caso, como pode ser visto comparando-se as figuras 3.9 e 3.4.

### 3.4 - CONCLUSÃO

Entre as várias alternativas possíveis para a estimativa de intervalos estudadas neste capítulo, nenhuma está de acordo com o que se espera de um resultado experimental. A primeira dessas expectativas é que a cada vez que indicamos um intervalo com conteúdo probabilístico  $(1-\alpha).100\%$  haja uma probabilidade de que o valor verdadeiro da grandeza esteja de fato no intervalo de  $(1-\alpha).100\%$ . A segunda, menos restritiva quanto ao comportamento de  $\beta(a)$ , é que pelo menos em média, em procedimento de longo termo,

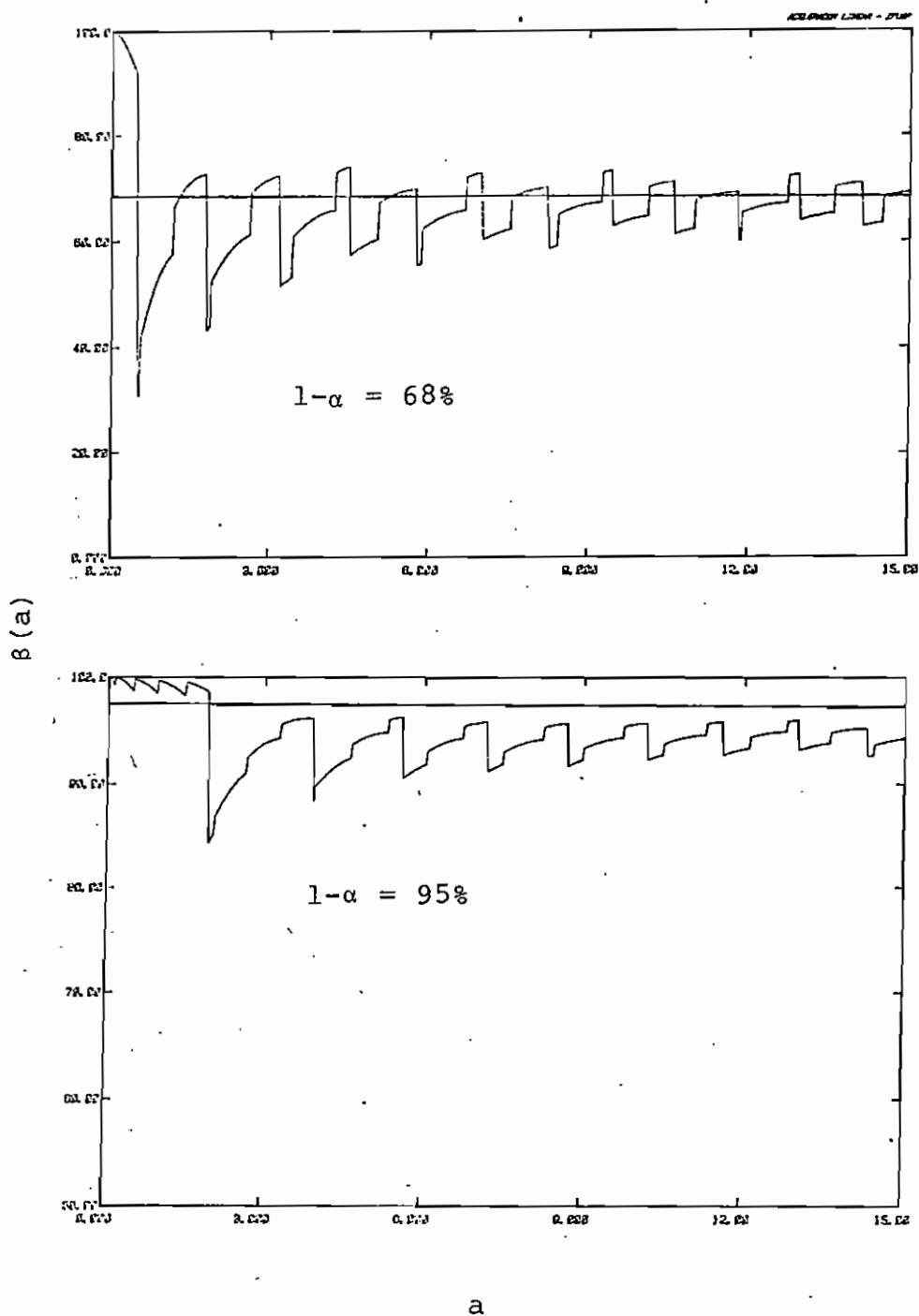


Figura 3.9 -  $\beta(a)$  para intervalos centrais determinado pela abordagem bayesiana com função densidade de probabilidade a priori  $1/\sqrt{a}$ .

o valor verdadeiro esteja contido no intervalo uma fração  $1-\alpha$  das vezes.

Com a excessão dos intervalos clássicos, que superestimam sempre o conteúdo probabilístico desejado, todos os outros intervalos estudados apresentam  $\beta(a)$  não muito diferente do valor desejado  $1-\alpha$  para  $a > 3$ . As diferenças observadas não são significativas a ponto de alterar os aspectos qualitativos das conclusões possíveis (está um resultado de acordo ou não com uma hipótese? São dois resultados compatíveis entre si?). A longo termo, repetindo-se diversas vezes experimentos com  $a > 3$ , as flutuações de  $\beta(a)$  em torno de  $1-\alpha$  tendem a diminuir ainda mais as consequências qualitativas das pequenas diferenças do conteúdo probabilístico em relação ao desejado.

O mesmo não ocorre para pequenos valores de  $a$ . A tabela 3.3 mostra o desvio padrão de  $\beta(a)$  em relação a  $1-\alpha$  para  $a \in (0,3)$  no caso de intervalos com  $1-\alpha = 0,683$ . Está também indicada nessa tabela a tendência dos intervalos nessa região. Parece-nos que os intervalos bayesianos com funções a priori uniforme ou  $1/\sqrt{a}$  são aqueles que mais se aproximam do que se deseja da representação de um resultado em física experimental sendo que os intervalos uniformes centrais são aqueles que apresentam menor desvio médio em relação ao valor  $1-\alpha$  esperado.

O "defeito" que os intervalos bayesianos uniformes centrais apresentam quando  $a < a_{(0)}$ , que implica em  $\beta(a) = 0$ , pode ser corrigido de forma tópica fazendo

Tabela 3.3 - Comportamento dos oito intervalos estudados para pequenos valores de  $a$ .

Intervalo	desvio médio $a \in (0,3)$	tendência
Aproximação gaussiana	25%	Subestimação
Clássico	20%	Superestimação
Bayesiano uniforme mínimo	15%	
Bayesiano uniforme central	11%	
Bayesiano $1/a$ central	22%	Subestimação
Bayesiano $1/a$ mínimo	30%	Subestimação
Bayesiano $1/\sqrt{a}$ central	13%	
Bayesiano $1/\sqrt{a}$ mínimo	18%	

$a_{(0)} = 0$  para qualquer nível de significância. A tabela 3.4 mostra esses intervalos "curados" dessa forma e a figura 3.10 mostra  $\beta(a)$  para esses intervalos. Com essa alteração o desvio médio de  $\beta(a)$  em relação a 68,3% para  $a \in (0,3)$ , que era de cerca de 11%, reduz-se para cerca da metade.

Tabela 3.4 - Intervalos "curados" baseados nos intervalos bayesianos centrais com distribuição a priori uniforme (limites inferiores - limites superiores).

n	68,3%		95,0%	
0	0,0	1,84	0,0	3,69
1	0,71	3,3	0,24	5,57
2	1,37	4,64	0,62	7,22
3	2,09	5,92	1,09	8,77
4	2,84	7,17	1,62	10,2
5	3,62	8,39	2,2	11,7
6	4,42	9,59	2,81	13,1
7	5,2	10,8	3,45	14,4
8	6,1	11,9	4,12	15,8
9	6,9	13,1	4,8	17,1
10	7,7	14,3	5,5	18,4
11	8,6	15,5	6,2	19,7
12	9,4	16,6	6,9	21,0
13	10,3	17,7	7,7	22,3
14	11,2	18,8	8,4	23,5
15	12,1	20,0	9,2	24,8
16	12,9	21,1	9,9	26,0
17	13,8	22,2	10,7	27,2
18	14,7	23,3	11,5	28,4
19	15,6	24,4	12,2	29,6
20	16,5	25,5	13,0	30,8

O grande defeito da função a priori divergente  $1/a$  é a forte tendência de subestimação do conteúdo probabilístico para pequenos valores de  $a$ . Esse defeito é inaceitável quando diversos experimentos recentes, relacionados com aspectos de fronteira em física, correspondem a valores pequenos do número de eventos observados e, conseqüentemente, dos parâmetros medidos.

É importante observarmos entretanto que se o número de eventos observado for nulo, o que em física ocorre com grande frequência em experimentos planejados para o estudo de eventos muito raros, a comparação com uma teoria é mais facilmente feita determinando-se a probabilidade de  $n = 0$  eventos na hipótese de sua validade. O mesmo procedimento é adequado quando a comparação é com outro resultado experimental. Para  $n > 0$  os intervalos apresentados na tabela 3.4 são perfeitamente utilizáveis. No entanto, em qualquer caso, a apresentação dos dados brutos é essencial quando se trata de experimentos com poucos eventos.

O diagnóstico apresentado aqui está de acordo com resultados de realização de muitos experimentos com poucos eventos (ES87), que parece privilegiar a abordagem bayesiana com função densidade de probabilidade a priori uniforme no caso de razão e diferença de pequenos números, cujo resumo é



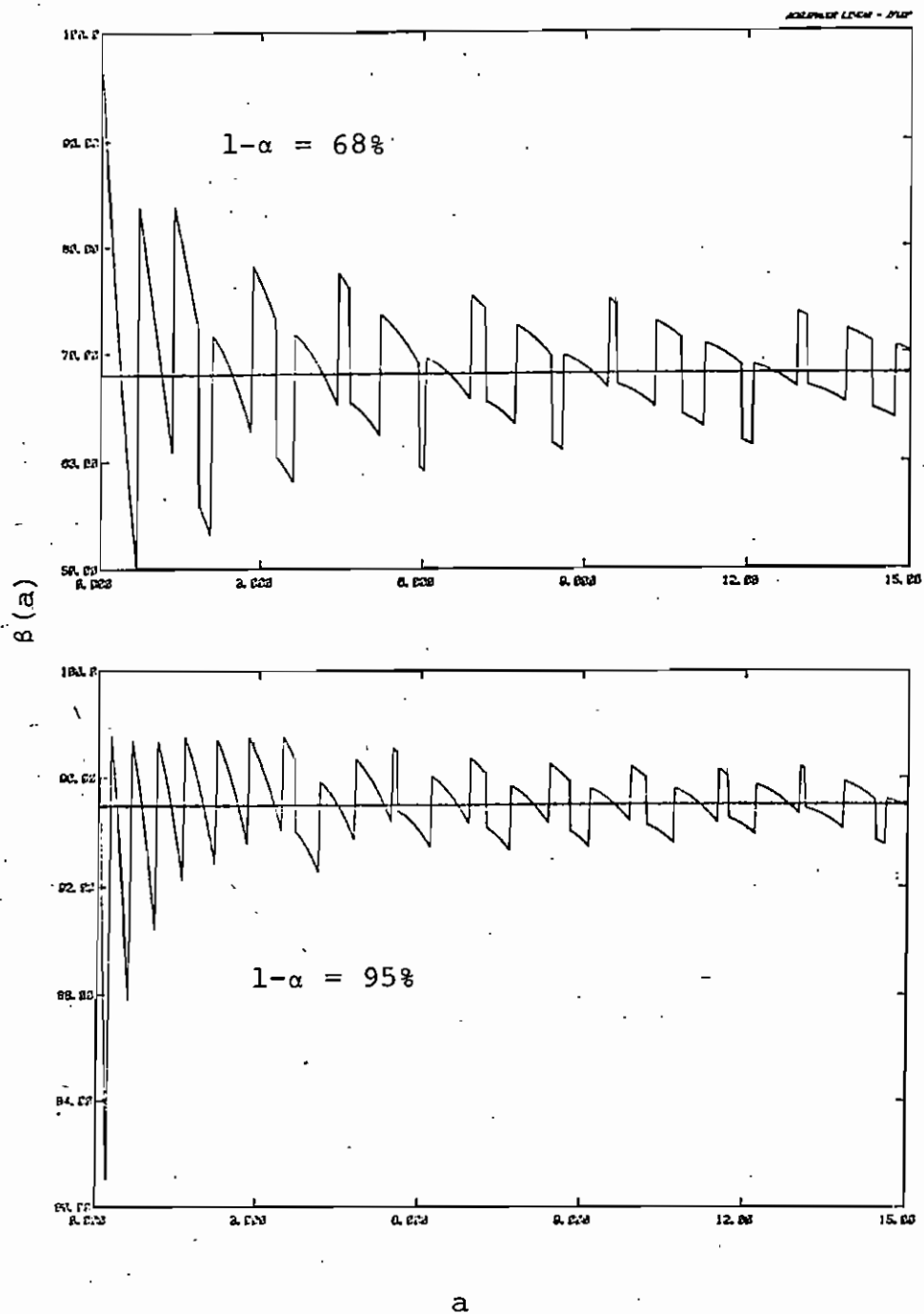


Figura 3.10 -  $\beta(a)$  para os intervalos apresentados na tabela 3.4.

apresentado no apêndice A. Pretende-se estender o presente diagnóstico para outros casos além da distribuição de Poisson. A conclusão que podemos chegar no momento é que os intervalos apresentados na tabela 3.4 e cujos valores (a) são mostrados na figura 3.10 estão bastante perto do que se espera em física experimental.

#### 4. PLANEJAMENTO E ANÁLISE DE EXPERIMENTOS

Em um experimento é desejável que o procedimento adotado, o arranjo experimental e a análise dos dados sejam tais que a precisão final obtida seja a máxima possível para dado esforço dispendido.

Neste capítulo analisaremos algumas situações reais chamando a atenção para aspectos da análise de dados. Essas situações são extremas no que diz respeito a quantidade de eventos envolvida: poucos eventos, muitos eventos e eventos comparáveis com a flutuação estatística do fundo. A primeira situação analisada corresponde a um experimento recentemente concluído no Laboratório do Acelerador Linear do Instituto de Física da USP, no qual usou-se procedimentos de vínculos entre parâmetros com o objetivo de minimizar o efeito de erros sistemáticos. A segunda situação corresponde a um trabalho em andamento no mesmo laboratório onde a maximização da qualidade experimental é feita com base na quantidade de informação. A terceira análise diz respeito a alteração de calibração em espectros multicanais com a preservação da flutuação

estatística. A última análise refere-se a quantidade de informação em diversos arranjos experimentais possíveis para o estudo do duplo decaimento beta, cujos experimentos são caracterizados pelo pequeno número de eventos, mesmo como fundo.

#### 4.1 - USO DE VÍNCULOS E COVARIÂNCIAS NA ANÁLISE DE DADOS

Diversos experimentos em física correspondem a medidas de grandezas vinculadas entre si, de maneira bem conhecida. Embora em alguns casos essas condições de vínculo sejam explicitamente impostas na análise dos dados, em outros casos isso não ocorre. No entanto há vantagens evidentes na imposição de vínculos, entre elas o aumento da quantidade de informação contida no resultado apresentado para cada parâmetro. Na primeira parte desta seção discutiremos a imposição das condições de vínculo em um caso concreto.

Outro ponto discutido nesta seção é o uso e geração de covariâncias entre parâmetros medidos. Mannhart (MA81) apontou a importância da consideração dessas covariâncias em determinações precisas de seções de choque; em trabalho recente foi mostrado como considerar tais

covariâncias bem como a importância delas no que diz respeito à preservação do conteúdo probabilístico dos intervalos (HE86), aplicando-se posteriormente o método a um problema de espectroscopia gama (MU87a).

#### a) Uso de vínculos

A determinação da energia de níveis nucleares excitados pode ser feita por meio de espectroscopia gama. Neste caso é usual a ocorrência de esquemas de decaimento gama nos quais a quantidade de níveis excitados é menor do que o número de transições gama mensuráveis. Esse fato impõe vínculos entre transições bem como permite a determinação das energias dos níveis por múltiplas combinações das energias das transições.

Assim o método dos mínimos quadrados com vínculo pode e deve ser usado, com vantagens. (Neste caso os vínculos podem corresponder inicialmente a hipóteses que devem ser testadas mas que são posteriormente adotadas como verdadeiras.)

Suponhamos que as energias de transições gama  $E_1, E_2, \dots, E_n$  tenham sido determinadas experimentalmente, sem a imposição de vínculos, como sendo

$$Y = \begin{pmatrix} E_1 \\ E_2 \\ E_3 \\ \vdots \\ E_n \end{pmatrix}, \quad (4.1)$$

sendo  $V$  a matriz de covariância desses dados. Os vínculos entre as energias são dadas por equações do tipo

$$\sum_j g_{ij} E_j = r_i \quad (4.2)$$

onde  $g_{ij}$  é  $\pm 1$  ou  $0$  e  $r_i$  as diferenças de energias de recuo do núcleo, sendo que  $i = 1, 2, \dots, n_g$  um índice que corresponde a relação de vínculo considerada. As equações (4.2) podem ser representada na forma matricial como

$$GA = R, \quad (4.3)$$

onde  $A$  é o vetor representando as energias em obediência às equações de vínculo, a ser ajustado pelo método dos mínimos quadrados,  $R$  o vetor contendo os resíduos das energias de recuo e  $G$  uma matriz retangular cujos elementos são os  $g_{ij}$  acima.

É direto demonstrar que o resultado do ajuste leva a (HE86)

$$A = MV^{-1}Y - BC^{-1}R \quad (4.4)$$

onde

$$B = VG^t, \quad (4.5)$$

$$C = -GVG^t \quad (4.6)$$

e

$$M = V + BC^{-1}B^t. \quad (4.7)$$

M é a matriz de covariâncias das energias ajustadas com vínculo. A equação (4.7) pode ser reescrita na forma

$$M = V - VG^t(GVG^t)^{-1}GV \quad (4.8)$$

O elemento  $M_{ii}$  é a variância da energia  $E_i$  ajustada em obediência às equações de vínculo. Uma inspeção da equação (4.8) mostra que a variância correspondente a medida de uma energia pode diminuir até de um fator igual ao número de equações de vínculo em que ela está envolvida. Como em

alguns esquemas de decaimento gama a quantidade de equações de vínculo pode ser bastante grande, como ilustrado em seguida, o ganho de informação na energia de uma transição (e dos próprios níveis excitados) é bastante importante.

O uso de vínculos na análise de alguns experimentos pode ser útil também para a redução e o controle de erros sistemáticos ligados ao arranjo experimental, mediante o uso de artifícios instrumentais (MU87a). O Apêndice C mostra algumas consequências do uso das equações de vínculo.

Como ilustração da utilidade do método dos mínimos quadrados com vínculo mostramos o resultado da espectroscopia gama do  $^{133}\text{Ba}$  com o uso de detectores semicondutores (de resolução típica de 2 keV), recentemente concluída (MU87a). A tabela 4.1 mostra na primeira coluna as energias das transições gama do  $^{133}\text{Ba}$ , medidas com o mesmo tipo de detectores usados naquele trabalho, sem o uso de condição de vínculo ou de artifícios que ela permite. Na segunda coluna mostramos as energias determinadas com a imposição das condições de vínculo. Na terceira coluna mostramos as energias e respectivos desvios padrões levando em conta os resultados mostrados na segunda coluna e dados de publicação recente no qual foram medidas as energias das



transições de 276 keV, 303 keV, 356 keV e 384 keV do  $^{133}\text{Ba}$  (KU85).

TABELA 4.1 - Energias e desvios padrões para as transições gama do  $^{133}\text{Ba}$ .

	E(keV) ( $\sigma$ (eV))	
(a)	(b)	(c)
53,161(1)	53,1632(1,0)	53,16230(0,92)
79,623(10)	79,61154(0,50)	79,61135(0,48)
80,997(3)	80,99744(0,48)	80,99770(0,46)
160,613(8)	160,6089(1,0)	160,60899(0,95)
223,234(12)	223,2396(1,5)	223,24164(0,98)
276,398(2)	276,4028(1,3)	276,40384(0,84)
302,853(1)	302,8510(1,4)	302,85284(0,79)
356,017(2)	356,0141(1,3)	356,01501(0,89)
383,851(3)	383,8483(1,4)	383,85034(0,85)

(a) LO85  
 (b) MU87a - sem incorporar dados de KU85  
 (c) MU87a - com dados de KU85

A matriz de correlação dos resultados apresentados na terceira coluna da tabela 4.1 é mostrada na tabela 4.2.

Esse exemplo de espectroscopia gama do  $^{133}\text{Ba}$  mostra como o uso de vínculos pode aumentar a quantidade

TABELA 4.2 - Matriz de correlação entre as energias das transições gama do  $^{133}\text{Ba}$  (MU87).

	52	80	81	161	223	276	303	356	384
52	1,0								
80	0,5	1,0							
81	0,5	1,0	1,0						
161	0,5	1,0	1,0	1,0					
223	-0,6	-0,6	-0,6	-0,6	1,0				
276	0,4	-0,2	-0,2	-0,2	0,5	1,0			
303	-0,5	-0,1	-0,1	-0,1	0,9	0,5	1,0		
356	0,6	0,4	0,4	0,4	0,2	0,9	0,4	1,0	
384	0,2	0,4	0,4	0,4	0,5	0,4	0,8	0,6	1,0

de informação em cada parâmetro independente (energia da transição): a quantidade de parâmetros independentes é igual à totalidade de parâmetros medidos menos a quantidade de equações de vínculo. Além disso, no trabalho citado, o uso de equações de vínculo permitiu artifícios instrumentais, basicamente a realização de diversas medidas curtas com diferentes arranjos experimentais, já testada anteriormente

(HE83b), que colaboraram no aumento de informação. O aumento da informação ocorrido nesse trabalho é ilustrado na figura 4.1. A figura 4.2 mostra as energias dos níveis do  $^{133}\text{Cs}$  determinadas no mesmo trabalho.

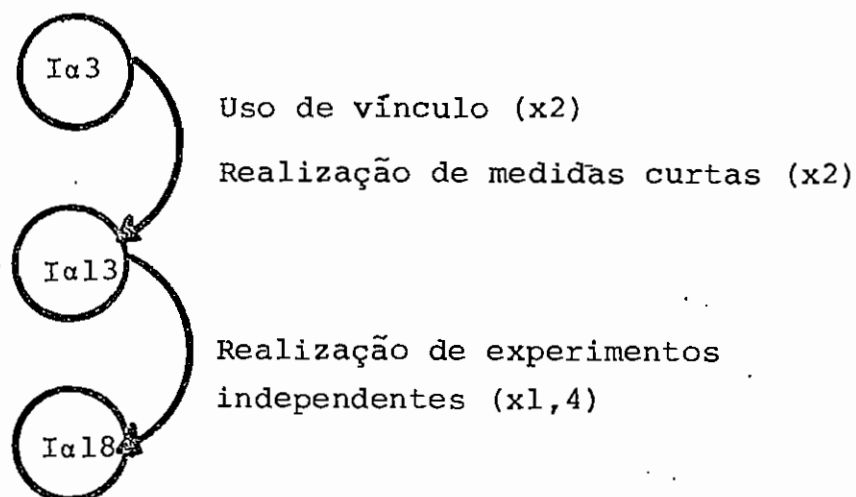


Figura 4.1 - Variação da quantidade de informação nas energias das transições gama do  $^{133}\text{Cs}$  conseguidas com o uso de vínculos bem como com o uso de artifícios instrumentais permitidos e facilitados pelo uso de técnicas de ajuste com vínculo.

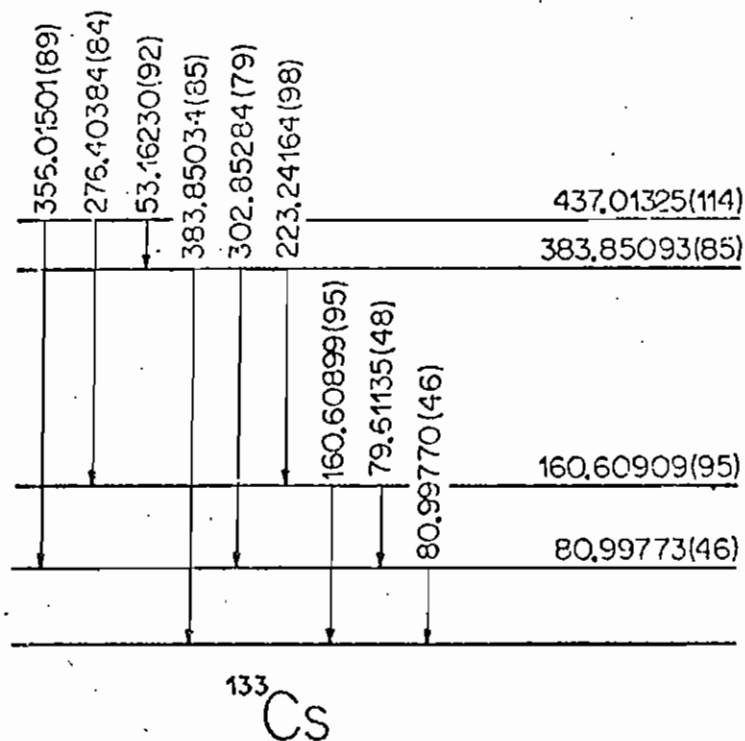


Figura 4.2 - Níveis e transições gama do  $^{133}\text{Cs}$  (MU87a).

**b) Uso de covariâncias**

O uso de covariâncias na análise de dados experimentais é indispensável se se quer manter o correto

conteúdo probabilístico dos resultados ou usá-los para o cálculo de outras grandezas físicas deles dependentes. Mannhart (MA81) mostrou a importância disso na análise de resultados de seção de choque, onde os dados são fortemente correlacionados devido aos arranjos experimentais usualmente adotados. No item a acima mostramos a importância do uso de vínculos em análises de dados, o que só pode ser feito corretamente se consideradas as matrizes de covariância em sua totalidade e não apenas os elementos diagonais, o que é infelizmente bastante comum.

Outro ponto bastante importante no uso e preservação das matrizes de covariância se refere à atualização dos resultados experimentais. Nova medida de uma única grandeza física pode alterar todas as outras com as quais ela está estatisticamente correlacionada (e não obrigatoriamente fisicamente ligada) bem como seus respectivos desvios padrões. Por exemplo, suponhamos duas grandezas -- funcionalmente independentes entre si -- estimadas experimentalmente como  $a$  e  $b$  e com a matriz de covariância

$$V = \begin{pmatrix} \sigma^2 & \text{cov}(a,b) \\ \text{cov}(a,b) & \sigma^2 \end{pmatrix} \quad (4.9)$$

Se  $b$  for medido posteriormente em outro experimento e de forma independente de  $a$ , obtendo-se os resultados  $b'$  e  $\sigma'_b$ , tanto o valor estimado  $a$  como sua variância terão que ser alterados. A alteração levaria a novos valores  $a$  e  $b$  dados por

$$\bar{a} = \frac{1}{1-\rho^2} \left| \left(1 - \frac{\rho^2}{1+\alpha}\right) (a-\rho b) + \frac{\rho\alpha}{1+\alpha} (b-a\rho) + \frac{b'\rho(1-\rho^2)}{1+\alpha} \right| \quad (4.10)$$

e

$$\bar{b} = \frac{1}{1-\rho^2} \left| \frac{\rho\alpha}{1+\alpha} (a-\rho b) + \frac{\alpha}{1+\alpha} (b-\rho a) + \frac{b'(1-\rho^2)}{1+\alpha} \right| \quad (4.11)$$

onde  $\alpha = \sigma'^2/\sigma^2$  e  $\rho = \text{cov}(a,b)/\sigma^2$ . Caso  $\sigma'^2 \ll \sigma^2$ , teríamos

$$\bar{a}' = a + \rho(b-b') \quad (4.12a)$$

$$\bar{b}' = b' \quad (4.12b)$$

com

$$\sigma_{\bar{a}'}^2 = \alpha^2(1-\rho^2) \quad (4.13)$$

As expressões 4.12 e 4.13 mostram quão significativas podem ser as alterações nos valores estimados das grandezas quando são realizados novos experimentos que não as medem diretamente mas que medem outras grandezas com as quais elas estão correlacionadas. E as determinações dessas alterações só pode ser feita de forma completa quando se conhece as matrizes de covariância.

No que diz respeito ao experimento citado no item a, pretende-se medir de forma precisa algumas das transições do  $^{133}\text{Cs}$ , como o dubleto 80-81 keV e eventualmente algumas transições mais fracas. O conhecimento da matriz de covariância entre as energias das transições gama do  $^{133}\text{Cs}$ , mostrada na tabela 4.2 permitirá a atualização correta das energias não medidas nesses novos experimentos.

Está implantado no computador PDP-11 do Laboratório do Acelerador Linear um programa para o método dos mínimos quadrados com formalismo matricial, o qual permite o uso e geração de covariâncias bem como imposição de condições de vínculos (lineares) de forma simples. Essa técnica foi recentemente usada em um experimento de espectroscopia gama do  $^{57}\text{Co}$  (SC87) com sucesso.

## 4.2 - PLANEJAMENTO DE UM EXPERIMENTO

Está em andamento no Laboratório do Acelerador Linear o estudo da transição  $0^+ \rightarrow 0^+$  no  $^{90}\text{Zr}$  pela emissão de dois gamas (MU87b) (veja figura 4.3). Nesse processo de decaimento as energias dos dois gamas emitidos obedecem às

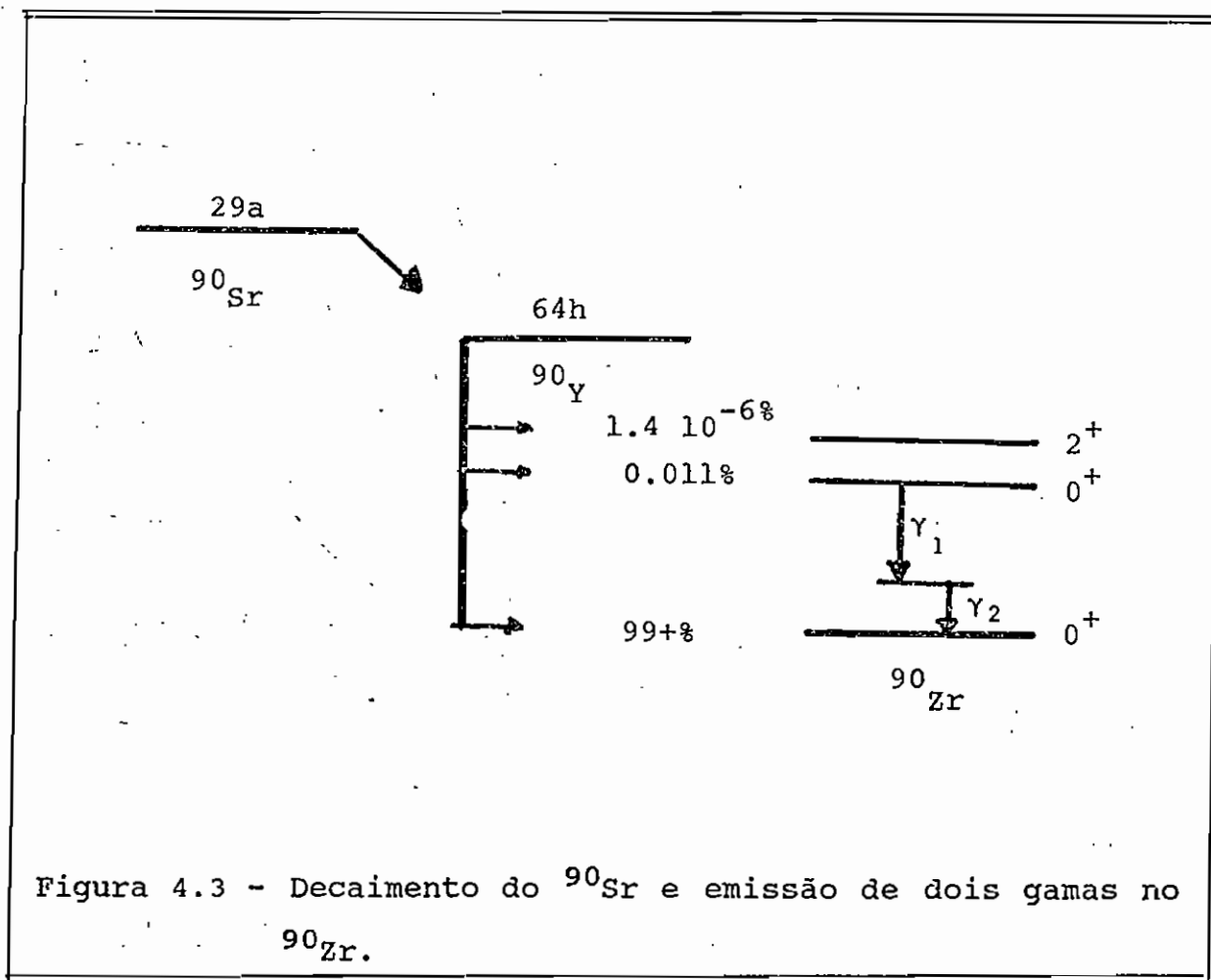


Figura 4.3 - Decaimento do  $^{90}\text{Sr}$  e emissão de dois gamas no  $^{90}\text{Zr}$ .



equações,

$$\begin{aligned} E_{\gamma_1} &\leq \Delta E \\ E_{\gamma_2} &\leq \Delta E \\ E_{\gamma_1} + E_{\gamma_2} &= \Delta E, \end{aligned} \tag{4.14}$$

onde  $\Delta E$  é a diferença energética entre os dois níveis  $0^+$  do  $^{90}\text{Zr}$ .

Alguns nuclídeos oferecem situações experimentais privilegiados para a observação desse fenômeno por apresentarem esquemas de desintegração com poucos gamas e um estado excitado  $0^+$  que não pode decair por um único gama para o nível fundamental, também  $0^+$ . Entre eles estão o  $^{90}\text{Zr}$ , o  $^{16}\text{O}$  e o  $^{40}\text{Ca}$ . No entanto até hoje o fenômeno não havia sido observado de forma não ambígua quando estudado a partir de decaimentos beta que alimentam os níveis  $0^+$  excitados desses nuclídeos (AS73), sendo o efeito identificado apenas em experimentos nos quais o estado  $0^+$  excitado era formado por reação (p,p') (WA75, BE73, SC84).

No experimento ora em andamento é usado um arranjo experimental equivalente àquele de Asano e Wu (AS73). Pretende-se com esse projeto confirmar resultados anteriores e estudar alguns pontos ainda não esclarecidos (SC84). Como o fenômeno é extremamente fraco, da ordem de  $10^{-4}$  do pro-

cesso de formação de pares, foram necessários diversos cuidados experimentais com o objetivo de maximizar-se a quantidade de informação contida nos dados. A figura 4.4 mostra um espectro de energia com o pico correspondente à soma das energias dos dois gamas emitidos ( $E_{\gamma_1} + E_{\gamma_2} = 1761$  keV) identificado de forma não ambígua.

Discussão detalhada dos cuidados experimentais são encontrados na referência MU87b. Entre os artifícios que permitiram a observação pela primeira vez do fenômeno com o uso de uma fonte estática está a colocação de "janelas" em espectros de energia e tempo, discutida a seguir.

A necessidade de separação de tipos diferentes de eventos em muitos experimentos em física implica frequentemente na colocação de janelas em espectros multicanais, separando sinais que se manifestam sob a forma de um pico gaussiano de um fundo constante, como ilustrado na figura 4.5. Quando o objetivo é determinar a intensidade do sinal, basicamente a área do pico, a precisão final do resultado depende da largura da janela a qual, para permitir máxima precisão, depende não só da largura do pico (que caracterizaremos pela sua variância  $\sigma^2$ ) mas também da relação entre sua área total,  $A_0$ , e o fundo,  $b_0$  (em eventos por canal). Admitiremos no que segue que  $b_0$  seja conhecido com erro desprezível, o que corresponde a sua estimativa do

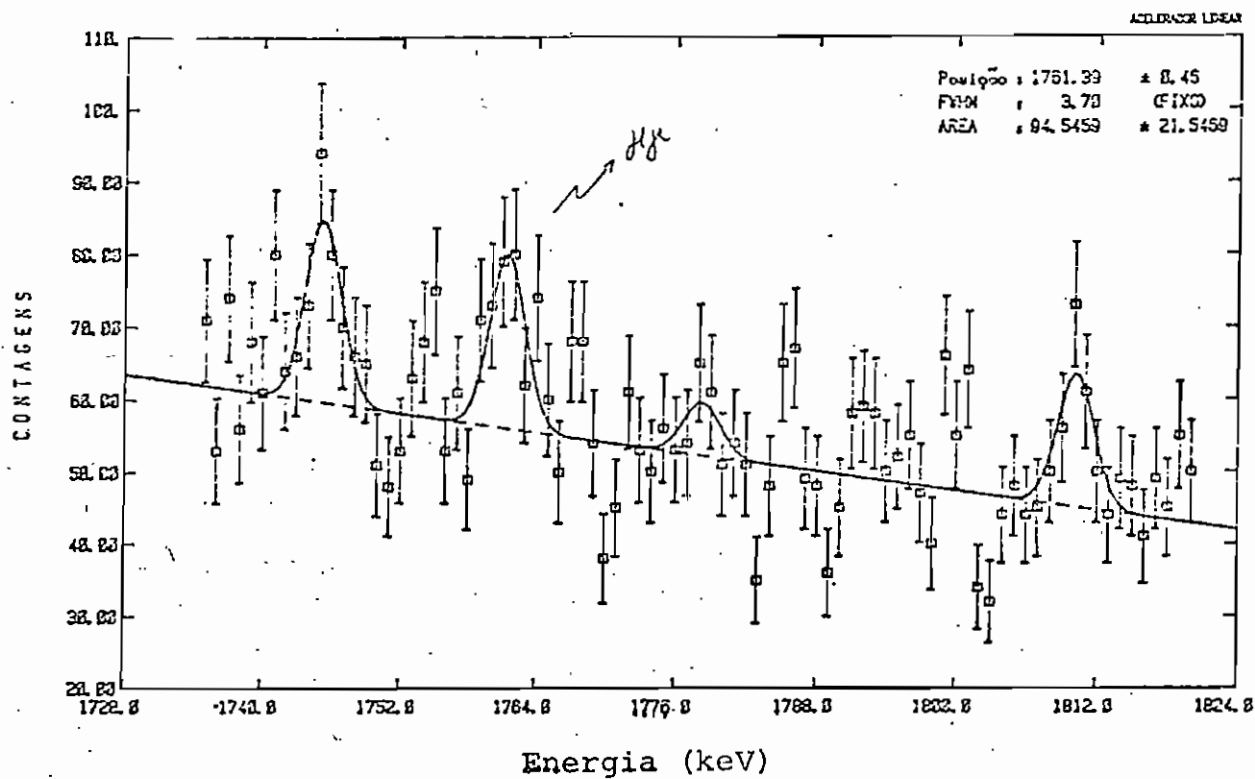


Figura 4.4 - Espectro correspondente a soma da energia de dois gamas em coincidência onde o pico de 1761 keV corresponde a transição  $0^+ \rightarrow 0^+$  no  $^{90}\text{Zr}$  pela emissão de duplo gama (reproduzida na referência MU87b).

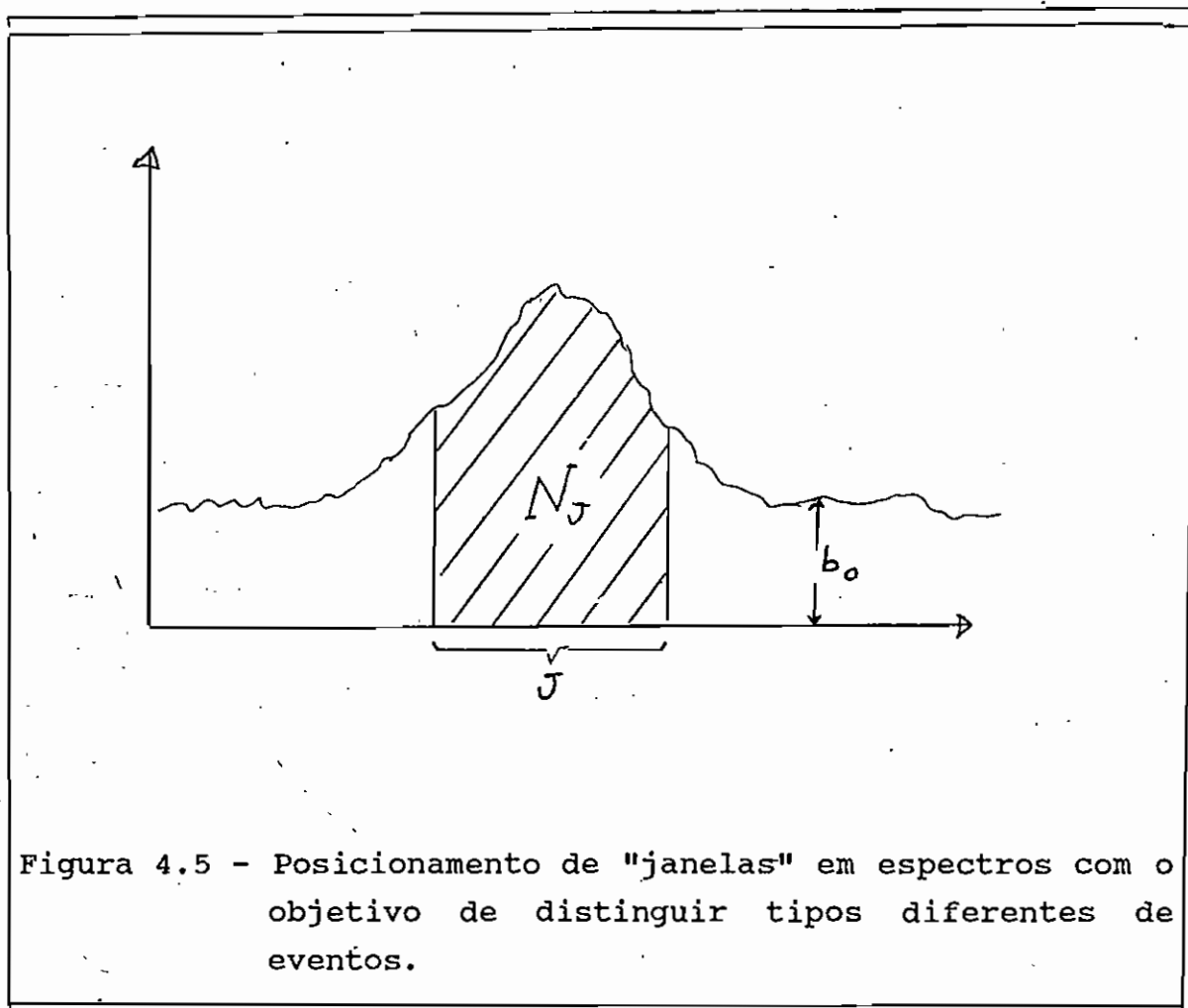


Figura 4.5 - Posicionamento de "janelas" em espectros com o objetivo de distinguir tipos diferentes de eventos.

fundo a partir de uma região suficientemente grande em relação a janela  $J$ .

A probabilidade de  $N_J$  eventos na região da janela  $J$  é dada por

$$P(N_J) = \frac{e^{-(A_0 \alpha_J + b_0 J)} (A_0 \alpha_J + b_0 J)^{N_J}}{N_J!} \quad (4.15)$$

onde  $\alpha_J$  é a integral de uma gaussiana normalizado representando o pico sobre a região da janela. Sendo observado  $N_J$  eventos na região,

$$A = \frac{N_J - b_0 J}{\alpha_J} \quad (4.16)$$

é um estimador não tendencioso de  $A_0$ ,  $E(A) = A_0$ , de mínima variância. A quantidade de informação contida nos dados da região  $J$ , calculada por intermédio de 2.11, é

$$I_J = \frac{\alpha_J^2}{A_0 \alpha_J + b_0 J} \quad (4.17)$$

dependente portanto de  $J$ . Assim dado  $A_0$  e  $b_0$  há uma janela ideal que permite a maximização de  $I_J$ . A tabela 4.3 mostra

os valores de  $I_J$  para, várias relações entre  $A_0$  e  $b_0$ . Por exemplo para  $b_0/A_0 = 0,1$  a janela ideal corresponde a aproximadamente 4,5 desvios padrões do pico (ou duas vezes a largura à meia altura). No entanto para  $b_0/A_0 = 10$  a janela ideal quando se deseja estimar  $A_0$  é cerca de 3 desvios padrões (1,3 vezes a largura à meia altura). Usar neste último caso a janela ideal para  $b_0/A_0 = 0,1$  implica em perda de cerca de 20% da informação máxima possível. Ou ainda, o

TABELA 4.3 - Quantidade de informação na estimativa da área de um pico em função da janela para alguns valores de  $b_0/A_0$ .

$J(\sigma)$	$b_0/A_0$				
	10	5	1	0,2	0,1
1,0	0,027	0,051	0,166	0,304	0,339
2,0	0,044	0,082	0,277	0,528	0,596
3,0	0,047	0,090	0,317	0,643	0,738
4,0	0,043	0,083	0,308	0,672	0,789
5,0	0,038	0,072	0,280	0,656	0,788
6,0	0,032	0,062	0,249	0,622	0,766

uso dessa janela inadequada implicaria em prolongar a experiência por um tempo 25% maior do que aquele necessário se fosse adotada a janela correta.

Procedimentos de escolha rigorosa de janelas -- além de outras melhorias instrumentais -- permitiram a observação clara do decaimento por duplo gama do  $^{90}\text{Zr}$  em um experimento cuja duração foi de 8 dias, o que não havia sido conseguido anteriormente em um experimento equivalente com duração de 60 dias (AS73). A performance atingida deverá permitir o estudo da distribuição angular e de energia do decaimento por duplo gama do  $^{90}\text{Zr}$ .

#### 4.3 - ALTERAÇÃO DE CALIBRAÇÃO EM ESPECTROS MULTICANAIS

A aquisição de dados sob a forma de espectros multicanais, tanto uniparamétricos como biparamétricos, é bastante usual em física experimental. Menos usual mas não raro é a alteração de calibração da abscissa (número do canal). Essa alteração pode decorrer, por exemplo, da

necessidade de comparação, de soma ou subtração de espectros diferentes para posterior análise .

Um espectro multicanal é basicamente um histograma cujas ordenadas correspondem a quantidade de eventos em intervalos da abscissa. Nesse caso o usual é que essas ordenadas obedecem a distribuições de Poisson. Assim num processo de alteração de calibração da abscissa é desejável que, além da conservação do número de eventos, os novos canais contenham eventos que obedecem à distribuições de Poisson.

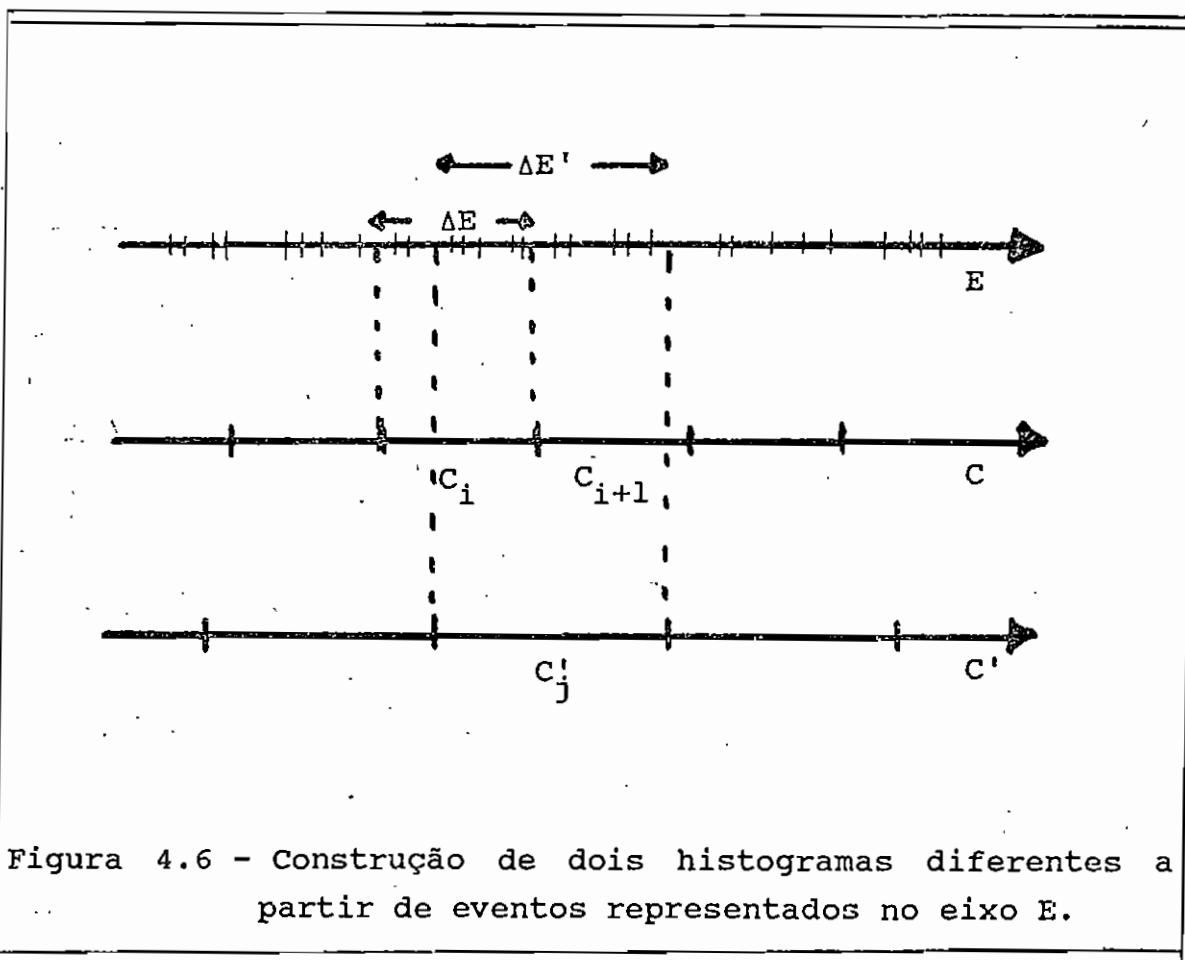
Nesta seção discutiremos como proceder a uma alteração de calibração preservando essas propriedades e ilustraremos com alguns exemplos onde isso não é feito, avaliando suas consequências.

#### a) Alteração de calibração

A figura 4.6 ilustra o esquema de um arranjo instrumental que elabora um espectro multicanal (unidimensional) a partir de eventos distribuídos continuamente. O eixo E representa a grandeza física observada, distribuída continuamente. O eixo C corresponde a uma conversão discreta de eventos contínuos. Assim no canal  $C_i$  representado na figura haverá um número de eventos igual



ao número de eventos na região  $\Delta E$ . O eixo  $C'$  representa um outro histograma com calibração diferente da de  $C$ . Assim no canal  $C'_j$  deve haver uma quantidade de eventos igual aquela observada no intervalo  $\Delta E'$  do eixo  $E$ .



No entanto não é usualmente conhecida a distribuição dos eventos no eixo  $E$ , mas apenas o espectro  $C$ .

Dessa forma o canal  $C'_j$  receberá uma quantidade de eventos correspondente a uma fração  $p$  dos eventos observados em  $C_i$  e outro,  $q$ , dos observados em  $C_{i+1}$ . Há portanto dois problemas distintos aqui: como determinar essas frações e, conhecendo-as, como calcular o número de eventos em  $C'_j$ .

É impossível determinar-se exatamente aquelas frações  $p$  e  $q$ . No entanto muitas vezes a forma de espectros multicanais não apresenta variações rápidas do número de eventos em função do canal. É possível então aproximar  $p$  e  $q$  -- e de uma maneira geral as frações de ocupação dos canais no espectro original pelos canais nos espectros relocados -- a partir do ajuste de funções polinomiais pelo espectro original. A quantidade de canais utilizados nesse ajuste bem como o grau do polinômio devem depender da forma e da intensidade (número de eventos por canal) do espectro original. No momento estão em andamento testes das limitações dos fatores de relocação em função desses parâmetros. No que segue vamos supor que  $p$  e  $q$  sejam conhecidos com precisão suficiente.

Supondo conhecidos  $p$  e  $q$  podemos discutir a questão de como proceder à alteração de calibração. Vamos supor que o espectro original seja uniforme e portanto que  $p$  e  $q$  sejam proporcionais à fração de ocupação de um canal do espectro original por um canal do espectro relocado.

Usando a convenção implícita na figura 4.6 e sabendo que o valor esperado e a variância de uma distribuição de Poisson são ambos iguais ao parâmetro dessa distribuição, temos

$$E(N_i) = E(N_{i+1}) = a \quad (4.18)$$

$$E(N_i^2) - a^2 = E(N_{i+1}^2) - a^2 = a, \quad (4.19)$$

onde  $E$  indica o valor esperado e  $a$  é o valor verdadeiro da média do número de eventos em cada canal do espectro  $C$ .

Uma possibilidade seria construir o espectro  $C'$  de forma determinística. Assim ao canal  $C'_j$  indicado na figura 4.6 seriam atribuídas

$$N'_i = p N_i + q N_{i+1} \quad (4.20)$$

contagens, não obrigatoriamente um número inteiro. Neste caso teríamos

$$E(N'_i) = E(pN_i + qN_{i+1}) = a(p+q) \quad (4.21)$$

e

$$E(N'_i{}^2) - (a(p+q))^2 = a(p^2 + q^2) < E(N_i). \quad (4.22)$$

Ou seja, essa última equação mostra que a variância do número de eventos em cada canal do novo espectro é menor do que deveria ser no caso de uma distribuição de Poisson. Assim a relocação determinística compromete a flutuação estatística esperada quando se analisa espectros multicanais.

A distribuição correta -- no sentido de preservar a flutuação estatística esperada --, quando  $p$  e  $q$  são conhecidos, é a binomial. Neste caso ao canal  $C'_j$  seriam atribuídas  $n_1$  contagens do canal  $C_i$  segundo a distribuição binomial

$$P_{N_i, p}(n_1) = \frac{N_i!}{(N_i - n_1)! n_1!} p^{n_1} (1-p)^{N_i - n_1}, \quad (4.23)$$

mais  $n_2$  contagens do canal  $C_{i+1}$  distribuídos também como uma binomial, neste caso com parâmetros  $N_2$  e  $q$ . A probabilidade de se observar  $n_1$  contagens no canal  $C'_j$  é dada por

$$P(n_1) = \sum_{N_i=n_1}^{\infty} P_a(N_i) \cdot P_{N_i, p}(n_1) \quad (4.24)$$

onde  $P_a(N_1)$  corresponde a uma distribuição de Poisson com parâmetro  $a$ . A partir da equação 4.24 é direto mostrar que

$$P(n_1) = \frac{e^{-ap} (ap)^{n_1}}{n_1!} \quad (4.25)$$

ou seja, o número de eventos no canal  $C'_j$  oriundos do canal  $C_i$  obedece a uma distribuição de Poisson com parâmetro  $a.p$ . Da mesma forma o número de eventos oriundos de  $C_{i+1}$  obedece a uma distribuição de Poisson com parâmetro  $a.q$ . Conseqüentemente a quantidade total de eventos em  $C'_j$  obedece a uma distribuição de Poisson com parâmetro  $a(p+q)$ .

Assim, desde que  $p$  e  $q$  sejam conhecidos o novo espectro apresentará as propriedades esperadas, não sendo perdida qualquer informação.

Está em implantação no computador do Laboratório do Acelerador Linear um programa para proceder a relocação de espectros de forma binomial. A publicação de um guia de uso desse programa e de suas limitações -- basicamente devidas a dificuldades na determinação dos parâmetros de probabilidade das distribuições binomiais -- deverá ser feita oportunamente (MU88). Testes preliminares indicam que é possível a relocação de espectros

gama como usualmente adquiridos por espectroscopistas sem comprometimento da qualidade dos resultados. Além do uso geral o programa será útil em experimentos de espectroscopia gama de alta precisão (como a melhor determinação das energias das transições gama do  $^{133}\text{Cs}$ ). Estudos e correções de não linearidades integral e diferencial de sistemas de espectroscopia gama e de variações incontrolláveis de ganho, já iniciadas anteriormente (HE83b, MU87), serão facilitadas por aquele programa. Da mesma forma medidas de baixa atividade, das quais o duplo decaimento gama é um exemplo, também serão facilitadas, uma vez que variações temporais de calibração poderão ser corrigidas a posteriori, não comprometendo assim a resolução dos sistemas detectores.

**b) Alteração de calibração de espectros: o duplo decaimento beta do  $^{76}\text{Ge}$ .**

O decaimento do  $^{76}\text{Ge}$  por dupla emissão beta sem neutrino (figura 4.7) tem sido estudado por diversos grupos experimentais.

No caso do decaimento para o estado fundamental  $0^+$  do  $^{76}\text{Se}$  o arranjo experimental mais comum consiste na observação de espectros de energia obtidas com detectores de Ge -- onde o  $^{76}\text{Ge}$  está presente numa proporção de cerca de

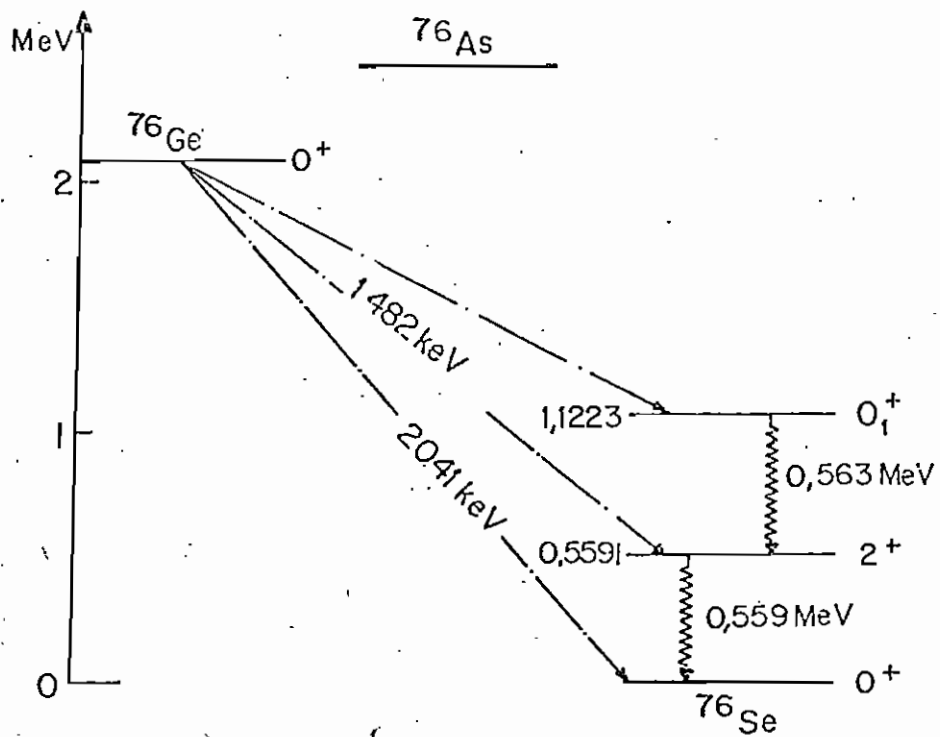


Figura 4.7 - Esquema possível para o decaimento  $\beta\beta$  do  $^{76}\text{Ge}$ .

8% -- em condições extremamente favoráveis de fundo radioativo. Para que essas condições sejam atingidas são usadas blindagens passivas, sistemas de anti-coincidência, laboratórios subterrâneos e seleção especial de materiais para fabricação dos detectores. As referências AV87, AV83, BE84, CA85, FO84 e SI84 descrevem com detalhes experimentos recentes do duplo decaimento  $\beta$   $0^+ \rightarrow 0^+$  do  $^{76}\text{Ge}$ . O sinal procurado que indicaria a existência dessa transição é um pico na energia de  $Q_{\beta\beta} = 2041$  keV, correspondente a deposição total da energia dos dois elétrons emitidos no detector.

Como em geral o número típico de eventos por canal nessa região, apesar da duração bastante longa dos experimentos, da ordem de ano, é muito pequena, é frequente uma padronização na apresentação dos resultados. Essa padronização corresponde a espectros com ganho de 1 keV/canal. Várias vezes entretanto essa padronização é feita a posteriori, de forma determinística. Como consequência as variâncias das ordenadas deixam de ser iguais, em média, aos valores das próprias ordenadas, como seria no caso de distribuições de Poisson, por um fator aproximadamente igual à fração de ocupação média de um canal original por um canal relocado. Isso pode ser visto a partir da equação 4.22 onde,





$$V(N') = a(p^2 + q^2), \quad (4.26)$$

$$E(N') = a(p + q) \quad (4.27)$$

e portanto, supondo  $\bar{p} = \bar{q} = 0,5$  e  $\overline{p^2} = \overline{q^2} = 1/3$ ,

$$\frac{V(N')}{E(N')} \cong \frac{2}{3}. \quad (4.28)$$

Assim os espectros relocados não apresentam a característica esperada de eventos que obedecem a distribuição de Poisson.

As referências CA85, AV83 e AV87 apresentam figuras de espectros aparentemente relocados de forma determinística. Supondo espectros uniformes nas regiões próximas a 2041 keV, obtemos valores para relações  $V(N)/E(N)$ , determinadas pelos pontos experimentais, mostrados na tabela 4.4. A proximidade desses valores com a estimativa da equação 4.28 indica que esses espectros tiveram de fato suas calibrações alteradas a posteriori. Os espectros mostrados nesses trabalhos não apresentam, por inspeção visual, qualquer estrutura que possa indicar um pico na energia de 2041 keV. No entanto é conveniente proceder a um

TABELA 4.4 - Flutuação das ordenadas e médias de alguns espectros de  $\beta\beta$   $0^+ \rightarrow 0^+$  do  $^{76}\text{Ge}$ .

Ref.	E(N)	V(N)	V(N)/E(N)	N <sub>pontos</sub>	P( $\chi^2$ )
CA85	3,2 $\pm$ 0,3	1,9	0,6	38	5%
AV83	31,7 $\pm$ 1,0	13	0,4	26	0,5%
AV87	14,7 $\pm$ 0,8	9,6	0,7	21	10%

teste estatístico objetivo para se decidir quanto a inexistência ou não de irregularidades dos espectros. A última coluna da tabela 4.4 mostra as probabilidades de testes de  $\chi^2$  associadas à hipótese de não regularidade dos espectros: como consequência do procedimento de relocação determinística dos espectros, todos eles são incompatíveis com essa hipótese, até mesmo mais incompatíveis do que se esperaria se de fato não há picos na região. De fato essa conclusão não é correta, mas apenas uma consequência do processo de relocação, que teve como efeito impossibilitar esse teste estatístico. Os espectros relocados tornaram-se mais regulares do que deveriam ser. Caso não haja picos nas regiões analisadas P( $\chi^2$ ) deveria ser da ordem de 50%;

havendo picos  $P(\chi^2)$  seria muito próximo de 100%. O processo de relocação determinística faria com que um teste de  $\chi^2$  pudesse indicar compatibilidade com a hipótese de não haver picos na região mesmo que houvesse.

Apresentamos em apêndice B análise do resultado do experimento de decaimento por duplo beta sem neutrino  $^{76}\text{Ge}(0^+) \rightarrow ^{76}\text{Se}(2^+)$  realizado pela colaboração Bordeaux - Zaragoza (M087). Na parte B.2 desse apêndice são apresentados resultados de testes estatísticos sobre o espectro original e sobre os espectros relocados deterministicamente e binomialmente. Uma comparação desses resultados mostra que a relocação determinística pode comprometer a conclusão quando se procede a um teste do tipo  $\chi^2$ .

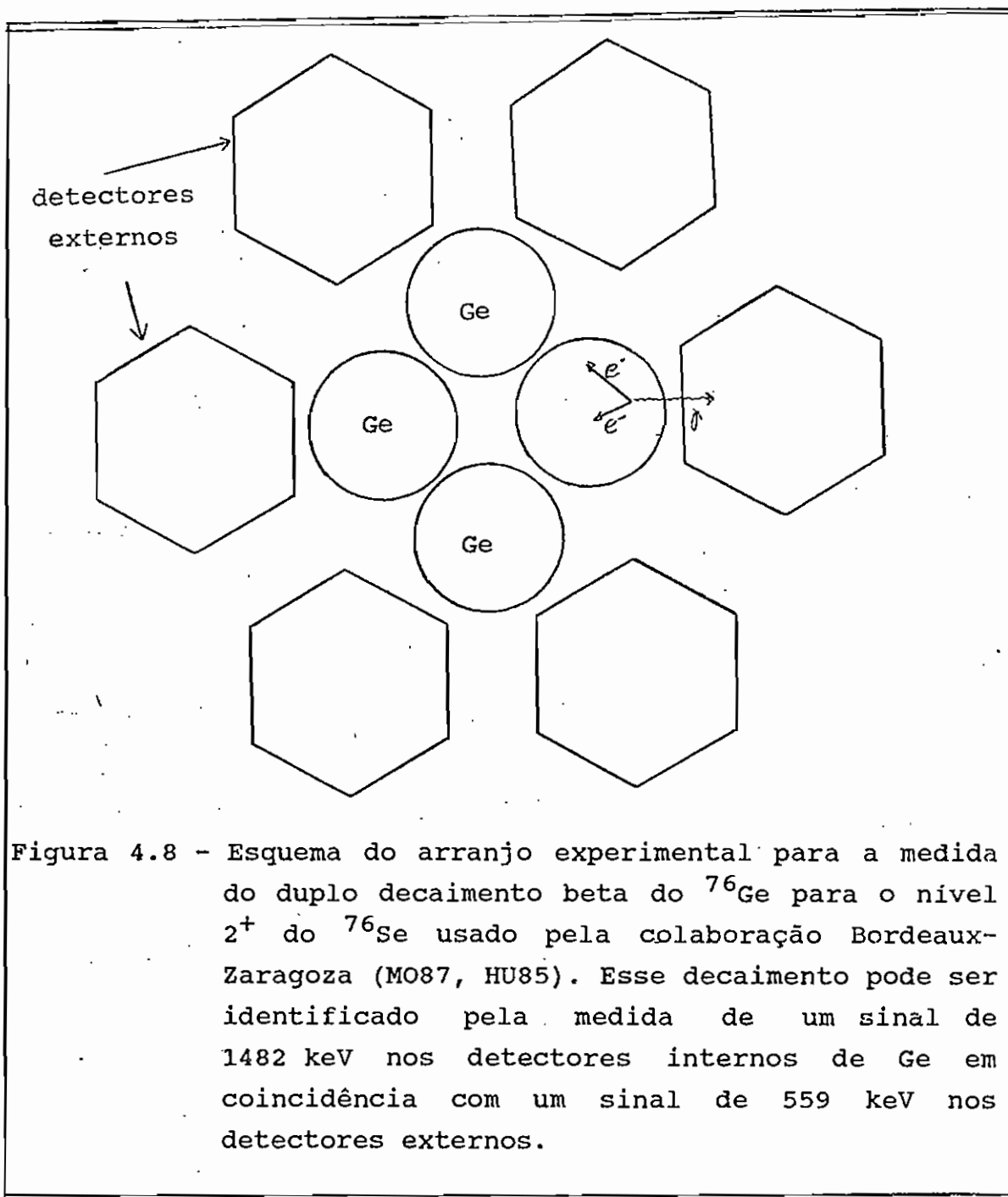
#### 4.4 - QUANTIDADE DE INFORMAÇÃO EM DIFERENTES ARRANJOS EXPERIMENTAIS

Na seção 4.2 discutimos como certos cuidados experimentais podem ser significativos no que diz respeito a maximizar, para dada duração do experimento, a quantidade de informação contida na estimativa da grandeza medida. Nesta seção compararemos diferentes arranjos experimentais para o

estudo do duplo decaimento beta sem neutrino do  $^{76}\text{Ge}$  para o nível  $2^+$  do  $^{76}\text{Se}$ .

Um possível arranjo experimental para o estudo desse decaimento está mostrado de forma esquemática na figura 4.8 (as referências M087 e HU85 contém os detalhes desse arranjo experimental). No caso da alimentação do nível  $2^+$  do  $^{76}\text{Se}$  os dois betas emitidos pelo  $^{76}\text{Ge}$  carregam uma energia total igual à diferença entre  $Q_{\beta\beta} = 2040,9 \pm 0,6 \text{ keV}$  (EL85) e a energia de excitação do primeiro nível  $2^+$  do  $^{76}\text{Se}$ ,  $E_{\gamma} = 559,11 \pm 0,05 \text{ keV}$  (LE78):  $E_{\beta\beta} = 1481,8 \pm 0,6 \text{ keV}$  (veja figura 4.7). No arranjo experimental esquematizado na figura 4.8 a evidência do decaimento em estudo é dada por um sinal correspondente à energia de 1481,8 keV nos detectores internos de Ge -- uma vez que os elétrons produzidos no decaimento perdem toda sua energia dentro do detector -- e outro correspondente a 559,11 keV nos detectores externos, simultaneamente. Esses sinais aparecem sobre um fundo radioativo.

O resultado experimental correspondente a um tempo de contagem de 6200h (M087, ME87) está mostrado no apêndice. Esse resultado foi obtido com o uso de detectores externos NaI. Nesse experimento foi observado um excesso



de cerca de 12 contagens na região onde seria esperado o sinal de 1481,8 keV, observado nos detectores de Ge, em

coincidência com uma janela em 559 keV nos detectores externos NaI. O fundo na região, no espectro biparamétrico projetado sobre os detectores de Ge é de 3,9 contagens por keV (cf. apêndice B). Como esse sinal é bastante fraco, sendo a energia ajustada na projeção Ge ligeiramente incompatível com a energia esperada a partir de  $Q_{\beta\beta}$  -- embora dados recentes de  $Q_{\beta\beta}$  sejam também ligeiramente incompatíveis entre si (cf. p. ex. EL85) -- há dúvidas quanto a corresponder ele ao decaimento estudado (M087, ME87). Essa dúvida implica no prosseguimento do experimento, preferivelmente em melhores condições. Assim vamos discutir aqui alguns dos possíveis projetos de continuidade.

No arranjo experimental usado os detectores externos tinham uma eficiência total de 60% para a detecção do gama de 559 keV, aí incluída a probabilidade que esses gamas saiam dos detectores internos de Ge. Os detectores internos eram quatro detectores HPGe de cerca de 100 cm<sup>3</sup> cada um. A tabela 4.5 mostra na coluna A alguns dados relevantes desse primeiro experimento. No que segue vamos supor que os 12 eventos observados em excesso em 6200h correspondem de fato ao decaimento estudado.

Um possível projeto de continuação do experimento (ME87) é substituir os quatro detectores internos de 100 cm<sup>3</sup>

TABELA 4.5 - Comparação entre a performance de três arranjos experimentais possíveis para a medida do duplo decaimento beta do  $^{76}\text{Ge}$  sem neutrino, alimentando o nível  $2^+$  do  $^{76}\text{Se}$ .

	Atual	Ge(e)-NaI	Ge(e)-Ge
	(A)	(B)	(C)
Eficiência dos detectores externos	60%	60%	6%
Volume equivalente de $^{76}\text{Ge}$ nos detectores internos ( $\text{cm}^3$ )	30	100	100
Fundo no espectro bidimensional na região 1480 keV x 559 keV (contagens/ano x $\text{keV}^2$ )	0,08	0,020	0,0020
Fundo sob 1,2 FWHM com janelas de 1,2 FWHM (contagem/ano)	16	4,0	0,015
Área do sinal (contagens/ano)	18	60	6
Variância da energia nos detectores internos ( $\sigma_i^2$ , $\text{keV}^2$ )	0,10	0,018	0,17
Variância da energia nos detectores externos ( $\sigma_e^2$ , $\text{keV}^2$ )	68	11	0,17
Variância na área sobre área quadrada ( $\sigma_A^2/A^2$ )	0,10	0,018	0,17

cada por um único detector de  $100 \text{ cm}^3$  de germânio enriquecido no isótopo 76. Os detectores externos continuariam sendo do tipo NaI. A segunda coluna da tabela 4.5 mostra as características deste possível arranjo experimental (B), bem como aquelas do experimento concluído (A).

Uma outra possibilidade aqui especulada é a substituição dos detectores externos de NaI por detectores HPGe. A desvantagem dessa troca é a perda de eficiência, de 60% para cerca de 6% na detecção do gama de 559 keV do  $^{76}\text{Se}$ . No entanto há duas vantagens nessa troca. A menor eficiência dos detectores externos implica numa diminuição da área do pico procurado mas também do fundo. Além disso detectores do tipo HPGe tem melhor resolução do que detectores NaI e, portanto, permitem a escolha de janelas mais estreitas e melhor definição da energia dos sinais.

Para escolher entre essas duas possibilidades de continuação do experimento é necessário saber o que se quer medir para identificar de forma não dúbia o fenômeno estudado.

A identificação do duplo decaimento beta nestes experimentos é feita pela energia observada nos detectores internos, que deve ser comparada com o valor esperado de



1481,8 ± 0,6 keV, e nos detectores externos, comparada com o valor 559,11 ± 0,05 keV. Identificado o fenômeno, a precisão na estimativa da meia vida do decaimento depende da variância relativa na área do pico. Como o fenômeno não foi até agora observado de maneira clara, não é relevante a precisão na determinação de sua meia vida, mas apenas a sua identificação.

Neste estágio a informação relevante é portanto aquela proveniente da comparação das energias observadas com as energias esperadas. Assim

$$I_1 = \frac{1}{\sigma_1^2 + \sigma_i^2} \quad (4.29)$$

corresponde a informação na medida do sinal de 1481 keV, sendo  $\sigma_i^2$  a variância da medida desse sinal no experimento em questão (nos detectores internos) e  $\sigma_1^2 = 0,36 \text{ keV}^2$  a variância correspondente ao valor hoje aceito para a energia da transição medida. Da mesma forma

$$I_2 = \frac{1}{\sigma_2^2 + \sigma_e^2} \quad (4.30)$$

é a informação contida na comparação da energia do sinal nos detectores externos (com variância  $\sigma_e^2$ ) e o valor aceito  $559,11 \pm 0,05$  keV, sendo portanto  $\sigma_2^2 = 2,510^{-3}$  keV<sup>2</sup>.

O duplo decaimento beta do  $^{76}\text{Ge}$  é igualmente caracterizado pelas energias medidas nos detectores internos e externos: em ambos está se caracterizando a formação do primeiro nível excitado do  $^{76}\text{Se}$ . Assim podemos adotar como a informação total no resultado.

$$I = I_1 + I_2 = \frac{1}{\sigma_1^2 + \sigma_i^2} + \frac{1}{\sigma_2^2 + \sigma_e^2} \quad (4.31)$$

Outra possibilidade seria comparar a soma das energias observadas nos detectores internos e externos com o valor experimental  $Q_{\beta\beta}$ . Neste caso a informação seria

$$I' = \frac{1}{\sigma_i^2 + \sigma_e^2 + \sigma_{\beta\beta}^2} \quad (4.32)$$

onde  $\sigma_{\beta\beta}^2$  é a variância do valor experimental de  $Q_{\beta\beta}$  (EL85).

Os valores  $\sigma_i^2$  e  $\sigma_e^2$  associados a cada um dos três arranjos estudados são mostrados na tabela 4.5. Essas variâncias foram determinadas a partir das resoluções dos detectores usados, às áreas dos sinais esperados e considerando o uso de janelas ideais de acordo com a discussão da seção 4.2. Foi também suposto que os 12 eventos observados com o arranjo experimental A correspondem ao fenômeno estudado.

Os valores de  $I_1$ ,  $I_2$ ,  $I$  e  $I'$  correspondentes aos três arranjos possíveis são mostrados na tabela 4.6. Os dados mostrados nessa tabela indicam que qualquer que seja a escolha para a decisão sobre ser ou não o pico observado correspondente ao duplo decaimento beta,  $I$  ou  $I'$ , o arranjo experimental C permite uma quantidade de informação significativamente maior do que o arranjo B. Isso deve-se basicamente ao fato que o arranjo B permite uma melhoria na informação sobre o sinal obtido nos detectores externos muito menor do que aquela conseguida com o arranjo C. Quanto a informação relativa à medida do sinal nos detectores internos, não há diferenças significativas entre os três arranjos considerados.

TABELA 4.6 - Quantidade de informação em três arranjos possíveis para a medida do duplo decaimento beta do  $^{76}\text{Ge}$ .

	ARRANJO		
	A	B	C
$I_1$	2,17	2,65	1,89
$I_2$	0,015	0,091	5,90
I	2,19	2,74	7,79
$I'$	0,015	0,088	1,43

A opção entre I ou  $I'$  para caracterizar o resultado desses experimentos parece irrelevante. Sob o aspecto físico, as comparações que dão origem a I (comparação das energias dos elétrons e do gama associado ao nível  $2^+$  do  $^{76}\text{Se}$ ) e a  $I'$  (comparação de  $Q_{\beta\beta}$  com a soma das energias medidas) são equivalentes no que diz respeito a identificar o sinal com o duplo decaimento beta sem neutrino do  $^{76}\text{Ge}$  para o nível  $2^+$  do  $^{76}\text{Se}$ . Sob o ponto de vista estatístico ambos parecem também equivalentes por

apresentarem o mesmo comportamento para as situações analisadas.

Até aqui admitimos que os cerca de 12 eventos observados com o arranjo A em uma medição de 6200h correspondiam de fato ao decaimento em questão. No entanto não está descartada a hipótese de serem esses eventos de outra origem. Caberia portanto comparar as performances dos arranjos B e C nas hipóteses de diferentes meias vidas para o duplo decaimento beta e, conseqüentemente, outras durações do experimento. A tabela 4.7 mostra as quantidades de informação para os arranjos C (linhas superiores) e B (linhas inferiores). Nessa tabela  $\alpha$  representa a relação entre o número de eventos realmente correspondentes ao duplo decaimento beta e o número de eventos que se admitiu corresponder a esse decaimento nos cálculos anteriores. Assim apenas para situações extremamente desfavoráveis,  $\alpha = 0,1$ , o arranjo B é melhor que o arranjo C. No entanto, se o tempo de contagem é suficientemente longo, maior que 2 anos para  $\alpha = 0,1$  ou maior que à 0,5 ano no caso  $\alpha = 0,5$ , o arranjo C torna a ser melhor que o arranjo B.

Na tabela 4.7 estão indicados com asteriscos as situações nas quais a probabilidade de que não ocorra nenhum evento no caso do arranjo C é maior ou cerca de 10%. Esta possibilidade deve ser levada em consideração, especialmente

TABELA 4.7 - Quantidade de informação nos resultados dos arranjos B (linhas inferiores) e C (linhas superiores) supondo diferentes atividades para o decaimento  $\beta\beta$  do  $^{76}\text{Ge}$  e diferentes durações do experimento. As situações que podem não apresentar nenhum evento com probabilidade significativa estão assinaladas por \*.

	DURAÇÃO (ANOS)			
$\alpha$	0,5	1,0	2,0	4,0
0,1	0,56*	1,07*	1,99*	5,0*
	1,10	1,57	2,02	2,36
0,5	2,46*	4,40	7,77	13,8
	2,31	2,55	2,73	2,89
1,0	4,41	7,79	13,9	25,1
	2,57	2,74	2,89	3,12
2,0	7,8	13,9	25,1	45,4
	2,74	2,90	3,12	3,52

no caso de valores de extremamente baixos.

Podemos concluir portanto que o uso de detectores de Ge em substituição aos detectores de NaI é, a menos de hipóteses bastante desfavoráveis, vantajosa. Mesmo nos casos em que o uso de detectores externos NaI são mais vantajosos,  $\alpha = 0,1$  com duração menor ou igual a 2 anos, a quantidade de

informação total no arranjo B é menor do que a quantidade de informação no arranjo A, correspondente ao experimento de 6200 h completado recentemente. Neste caso a continuação do experimento não trará novas informações sobre o fenômeno estudado, pelo menos se mantidas as outras condições experimentais.

## 5. CONCLUSÃO

Neste trabalho foram discutidos alguns aspectos sobre tratamento de dados e planejamento e análise de experimentos em casos extremos em física. As situações analisadas correspondiam a experimentos com poucos eventos, quer em termos absolutos, quer em comparação com a flutuação estatística do fundo, assim como experimentos com muitos eventos.

No primeiro caso, poucos eventos, foram estudados aspectos relativos a determinação de intervalos de confiança e ao planejamento de experimentos. A determinação de intervalos de confiança para a distribuição de Poisson foi estudada comparando-se diferentes abordagens estatísticas. A abordagem bayesiana com o uso de uma função densidade de probabilidade a priori uniforme mostrou-se em melhor acordo com as expectativas em física experimental. Esse resultado está de acordo com resultado anterior (ES87), mostrado no apêndice A, no qual comparou-se as abordagens bayesiana e clássica. Pode-se então afirmar que, pelo menos no caso da distribuição de Poisson, é lícito e desejável a indicação de



intervalos de confiança mesmo em experimentos com pouquíssimos eventos, sendo recomendada a abordagem bayesiana cujos intervalos "curados" estão indicados na tabela 3.4. Esse estudo deverá ser estendido ao caso da diferença e razão entre pequenos números que obedecem a distribuições de Poisson.

A quantidade de informação, discutida no capítulo 2, foi usada no capítulo 4 no planejamento de experimentos e comparação entre diferentes arranjos experimentais. As conclusões apresentadas mostram-se de acordo com o que usualmente consideramos em física como precisão de um experimento. Entre os exemplos abordados estava a colocação de "janelas" de discriminação em espectros multicanais. Foi mostrado neste caso que as janelas dependem não apenas da largura do pico criado pelo sinal que se deseja discriminar, mas também da relação entre as intensidades do sinal e do ruído de fundo. Uma janela mal posicionada pode comprometer significativamente o resultado de um experimento. Como exemplo de aplicação desse procedimento foi citado o estudo do duplo decaimento gama do nível excitado  $0^+$  do  $^{90}\text{Zr}$ , ora em andamento no Laboratório do Acelerador Linear e que exige o estabelecimento de duas janelas, uma em tempo e outra em energia, com o objetivo de discriminar eventos de origens diferentes (além do uso de filtros de absorção gama, que

também exige o mesmo tipo de tratamento). Neste caso foi possível a identificação clara do fenômeno, pela primeira vez em experimento estático (MU87b), usando-se arranjos experimentais equivalentes a outros já usados anteriormente.

Outra situação estudada com base no conceito de quantidade de informação foi o experimento do duplo decaimento beta do  $^{76}\text{Ge}$  para o nível  $2^+$  do  $^{76}\text{Se}$ . Neste caso foi comparada a performance de dois arranjos experimentais onde a diferença básica era o uso de detectores semicondutores de Ge ou detectores NaI para identificar o gama de 559 keV correspondente ao decaimento de nível  $2^+$  do  $^{76}\text{Se}$ . Neste caso como a evidência do decaimento em questão envolve a observação de duas grandezas distintas -- energia dos elétrons e do gama emitidos -- estudou-se duas possibilidades para a quantidade de informação gerada no experimento. Uma delas era baseada na identificação das duas energias independentemente e a outra na identificação da soma dessas energias com o valor atual de  $Q_{\beta\beta}$ . Essas duas possibilidades para determinação da quantidade de informação apresentaram o mesmo comportamento, indicando ambas ser o uso de detectores de Ge mais convenientes que o uso de detectores NaI.

No caso de experimentos com muitos eventos o desvio padrão pode se tornar menor que eventuais erros

sistemáticos. Foi mostrado que a inclusão na análise de dados de vínculos entre as grandezas medidas contribui para a redução desses erros sistemáticos assim como pode facilitar o uso de artifícios instrumentais que têm esse mesmo efeito. Esse procedimento permitiu a obtenção de precisões na determinação de energias de transições gama com o uso de detectores Ge(Li) ou HPGe comparáveis aquelas obtidas em espectroscopia com o uso de cristais curvos. O uso de vínculos bem como a adoção de artifícios experimentais baseados na realização de muitas medidas curtas com diferentes arranjos (HE83b, MU87a), cuja análise é bastante facilitada pelo uso do método dos mínimos quadrados com vínculos lineares, pode abrir perspectivas significativas na espectroscopia gama com detectores semicondutores.

De forma geral é necessário o uso das covariâncias entre os resultados experimentais para que se preserve, na manipulação de dados, o conteúdo probabilístico dos intervalos indicados. Particularmente no caso de experimentos com muitos eventos, com pequenos desvios padrões e vínculo entre as grandezas medidas, o uso das covariâncias se mostra essencial. Além disso o uso de covariâncias é fundamental para a atualização de estimativas e variâncias de grandezas correlacionadas com outras que são

posteriormente medidas de forma independentes, como mostrado na seção 4.1.b.

Foi discutida ainda a manipulação de espectros multicanais, especialmente quanto à preservação da flutuação estatística. A manipulação determinística de espectros quando se procede a uma alteração de calibração (da abscissa) compromete a flutuação estatística do número de eventos em cada canal e, portanto, pode inviabilizar um teste estatístico. É mostrado no capítulo 4 como proceder a uma alteração de calibração não determinística. O apêndice B mostra um espectro multicanal sobre o qual se procedeu a alterações de ganho determinística e estatística, comparando-se então o resultado de testes de  $\chi^2$  sobre os dois espectros criados. O resultado desse teste no caso do espectro determinístico mostrou-se em desacordo com o resultado do teste no espectro original, o mesmo não acontecendo no caso do espectro relocado estatisticamente. A relocação estatística de espectros por meio de um programa de computador, ora em implantação no Laboratório do Acelerador Linear, deverá viabilizar alguns trabalhos de espectroscopia gama de alta precisão (muitos eventos) bem como trabalhos que envolvam baixa estatística e portanto demandam tempos longos de medida com a consequente variação

de ganho do sistema detector ou mesmo a troca de equipamentos.

A conclusão geral é que o tratamento e a interpretação cuidadosa dos dados de um experimento permitem uma significativa melhora na quantidade e no aproveitamento da informação por ele gerada. Essa melhora pode tornar realizável um experimento que não seria conclusivo de outra forma. Além disso pode-se obter uma significativa redução do tempo necessário para a observação de um efeito com a escolha adequada do arranjo experimental.

## APÊNDICE A

No capítulo 3 comparamos o comportamento de intervalos de confiança para a distribuição de Poisson determinadas a partir de diversas abordagens estatísticas diferentes. Em trabalho recente (ES87) havíamos feito uma comparação experimental de intervalos clássicos e bayesianos com distribuição a priori uniforme tanto para a distribuição de Poisson como para a razão entre números que obedecem a distribuições de Poisson diferentes. O resultado daquela comparação é que a abordagem bayesiana está em melhor acordo com as expectativas (i) e (ii) discutidas no capítulo 3. Neste apêndice apresentamos um resumo daquele trabalho.

### A.1 - INTERVALOS DE CONFIANÇA

#### a) Distribuição de Poisson

A determinação de intervalos de confiança para a distribuição de Poisson no esquema clássico é mostrada nas equações 3.6. Aquelos intervalos correspondem a intervalos

centrais. É possível definir-se também intervalos mínimos com mesmo conteúdo probabilístico  $1-\alpha_1 - \alpha_2 = 1-\alpha$  fazendo

$$\alpha_1 = \sum_{m=n}^{\infty} P_{a_-}(m) \quad (\text{A.1a})$$

$$\alpha_2 = \sum_{m=0}^n P_{a_+}(m) \quad (\text{A.1b})$$

com  $(a_+ - a_-)$  mínimo, onde  $n$  é o número de eventos observados.

A tabela A.1 mostra os limites mínimos  $a_-$ ,  $a_+$  para  $1-\alpha_1 - \alpha_2 = 0,683$  que foram usadas no teste experimental da referência ES87.

Os intervalos bayesianos mínimos com densidade a priori uniforme aparecem na tabela 3.1.

## **b) Razão entre pequenos números**

### **b1) Abordagem clássica**

Suponha que  $N$  e  $D$  sejam dois resultados de medidas de grandezas que obedecem a distribuições de Poisson com parâmetros  $\nu$  e  $\delta$  respectivamente. A probabilidade conjunta

TABELA A.1 - Limites inferiores e superiores (mínimos) para a distribuição de Poisson no esquema clássico para  $1 - \alpha = 0,683$ .

n	$a_-$	$a_+$
0	0,00	1,15
1	0,00	2,50
2	0,27	3,86
3	0,86	5,15
4	1,55	6,40
5	2,29	7,63
6	3,06	8,84
7	3,85	10,0
8	4,65	11,2
9	5,47	12,4
10	6,30	13,5
12	7,99	15,8
14	9,70	16,1
16	11,4	20,4
18	13,2	22,6
20	15,0	24,8



para  $N = n$  e  $D = d$  é

$$P(n, d) = \frac{e^{-v} v^n}{n!} \frac{e^{-\delta} \delta^d}{d!} \quad (\text{A.2})$$

Seja  $p$  a razão entre  $v$  e  $\mu$ ,  $\mu$  sendo o valor esperado de  $M = N + D$  ( $\mu = v + \delta$ ),

$$p = \frac{v}{\mu}, \quad 1-p = \frac{\delta}{\mu} \quad (\text{A.3})$$

A equação A.2 pode ser escrita então na forma

$$\begin{aligned} P(n, d = M-n) &= e^{-p\mu} \frac{(p\mu)^n}{n!} e^{-(1-p)\mu} \frac{((1-p)\mu)^d}{d!} = \\ &= \frac{e^{-\mu} \mu^M}{M!} \left( \frac{M!}{n!(M-n)!} p^n (1-p)^{M-n} \right). \end{aligned} \quad (\text{A.4})$$

Supondo  $M$  fixo e  $p$  conhecido, a probabilidade de observar-se  $n$  eventos é

$$P_{M,p}(n) = \frac{M!}{n!(M-n)!} p^n (1-p)^{M-n}. \quad (\text{A.5})$$

A partir dessa equação pode-se determinar os limites para  $p$  de forma análoga à equação A.1:

$$\alpha_1 = \sum_{n=0}^N P_{M, p_+}(n) \quad (\text{A.6a})$$

$$\alpha_2 = \sum_{n=N}^{\infty} P_{M, p_-}(n) \quad (\text{A.6b})$$

onde  $1 - \alpha_1 - \alpha_2$  é o conteúdo probabilístico do intervalo  $(p_-, p_+)$  mínimo.

Os limites clássicos correspondentes para  $R = v/\delta$ ,  $(R_-, R_+)$ , podem ser calculados a partir das equações A.3, que relacionam  $R$  e  $p$ :

$$R_{\pm} = \frac{p_{\pm}}{1-p_{\pm}} \quad (\text{A.7})$$

## b2) Abordagem Bayesiana

No esquema bayesiano a função densidade de probabilidade para  $R = v/\delta$  é dada por (ES87, HE84),

$$f_p(R) = \frac{(N+D+1)! R^N}{N! D! (1+R)^{N+D+2}} \quad (\text{A.8})$$

supondo-se funções densidade de probabilidade a priori para  $\nu$  e  $\delta$  uniformes. Os limites mínimos ( $R_-$ ,  $R_+$ ) são determinados de forma análoga às equações 3.8 e 3.10.

A tabela A.2 mostra os limites inferiores e superiores, para os esquemas clássico e bayesiano (uniforme), dos intervalos mínimos para a razão  $N/D$ .

#### A.2 - TESTE EXPERIMENTAL

A comparação experimental dos intervalos clássico e bayesiano foi feita a partir da realização de experimentos com poucos eventos. Um dos experimentos correspondeu a medida de energia de prótons de 1,5 MeV espalhados em alvos de silício. O outro experimento correspondeu a medidas de energia gama de fontes típicas de calibração com detectores HPGe. Em ambos os casos os eventos foram armazenados em espectros multicanais. Ambos os espectros foram obtidos em duas situações diferentes: uma medida longa

TABELA A.2 - Limites mínimos para a razão N/D para  $1-\alpha = 0,683$  nas abordagens clássicas (tabela superior) e bayesiana (tabela inferior).

N	0	1	2	3	4	5	6	7	8	9	10
D											
1	0.000	0.000	0.038	0.150	0.284	0.428	0.576	0.727	0.879	1.033	1.187
	2.155	4.762	7.381	10.011	12.648	15.286	17.926	20.567	23.207	25.848	28.489
2	0.000	0.000	0.037	0.137	0.255	0.380	0.508	0.639	0.771	0.905	1.039
	0.776	1.683	2.602	3.516	4.426	5.333	6.240	7.145	8.049	8.953	9.856
3	0.000	0.000	0.034	0.122	0.224	0.333	0.445	0.558	0.673	0.789	0.906
	0.467	1.008	1.556	2.097	2.635	3.170	3.704	4.237	4.769	5.301	5.832
4	0.000	0.000	0.031	0.108	0.198	0.294	0.392	0.492	0.593	0.696	0.798
	0.333	0.718	1.108	1.492	1.872	2.249	2.626	3.002	3.376	3.746	4.121
5	0.000	0.000	0.028	0.097	0.177	0.262	0.350	0.439	0.529	0.620	0.712
	0.258	0.557	0.860	1.157	1.450	1.741	2.031	2.320	2.608	2.896	3.183
6	0.000	0.000	0.026	0.088	0.160	0.236	0.315	0.395	0.477	0.559	0.641
	0.211	0.456	0.703	0.945	1.183	1.420	1.655	1.890	2.123	2.356	2.589
7	0.000	0.000	0.024	0.080	0.145	0.215	0.286	0.359	0.433	0.508	0.583
	0.178	0.385	0.594	0.798	0.999	1.198	1.396	1.593	1.790	1.985	2.181
8	0.000	0.000	0.022	0.074	0.133	0.197	0.262	0.329	0.397	0.465	0.534
	0.154	0.334	0.515	0.691	0.865	1.037	1.207	1.377	1.546	1.715	1.883
9	0.000	0.000	0.020	0.068	0.123	0.181	0.242	0.304	0.366	0.429	0.493
	0.136	0.294	0.454	0.609	0.762	0.913	1.063	1.212	1.361	1.509	1.656
10	0.000	0.000	0.019	0.063	0.114	0.168	0.224	0.282	0.340	0.398	0.457
	0.122	0.263	0.406	0.545	0.681	0.816	0.950	1.082	1.215	1.347	1.478
1	0.000	0.038	0.150	0.284	0.428	0.576	0.727	0.879	1.033	1.187	1.342
	0.776	1.683	2.602	3.516	4.426	5.333	6.240	7.145	8.049	8.953	9.856
2	0.000	0.037	0.137	0.255	0.380	0.508	0.639	0.771	0.905	1.039	1.173
	0.467	1.008	1.556	2.097	2.635	3.170	3.704	4.237	4.769	5.301	5.832
3	0.000	0.034	0.122	0.224	0.333	0.445	0.558	0.673	0.789	0.906	1.023
	0.333	0.718	1.108	1.492	1.872	2.249	2.626	3.002	3.376	3.746	4.121
4	0.000	0.031	0.108	0.198	0.294	0.392	0.492	0.593	0.696	0.798	0.902
	0.258	0.557	0.860	1.157	1.450	1.741	2.031	2.320	2.608	2.896	3.183
5	0.000	0.028	0.097	0.177	0.262	0.350	0.439	0.529	0.620	0.712	0.804
	0.211	0.456	0.703	0.945	1.183	1.420	1.655	1.890	2.123	2.356	2.589
6	0.000	0.026	0.086	0.160	0.236	0.315	0.395	0.477	0.559	0.641	0.724
	0.178	0.385	0.594	0.798	0.999	1.196	1.396	1.593	1.790	1.985	2.181
7	0.000	0.024	0.080	0.145	0.215	0.286	0.359	0.433	0.508	0.583	0.659
	0.154	0.334	0.515	0.691	0.865	1.037	1.207	1.377	1.546	1.715	1.883
8	0.000	0.022	0.074	0.133	0.197	0.262	0.329	0.397	0.465	0.534	0.604
	0.136	0.294	0.454	0.609	0.762	0.913	1.063	1.212	1.361	1.509	1.656
9	0.000	0.020	0.068	0.123	0.181	0.242	0.304	0.366	0.429	0.493	0.557
	0.122	0.263	0.406	0.545	0.681	0.816	0.950	1.082	1.215	1.347	1.478
10	0.000	0.019	0.063	0.114	0.168	0.224	0.282	0.340	0.398	0.457	0.517
	0.110	0.238	0.367	0.493	0.616	0.737	0.858	0.978	1.097	1.216	1.334

na qual podia-se determinar com excelente precisão o número médio de eventos por canal por unidade de tempo; outra medida curta onde o número de eventos por canal era pequeno. Os resultados obtidos nas medidas curtas eram analisadas com base nas abordagens clássica e bayesiana; o resultado dessa

análise era então comparado com o resultado da medida longa, com erro relativo desprezível, podendo ser considerada como o valor verdadeiro da grandeza (desde que corrigido pelas diferenças de tempo de contagem). Em seguida testava-se se o "valor verdadeiro" estava ou não contido nos intervalos indicados pelas duas abordagens. (A referência ES87 apresenta os detalhes da análise e dos experimentos.)

A tabela A.3 mostra o resultado para a distribuição de Poisson. Estão indicados nessa tabela o número de experimentos analisados para cada valor de  $n$  e as frações de vezes que os "valores verdadeiros" das grandezas estavam contidos nos intervalos determinados pelas abordagens clássica e bayesiana. (A classificação dos experimentos em relação ao número de eventos  $n$  foi feita pelo número esperado de eventos no caso dos experimentos de espalhamento de prótons e pelo número observado no caso de espectroscopia gama. Na tabela A.3 mostramos o resultado considerando os dois experimentos. Em ES87 os resultados para os dois experimentos são mostrados separadamente.)

A tabela A.4 mostra a mesma análise para a razão

TABELA A.3 - Número de experimentos analisados e fração (%) de vezes que o "valor verdadeiro" da grandeza estava contido nos intervalos determinados pelos esquemas bayesiano e clássico, para a distribuição de Poisson.

n	Experimentos	Bayesiano	Clássico
1	22160	69,9	79,2
2	1893	82,9	89,4
5	21180	68,0	74,3
10	10435	68,2	73,5
20	2000	68,3	72,5
50	2000	67,0	69,8
Total	59628	69,2	76,2

entre pequenos números. Também aqui foram adicionados os resultados das análises dos dois experimentos.

Em todas as situações analisadas -- espalhamento de prótons e espectroscopia gama, distribuição de Poisson simples e razão de pequenos números -- a abordagem bayesiana mostrou em melhor acordo com o que se espera de um intervalo em física experimental.

TABELA A.4 - Número de experimentos analisados (linhas superiores) e fração (%) de vezes que o "valor verdadeiro" estava contido nos intervalos bayesianos (linhas intermediárias) e clássicos (linhas inferiores) determinadas a partir de N e D.

D	N			
	1	2	3	4
	95215	100712	88606	82034
1	72,3	70,5	68,8	64,7
	95,5	97,6	98,0	98,8
	92546	94656	83734	78710
2	64,2	67,1	62,8	60,3
	82,6	83,2	82,8	81,1
	92346	90726	86974	84498
3	73,8	71,1	64,9	67,8
	83,4	84,0	82,1	83,0
	85470	79457	80939	77656
4	74,1	68,0	67,4	67,0
	81,7	72,4	81,6	80,5
Total - 1.394.356				
	67,9			
	80,1			

### APÊNDICE B

O duplo decaimento beta sem neutrino do  $^{76}\text{Ge}$  para o nível  $2^+$  do  $^{76}\text{Se}$  vem sendo estudado pela colaboração Bordeaux-Zaragoza há alguns anos (M087, HU85). Alguns detalhes desse experimento foram apresentados e discutidos na seção 4.4. Neste apêndice discutiremos especificamente o problema de decisão quanto a ser ou não significativo o excesso de contagens em uma região do espectro bidimensional onde se esperava o sinal do fenômeno estudado em um experimento cuja duração foi de 6200 h.

Como discutido na seção 4.4 a evidência do duplo decaimento beta em questão é a existência de um pico em um espectro bidimensional,  $E(\text{Ge}) \times E(\text{NaI})$ , na região 1481 keV - 559 keV. A tabela B.1 mostra os dados brutos observados nessa região em um experimento de 6200 h de duração (ME87) onde, na região central há um acúmulo de eventos sem haver no entanto a formação de um pico nítido. Seria esse acúmulo apenas fruto de flutuação estatística ou seria de fato um sinal? (A questão que segue naturalmente a



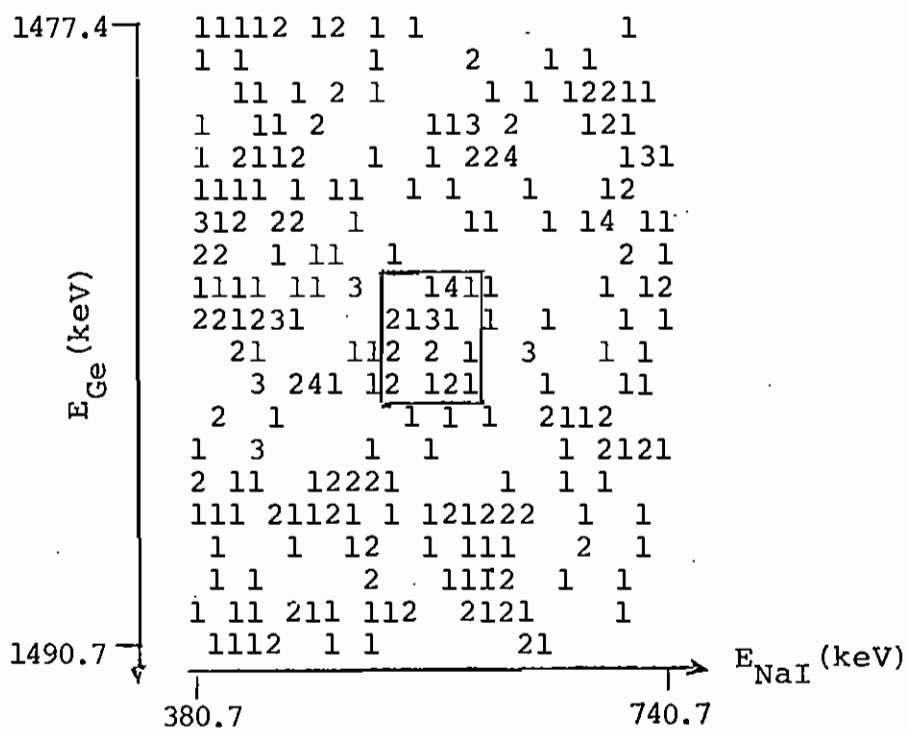


TABELA B.1 - Espectro bidimensional  $E(Ge) \times E(NaI)$  na região  
1484 keV - 559 keV.

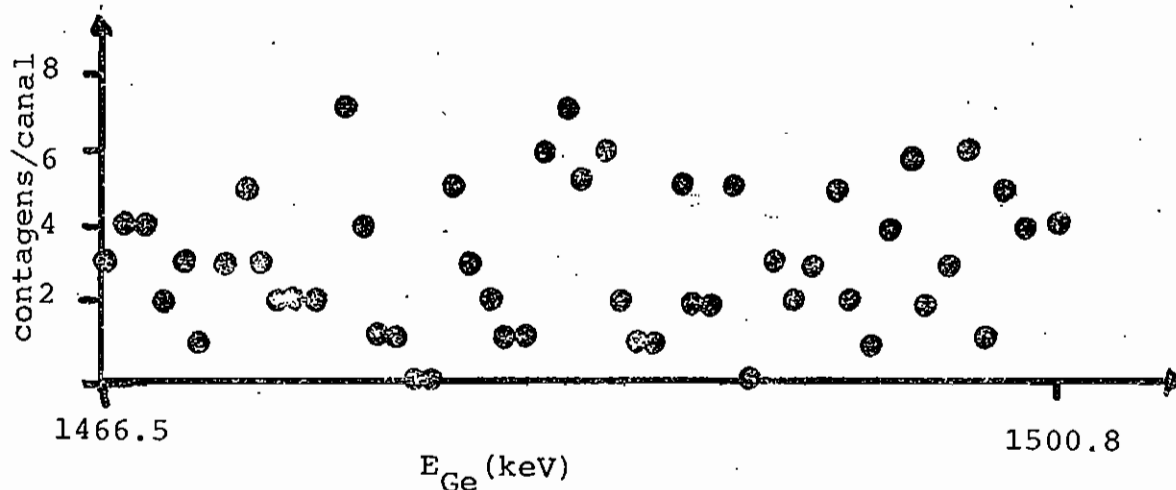


Figura B.1 - Projeção do espectro bidimensional na direção  
dos detectores Ge com janela 525 keV - 595 keV  
nos detectores NaI.

esta, "sendo um sinal, seria evidência do duplo beta estudado ou teria outra origem?", não é discutida aqui.)

Há portanto duas hipóteses a testar: há um pico e não há um pico na região. Vamos atacar esse problema de algumas maneiras diferentes

### B.1 - ANÁLISE DO ESPECTRO BIDIMENSIONAL

Uma primeira possibilidade de análise é a tentativa de ajuste de um pico bidimensional ao espectro com  $\sigma_x^2 = 625 \text{ keV}^2$  (onde o eixo x corresponde à direção NaI, cujos detectores apresentam picos com resolução -- largura a meia altura -- de 60 keV) e  $\sigma_y^2 = 1 \text{ keV}^2$  (correspondente à direção dos detectores de Ge de resolução 2,4 keV). O espectro ajustado era da forma

$$n(x,y) = A+Bx + P e^{-\frac{(x-x_0)^2}{2\sigma_x^2}} e^{-\frac{(y-y_0)^2}{2\sigma_y^2}} \quad (B.1)$$

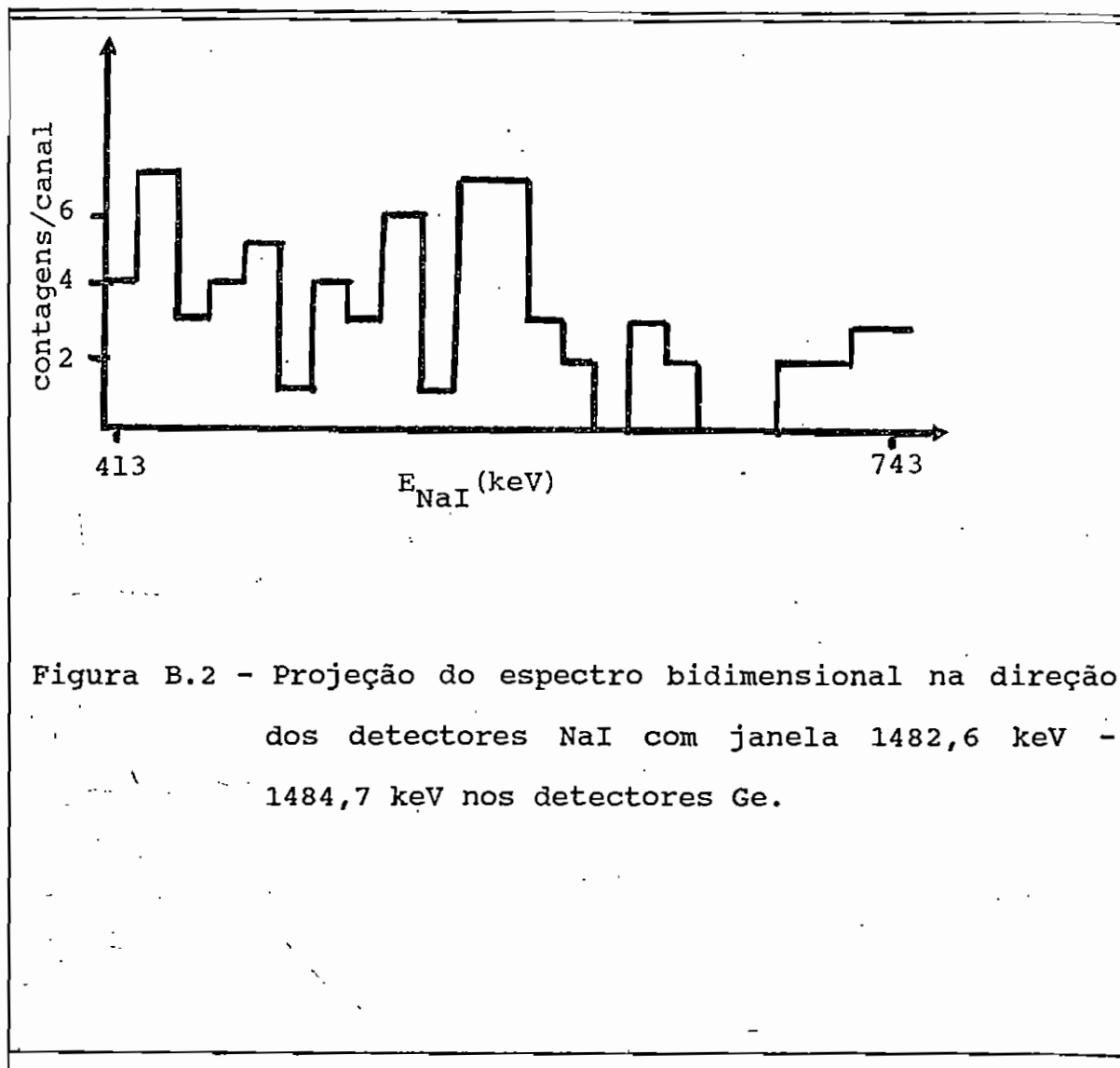


Figura B.2 - Projeção do espectro bidimensional na direção dos detectores NaI com janela 1482,6 keV - 1484,7 keV nos detectores Ge.

correspondendo a um pico bidimensional e um fundo constante na direção y e linear na direção x. As figuras B.1 e B.2

mostram as projeções do espectro bidimensional nas direções dos detectores de Ge e NaI respectivamente, a partir das quais pode-se julgar melhor a adequação ou não da forma escolhida do fundo ajustado.

O ajuste foi feito pelo método da máxima verossimilhança, mais adequada a situação de um espectro com poucos eventos do que o método dos mínimos quadrados: esse método fornece estimadores consistentes e não tendenciosos (KE79) e tem sido usado e testado em situações reais (G087).

Os resultados do ajuste são mostrados na tabela B.2.

TABELA B.2 - Resultado da análise do espectro bidimensional mostrado na tabela B.1.

$$A = 0,749$$

$$B = - 0,0119$$

$$x = 12,79 \pm 0,88 \quad E_{\text{NaI}} = 557 \pm 13 \text{ keV}$$

$$y = 10,46 \pm 0,60 \quad E_{\text{Ge}} = 1484,0 \pm 0,4$$

$$\text{Área} = 11,9 \pm 5,5$$

$$= 530 \text{ (495 graus de liberdade)}$$

$$P(\chi^2) = 14\%$$

A análise dos resultados mostra que a energia observada na direção NaI,  $557 \pm 13$  keV, está em bom acordo com o valor esperado  $559,11 \pm 0,05$  keV (LE78). No entanto o mesmo não ocorre com a energia observada na direção Ge,  $1484,0 \pm 0,4$  keV em comparação ao valor  $1481,8 \pm 0,6$  keV (EL85, LE78) esperado. Um teste t indicaria neste caso um nível de significância de apenas cerca de 0,3%. No entanto, como apontado na seção 4.2, dados experimentais de  $Q_{\beta\beta}$  do  $^{76}\text{Ge}$  disponíveis na literatura (EL85) mostram-se parcialmente incompatíveis entre si; além disso não é nula a possibilidade de uma diferença sistemática entre os padrões de referência para medidas de espectroscopia gama e de diferenças de massa da ordem dos seus próprios desvios padrões. Assim não discutiremos aqui a questão de ser esse o pico procurado mas apenas a questão de constituírem esses eventos um pico ou não.

O nível de probabilidade do  $\chi^2$  obtido no ajuste feito, 14%, é bastante bom, especialmente se considerarmos a existência de outros possíveis picos na região -- e ao observarmos a tabela B.1 vemos que de fato há outras regiões nas quais há acúmulo de contagens -- não incluídos no ajuste. Além disso devemos considerar que o teste de  $\chi^2$  em casos de poucos eventos é apenas uma aproximação (CO54, EA71).

Esse teste poderia ser complementado de duas formas. Uma delas é o ajuste apenas de um espectro de fundo na região. Isto foi feito e o resultado obtido corresponde a um nível de probabilidade de  $\chi^2$  de cerca de 0,01%, mostrando a incompleta inadequação da hipótese: a região não é compatível apenas com um fundo radioativo na forma considerada.

Outra possibilidade seria analisarmos a região fazendo um mapa de probabilidades de  $\chi^2$ . Isso é mostrado na tabela B.3 tanto para a hipótese de existência de pico como para a hipótese de ser o espectro apenas provenientes de um fundo. No primeiro caso o nível de significância relativamente baixo, de 14%, é causado por duas regiões que apresentam níveis baixos de probabilidade 0,1% e 0,3%, sendo que o restante do espectro, inclusive a que contém o excesso de contagens em estudo, apresenta níveis aceitáveis de  $\chi^2$ . O mesmo mapa para a hipótese de fundo mostra que o baixo nível de probabilidade de  $\chi^2$ , cerca de 0,01%, é basicamente comprometida pela região central, onde se localiza o excesso de contagens.

Outra possibilidade de testar-se a hipótese de ser o excesso de contagens observado causado por flutuação estatística é calcular a probabilidade de que apenas como consequência dessa flutuação seja observado na região um

TABELA B.3 - Probabilidades de testes de  $\chi^2$  em regiões do espectro (A) considerando o ajuste de um pico bidimensional e (B) considerando o ajuste apenas de um fundo.

E (Ce)	E (NaI)			
	380,7	500,7	620,7	740,7
1477,4				
1480,1	54%	30%	76%	
1482,7	76%	83%	01%	
1485,4	0,3%	31%	22%	A
1488,0	38%	84%	5%	
1490,7	59%	62%	34%	
1477,4	63%	26%	80%	
1480,1	81%	86%	0,4%	
1482,7	1%	10 <sup>-3</sup> %	29%	B
1485,4	50%	88%	12%	
1488,0	69%	70%	27%	

número igual ou maior de eventos do que a quantidade de eventos que de fato há,

$$P(n \geq n_{\text{obs}}) = \sum_{n=n_{\text{obs}}}^{\infty} \frac{e^{-a} a^n}{n!} \quad (\text{B.2})$$

onde  $a$  é o número médio de eventos esperado na região central de 5x4 canais assinalada na tabela B.1 e  $n_{obs}$  o número de eventos observados nessa região, 26. Supondo não haver pico  $a$  é determinado a partir do número médio de eventos no espectro, 12,5. Neste caso  $P(n > n_{obs}) \cong 0,03\%$ . Supondo haver um pico  $a$  é determinado pelo fundo do espectro ajustado mostrado na tabela B.2. Neste caso  $P(n > n_{obs}) \cong 0,01\%$ . Em ambos os casos os baixos níveis de probabilidade estão de acordo com a análise de  $\chi^2$ .

Outro procedimento de análise possível seria a determinação dos níveis de probabilidade associados a limites superiores da área do pico dados os resultados observados. Uma análise bayesiana, nos moldes da referência HE83c, forneceria os níveis de probabilidades associados aos limites superiores de área do sinal indicados na tabela B.4. Por exemplo, a probabilidade que o pico que originou o excesso de contagens tenha área verdadeira menor que 2 é 0,1%.

A conclusão básica é que todas as análises estatísticas indicam níveis de probabilidade extremamente baixos para a hipótese de não existir um pico na região bem como para a hipótese desse pico ser extremamente fraco, como mostra a tabela B.4.



## B.2) ANÁLISE DA PROJEÇÃO NOS DETECTORES HPGe

A figura B.1 mostra a projeção do espectro bidimensional na direção dos detectores de Ge com uma janela de 70 keV nos detectores NaI em torno da energia de 559 keV. A análise desse espectro leva a resultados inteiramente compatíveis com aqueles do item B.1 acima. O que pretendemos discutir aqui são as consequências de um processo de relocação do espectro, discutido na seção 4.2. As figuras B.3 e B.4 mostram o mesmo espectro da figura B.1 relocados

TABELA B.4 - Probabilidade que a área verdadeira do pico observado seja menor que o limite indicado, determinada a partir de uma análise bayesiana.

Limite superior da área	Probabilidade
1	0,04%
2	0,1%
5	2,6%
10	24%

para um ganho de 1 keV/canal de forma determinística e de forma estatística, respectivamente.

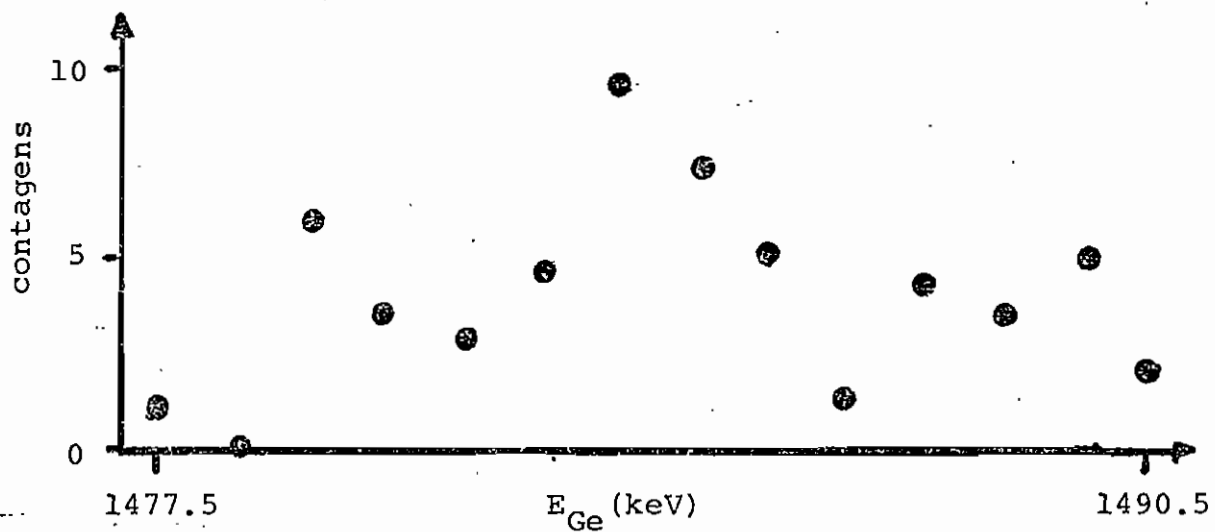


Figura B.3 - O mesmo espectro da figura B.1 com alteração de calibração para 1 keV/canal feita de forma determinística.

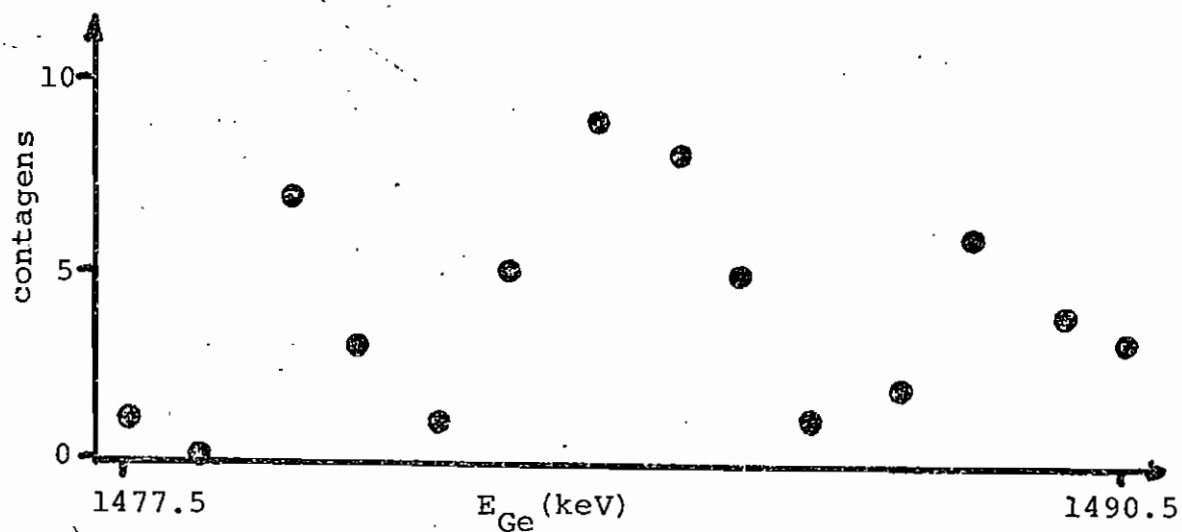


Figura B.4 - A mesma alteração de calibração da figura B.3 mas, neste caso, feita de forma estatística.

Uma possível análise que poder-se-ia fazer em um espectro unidimensional é o de  $\chi^2$  (aproximado) dado por (EA71)

$$\chi^2 = \sum \frac{(n_i - \bar{n})^2}{\bar{n}} \quad (\text{B.3})$$

onde  $n_i$  é o número de eventos no canal  $i$  e  $\bar{n}$  o número médio de eventos na região. (Não se pretende aqui comparar este teste com aqueles apresentados no item anterior. O discutido aqui é mais restrito por ser feito em uma região limitada do espectro -- não considera a região vizinha à janela  $559 \pm 35$  keV nos detectores NaI -- e por incluir automaticamente no cálculo do fundo o próprio pico.)

A tabela B.5 mostra o resultado do teste no qual foi usada uma região de 14 keV em torno da energia 1482 keV. O espectro original indica um nível de probabilidade (associado à hipótese de inexistência do pico) bastante baixo. O espectro relocado deterministicamente apresenta um nível de probabilidade significativamente maior, 5%, a ponto mesmo de comprometer a decisão sobre a existência ou não de um pico.

O espectro relocado estatisticamente preserva as características do espectro original, apresentando um nível

de probabilidade comparável ao apresentado pelo espectro original. Acreditamos que a diferença observada, de 0,5% para 1,5%, seja apenas consequência de flutuação estatística. Os testes do processo de relocação aleatória

TABELA B.5 - Resultados de testes de  $\chi^2$  nos espectros original, relocado deterministicamente e relocado estatisticamente.

Espectro	$P(\chi^2)$
original	0,5%
relocado deterministicamente	5%
relocado aleatoriamente	1,5%

ora em andamento, citado na seção 4.2, poderão esclarecer essa questão.

APÊNDICE C

O uso de vínculos entre parâmetros ajustados pode ter como consequência a redução de erros estatísticos assim como, em certos casos, a redução de eventuais erros sistemáticos.

Suponha uma situação na qual a energia de três transições gama  $E_{10}$ ,  $E_{20}$ ,  $E_{30}$ , que obedecem à relação de vínculo

$$E_{10} = E_{20} + E_{30} \quad (C.1)$$

(que corresponde a uma cascata  $E_{20} - E_{30}$  e a uma transição tipo "cross-over"  $E_{10}$  onde as energias de recuo foram desprezadas) tenham sido medidas como

$$Y = \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} \quad (C.2)$$

sendo

$$V = \sigma^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{C.3})$$

a matriz de covariância desse resultado.

Suponha também que as energias sejam medidas com erros sistemáticos

$$\delta = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix} \quad (\text{C.4})$$

Em espectroscopia gama esses erros sistemáticos são devidos a não linearidade do sistema detector bem como a erros sistemáticos das energias usadas no processo de calibração. Podemos supor que, ao menos aproximadamente

$$\langle \delta_i \rangle = 0 \quad (\text{C.5})$$

$$\text{Cov}(\delta_i, \delta_j) = 0 \quad \text{se } i \neq j \quad (\text{C.6})$$

$$\langle \delta_i^2 \rangle = \sigma_\delta^2$$

para qualquer  $i$ . O símbolo  $\langle \rangle$  corresponde à média sobre todo o espectro de energia. A equação C.5 pode ser

justificada tendo-se em vista o fato que a função de calibração usualmente ajustada segue, em média, os pontos experimentais, sendo os efeitos de não linearidade do sistema de detecção e análise distribuídas uniformemente em torno dela.

Usando o método dos mínimos quadrados com vínculo (equações 4.4 e 4.8) obtemos

$$A = \begin{pmatrix} E_{1v} \\ E_{2v} \\ E_{3v} \end{pmatrix} = 1/3 \begin{pmatrix} 2E_1 + E_2 + E_3 \\ E_1 + 2E_2 - E_3 \\ E_1 - E_2 + 2E_3 \end{pmatrix}, \quad (C.7)$$

sendo

$$V_v = \frac{\sigma^2}{3} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \quad (C.8)$$

a matriz de covariância das energias ajustadas em obediência às condições de vínculo.

Os erros sistemáticos dos elementos de A são dadas por

$$\begin{pmatrix} \delta_{A1} \\ \delta_{A2} \\ \delta_{A3} \end{pmatrix} = 1/3 \begin{pmatrix} 2\delta_1 + \delta_2 + \delta_3 \\ \delta_1 + 2\delta_2 - \delta_3 \\ \delta_1 - \delta_2 + 2\delta_3 \end{pmatrix}. \quad (C.9)$$

A equação C.7 mostra que na estimativa da energia de uma transição são incorporadas informações relativas aos resultados das medidas das outras transições. A consequência disso é a redução da variância de cada uma das energias ajustadas em acordo com a equação de vínculo em relação à variância original da medida,  $\sigma^2$ , de um fator 2/3, como pode ser visto observando-se os valores da diagonal principal de  $V_v$  da equação C.8.

A equação C.9 mostra que o erro sistemático em cada energia ajustada em obediência à condição de vínculo passa a depender dos erros sistemáticos em diferentes regiões do espectro. Pela sua origem estes erros são ao menos parcialmente independentes uns dos outros. Portanto a imposição da condição de vínculo tende a tornar estatístico o efeito de erros sistemáticos. Isso pode ser visto de outra forma uma vez que

$$\langle \delta_{Ai}^2 \rangle = 6/9 \sigma_\delta^2 \langle \delta_i^2 \rangle. \quad (C.10)$$



Ou seja, é mais provável um erro sistemático menor na energia ajustada de acordo com a condição de vínculo do que na energia inicialmente determinada.

REFERÊNCIAS

- AN53 - M. Annis, W. Cheston e H. Primakoff, Rev. Mod. Phys. 25, 818 (1953).
- AS73 - Y. Asano e C.S. Wu, Nucl. Phys. 215, 557 (1973).
- AV83 - F.T. Avignone III et al., Phys. Rev. Lett. 50, 721 (1983).
- AV87 - F.T. Avignone III et al., Phys. Rev. D-35, 1713 (1987).
- BE73 - E. Beardsworth et al., Phys. rev. C8, 216 (1973).
- BE82 - J.P. Benzécri "Histoire et prehistoire de l'analyse des donnees", Dunod, Paris (1982).
- BE84 - E. Bellotti et al., Phys. Lett. 146B, 450 (1984).
- BI87 - R.M. Bionta et al., Phys. Rev. Lett. 58, 1494 (1987).
- CA82 - B. Cabrera, Phys. Rev. Lett. 48 1378 (1982).
- CA85 - D.O. Caldwell et al., Phys. Rev. Lett. 54, 281 (1985).
- CO54 - W.G. Cochran, Biometrics 10, 417 (1954).
- DO77 - J.J. Donn e R.L. Wolke, Health Phys. 32, 1 (1977).

- EA71 - W.T. Eadie et al., "Statistical Methods in Experimental Physics", North-Holland, Amsterdam, (1971).
- EL85 - R.J. Ellis et al., Nucl. Phys. A435, 34 (1985).
- ES87 - B. Escoubes, S. De Unamuno e O. Helene, Nucl. Instr. and Meth. A257, 346 (1987).
- FI50 - R.A. Fisher, "Contributions to Mathematical Statistics", editado por W.A. Shewhart, John Wiley & Sons, New York (1950).
- FO84 - A. Forster et al., Phys. Lett. 138B, 301 (1984).
- GE86 - N. Gehrels, Astroph. J. 44F12, 303 (1986).
- GO83 - P. Gouffon, "Idefix: manual do usuário", Laboratório do Acelerador Linear (1983), não publicado.
- HE80 - O. Helene e V.R. Vanin, "Tratamento Estatístico de Dados em Física Experimental", Edgard Blucher, São Paulo (1980).
- HE82 - O. Helene, "Upper Limit of Peak Area", Publicações IFUSP/P-354 (1982).
- HE83a - O. Helene, "Errors in Experiments with Small Numbers of Events", Publicações IFUSP/P-383 (1983).
- HE83b - O. Helene, "Precise Determination of Photopeak Width", Publicações IFUSP/P-387 (1983).

- HE83c - O. Helene, Nucl. Instr. and Meth. 212, 319 (1983).
- HE84 - O. Helene, Nucl. Instr. and Meth. 228, 120 (1984).
- HE86 - O. Helene, "Método dos mínimos quadrados com formalismo matricial", IFUSP (1986), não publicado.
- HI87 - K. Hirata et al., Phys. Rev. Lett., 58, 1490 (1987).
- HU85 - P. Hubert et al., Nuovo Cimento 85A, 19 (1985).
- JA80 - F. James e M. Roos, Nucl. Phys. B171, 475 (1980).
- JA68 - E.T. Jaynes, IEEE Trans. Nucl. Sci. SSC-4, 227 (1968).
- JE61 - H. Jeffreys, "Theory of Probability", Claredon, Oxford (1961).
- KE79 - M.G. Kendall e A. Stuart, "The Advanced Theory of Statistics", vol. 2, 4ª edição, Griffin, London (1979).
- KU85 - H. Kumahora, Nucl. Instr. and Meth. A228, 431 (1985).
- LE78 - C.M. Lederer e V.S. Shirley, "Table of Isotopes", 7ª edição, Wiley, New York (1978).
- LE82 - H.W. Lewis, Am. J. Phys. 50, 1164 (1982).
- LE85 - H.W. Lewis, Am. J. Phys. 53, 592 (1985).
- LI82 - J.A. Little, Health Phys. 43, 693 (1982).

- LO85 - A. Lorenz, IND C(NDS)-171/GE (1985).
- MA81 - W. Mannhart, "A Small Guide to Generating Covariances", PTB - FMRB-84 (1981).
- ME87 - P. Mennrath, Comunicação privada (1987).
- MO87 - A. Morales et al., An. Fis. Ses. B, 83, 28 (1987).
- MU80 - J.W. Muller, "Un Nouveau Regard sur les Probabilités a Priori", BIPM-8016 (1980).
- MU87a - E.R. Mucciolo e O. Helene, Nucl. Instr. and Meth. A256, 153 (1987).
- MU87b - E.R. Mucciolo, Relatório FAPESP, nº 3 (1987), não publicado.
- MU88 - E.R. Mucciolo e O. Helene, manual e testes de um programa para relocação estatística de espectros, a ser publicado.
- OL86 - R.A.M. de Oliveira, "Análise Estatística de Medidas de Correlação Angular", Dissertação de Mestrado, IFUSP (1986).
- PA71 - B.S. Pasternack e N.H. Harley, Nucl. Instr. and Meth. 91, 533 (1971).
- PR85a - H.B. Prosper, Nucl. Instr. and Meth. A238, 500 (1985).