Introdução à Análise de dados em FAE **(4 de Outubro de 2024)**

Lista 2

Professores: Dilson Damião, Eliza Melo e Maurício Thiel *Nome:* José Gonçalves Chaves Junior

EXERCÍCIO 1

Deduza as equações para o ajuste linear para o caso que as incertezas em y são diferentes para cada ponto.
Resolução
Como vimos, a função que melhor se ajusta aos dados é aquela a qual a soma dos quadrados dos desvios é mínima. Para uma reta, temos que:

$$y = ax + b \tag{0.1}$$

E portanto, podemos escrever uma função $\chi^2(a, b)$, para diferentes pontos (x_i, y_i) temos que:

$$\chi^2 = \sum_{k=1}^{N} \frac{[y_i - (ax_i + b)]^2}{\sigma_i^2} \tag{0.2}$$

onde σ_i é a variância dos dados. Abrindo o quadrado da diferença, podemos reescrever χ^2 como:

$$\chi^2 = \sum_{k=1}^{N} \frac{[y_i^2 - 2y_i(ax_i + b) + (ax_i^2 + 2ax_ib + b^2)]}{\sigma_i^2} \tag{0.3}$$

e, separando os termos,

$$\sum_{k=1}^{N} \left(\frac{y_i}{\sigma_i}\right)^2 - 2a\sum_{k=1}^{N} \frac{y_ix_i}{\sigma_i^2} - 2b\sum_{k=1}^{N} \frac{y_i}{\sigma_i^2} + a^2\sum_{k=1}^{N} \left(\frac{x_i}{\sigma_i}\right)^2 + 2ab\sum_{k=1}^{N} \frac{x_i}{\sigma_i^2} + \sum_{k=1}^{N} \left(\frac{b}{\sigma_i}\right)^2 \tag{0.4}$$

como nossas incógnitas são a e b, para minimizarmos a função, calcularemos as derivadas parciais em relação à a e b, portanto:

$$\begin{cases} \frac{\partial\chi^2}{\partial a} = -2\sum_{k=1}^{N} \frac{y_ix_i}{\sigma_i^2} + a\sum_{k=1}^{N} \left(\frac{x_i}{\sigma_i}\right)^2 + 2b\sum_{k=1}^{N} \frac{x_i}{\sigma_i^2} = 0 \\ \frac{\partial\chi^2}{\partial b} = -2\sum_{k=1}^{N} \frac{y_i}{\sigma_i^2} + 2a\sum_{k=1}^{N} \frac{x_i}{\sigma_i^2} + 2b\sum_{k=1}^{N} \frac{1}{\sigma_i^2} = 0 \end{cases} \tag{0.5}$$

Podemos então, reescrever os termos das derivadas como:

$$\begin{cases} a\overline{x^2} + b\overline{x} = \overline{xy} \\ a\overline{x} + b = \overline{y} \end{cases} \tag{0.6}$$

Isolando os termos, temos finalmente, para os coeficientes:

$$\boldsymbol{a = \frac{\overline{xy} - \overline{x}\overline{y}}{\overline{x^2} - \overline{x}^2}} \tag{0.7}$$

$$\boldsymbol{b = \overline{y} - a\overline{x}} \tag{0.8}$$

E, portanto, podemos escrever a equação em função apenas dos pontos.

EXERCÍCIO 2

A seção de choque de um processo definido como sinal pode ser obtida experimentalmente pela equação:

$$\rho = \frac{N_{total} - N_{background}}{\mathcal{L}}$$

onde N_{total} é o número de eventos observados; $N_{background}$ é o número de eventos de fundo esperados e $\mathcal{L}$ é a luminosidade integrada. Considerando para este caso que o número total de eventos observados foi de 2567, o número de eventos de fundo esperado é 1223.5, e a luminosidade integrada de 25fb^{-1} com uma incerteza sistemática de 10%. Calcule o valor da seção de choque propagando separadamente as incertezas estatísticas (lembrando que aqui se trata de distribuição de Poisson) e sistemáticas.

Resolução

Substituindo os valores da questão para o cálculo de ρ, temos:

$$\rho = \frac{2567 - 1223.5}{25}\frac{1}{fb^{-1}} = 53.74 \text{ fb} \tag{0.9}$$

Para o cálculo da incerteza, temos:

$$\sigma^2 = \left(\frac{\sqrt{N}}{L}\right) + \left(\frac{\sqrt{N_{bg}}}{L}\right) \tag{0.10}$$

substituindo os valores, temos para $\sigma = 2.463$.

Como temos uma incerteza sistemática de 10%, precisamos calculá-la e propagar os erros novamente.

$$\sigma_{sis} = 0.10\sigma = 0.10 \times 53.74 \text{ fb} = 5.37 \text{ fb} \tag{0.11}$$

E, propagando os erros, temos o valor final para a seção de choque:

$$\boldsymbol{\rho = 53.74 \pm 5.90} \text{ fb} \tag{0.12}$$

EXERCÍCIO 3

A seleção de eventos em uma análise em FAE por busca de acoplamentos anômalos previu um número de eventos de fundo de 0.07 eventos após todos os cortes. Ao olhar para os dados, também após todos os cortes, se observou zero eventos. Diferentes modelos para acoplaments anômalos previram diferentes números de eventos a serem observados, desde 0.09 até 35 eventos. Considerando que a contagem de eventos se dá através de uma pdf de Poisson, calcule até quantos eventos esperados podemos excluir com essa análise com 95% de C.L.

Resolução

Como o número de eventos segue uma distribuição de Poisson, e o número de eventos observados foi zero, podemos escrever que $P(\lambda)$ é:

$$P(\lambda) = e^{-\lambda} \tag{0.13}$$

onde λ é o número esperado de eventos. Para um limite de 95 % de confiança, temos que $P(\lambda) = 0.05$. Dessa forma,

$$P(\lambda) = e^{-\lambda} = 0.05 \tag{0.14}$$

utilizando ln em ambos os lados da equação:

$$ln(e^{-\lambda}) = ln(0.05) \tag{0.15}$$

$$\lambda = 2.9957 \tag{0.16}$$

Portanto, o número esperado de eventos é de aproximadamente 2996 eventos.

EXERCÍCIO 4

Mostre que a melhor função que se adequa aos dados é quando $\chi^2/\text{ndf} \to 1$.
Resolução
Como vimos anteriormente, a função que melhor se ajusta aos dados é aquela a qual a soma dos quadrados dos desvios é mínima, ou seja, quando $\chi^2 \to 1$. A demonstração pode ser feita com base no princípio de máxima verossimilhança, que nos leva à distribuição para dados com incertezas gaussianas.
Escrevemos χ^2 como:

$$\chi^2 = \sum_{i}^{N} \frac{(y_i - f(x_i)}{\sigma_i^2} \tag{0.17}$$

Para avaliar o ajuste, precisamos levar em consideração o número de graus de liberdade (ndf), além disso, que o valor esperado de χ^2 deve ser aproximadamente igual ao número de graus de liberdade [1].
Dessa forma, se definirmos o valor esperado para um modelo ajustado, temos que:

$$E\left[\frac{\chi^2}{ndf}\right] = 1 \tag{0.18}$$

Portanto, para uma função bem ajustada, esse valor deve tender à 1.

[1] James, F. (2006). *Statistical methods in experimental physics.*