Dosimetria

Estatística de Contagens

Dalila Mendonça

1 Introdução

Como a Radioatividade se trata de um processo aleatório, realizar medidas das partículas emitidas está relacionada a uma certa flutuação estatística que está associada a imprecisão nas contagens.

A estatística de contagens é utilizada para determinar uma medida adequada da radiação emitida em processos nucleares juntamente com a incerteza associada a medida. A estatística de contagens pode ser dividida em duas categorias:

1. É utilizada para verificar o correto funcionamento do equipamento de medida onde diversas medidas são realizadas sob as mesmas condições, mantendo as características do equipamento o mais constante possível. Devido às flutuações das leituras, as medidas obtidas terão uma certa variação interna que pode ser quantificada e comparada com valores determinados em modelos estatísticos preditivos e caso exista uma divergência pode-se inferir que o equipamento está com anormalidades nas contagens;

2. É utilizada para determinar a incerteza estatística inerente da medição quando apenas uma medida é realizada;

2 Caracterização dos Dados

Assumindo uma coleção de N medidas independentes da mesma quantidade física:

$$x_1, x_2, x_3, ..., x_i, ..., x_N$$

onde os valores de x_i são inteiros, medidos em intervalos de tempo com mesmo tamanho.

- A *Soma $\mathcal{S}$* das medidas é dada por:

$$\mathcal{S} \equiv \sum_{i=1}^{N} x_i \qquad \text{(Eq. 1)}$$

- A *Média Experimental $\bar{x}_e$* das medidas é dada por:

$$\bar{x}_e \equiv \frac{\mathcal{S}}{N} \qquad \text{(Eq. 2)}$$

onde N é o total de medidas realizadas.

- A *Função de Distribuição de Frequência $F(x)$* representa a frequência relativa com o qual cada medida aparece em uma coleção de dados, ou seja:

$$F(x) = \frac{Numero\ de\ ocorrencias\ de\ x}{numero\ total\ de\ medidas\ N} \qquad \text{(Eq. 3)}$$

A distribuição de frequências é um valor normalizado, ou seja:

$$\sum_{x=0}^{\infty} F(x) = 1 \qquad \text{(Eq. 4)}$$

- ☐ A média experimental $\bar{x}_e$ pode ser obtida em função da distribuição de frequência através da seguinte relação:

$$\bar{x}_e = \sum_{x=0}^{\infty} x \cdot F(x) \qquad \text{(Eq. 5)}$$

- ☐ O *resíduo d_i* de qualquer dado do conjunto, que é a diferença do valor medido para a média do conjunto de dados, ou seja

$$d_i \equiv x_i - \bar{x}_e \qquad \text{(Eq. 6)}$$

de modo que:

$$\sum_{i=1}^{\infty} d_i = 0 \qquad \text{(Eq. 7)}$$

- ☐ O desvio, é definido semelhantemente ao resíduo, com a diferença que para o desvio é utilizado a média experimental $\bar{x}_e$ e para o desvio é utilizado o valor verdadeiro da média $\bar{x}$, ou seja:

$$\epsilon_i \equiv x_i - \bar{x} \qquad \text{(Eq. 8)}$$

- ☐ A variância amostral quantifica a quantidade de flutuação interna em conjunto de dados. É definida como a média do quadrado do desvio:

$$s^2 \equiv \epsilon^2 = \frac{1}{N} \sum_{i=1}^{\infty} (x_i - \bar{x})^2 \qquad \text{(Eq. 9)}$$

A variância amostral é útil para avaliar o grau de dispersão de um conjuntos de dados ou avaliar o quando um valor é diferente de outro.

- ☐ Uma vez que não se sabe o valor verdadeiro da média, a expressão alternativa para obter a variância é dada por:

$$s^2 = \frac{1}{N-1} \sum_{i=1}^{\infty} (x_i - \bar{x}_e)^2 \qquad \text{(Eq. 10)}$$

onde é aplicada a uma amostra onde é utilizada a média experimental. Como a variância amostral é dada pela média do quadrado dos desvios do valor medido para a média, a variância amostral é uma medida da flutuação do conjunto de dados.

- ☐ A variância amostral em termos da função de distribuição de frequência é dada por:

$$s^2 = \sum_{x=0}^{\infty} (x - \bar{x})^2 F(x) \qquad \text{(Eq. 11)}$$

o que permite chegar na relação:

$$s^2 = \bar{x^2} - \bar{x}^2 \quad \text{(Eq. 12)}$$

3 Modelos Estatísticos

Sob certas condições, é possível predizer a função de distribuição que irá descrever o resultado de diversas medições feitas repetidamente. As medições são definidas como a contagem do número de sucessos em um ensaio, onde só existem duas possibilidades: sucesso ou falha (processo binário).

Processos Binários

- Jogar uma moeda: onde a Cara pode ser definida como a opção de sucesso e a probabilidade de sucesso é de $p = 1/2$;
- Jogar um dado: Onde uma das faces pode ser escolhida como a opção de sucesso, por exemplo a face 6 e a probabilidade de sucesso é de $p = 1/6$;
- Decaimento radioativo: onde o decaimento pode ser definido como a opção de sucesso e a probabilidade de sucesso é de $1 - e^{-\lambda t}$

Existem três principais modelos estatísticos que abordam processos binários:

1. ***Distribuição Binomial:*** Este modelo amplamente utilizado para todos os casos em que a probabilidade p de ocorrência de sucesso é constante. Não é computacionalmente aplicável à decaimentos radioativos pois o número de núcleos é sempre muito grande o que o torna raramente aplicável à contagens nucleares.
2. ***Distribuição de Poisson:*** Este modelo é uma aproximação matemática direta da distribuição binomial quando é considerado que a probabilidade de sucesso p é pequena e constante. Isto significa que é escolhida um tempo de observação muito menor que o tempo de meia-vida da fonte, de forma que o número de átomos decaindo em função do tempo é constante e a probabilidade de registrar uma contagem para um dado núcleo é sempre pequena;
3. ***Distribuição Gaussiana:*** É também conhecida como distribuição normal; e faz a aproximação para os casos em que o número médio de sucessos é relativamente grande. Esta condição se aplica a qualquer situação em que será acumulado diversas contagens durante a medição e não apenas algumas poucas contagens. Portanto, este é o modelo amplamente utilizado em estatística de contagens.

3.1 Distribuição Binomial

Se n é o número de ensaios e para cada ensaio existe a probabilidade p de ocorrer um sucesso então podemos predizer que a probabilidade de contarmos x sucessos é dada por:

$$P(x) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} \quad \text{(Eq. 13)}$$

onde $P(x)$ é a função preditiva da distribuição de probabilidade, definida apenas para valores inteiros de x e n.

- A Distribuição Binomial é normalizada de forma que:

$$\sum_{x=0}^{n} P(x) = 1 \qquad \text{(Eq. 14)}$$

- O valor médio $\bar{x}$ da distribuição é dado por:

$$\bar{x} = \sum_{x=0}^{n} x \cdot P(x) \qquad \text{(Eq. 15)}$$

que pode ser simplificado para a relação:

$$\bar{x} = p \cdot n \qquad \text{(Eq. 16)}$$

- A variância é dada por:

$$\sigma^2 \equiv \sum_{x=0}^{n} (x - \bar{x})^2 \cdot P(x) \qquad \text{(Eq. 17)}$$

que pode ser simplificada para a relação:

$$\sigma^2 = np(1 - p) \qquad \text{(Eq. 18)}$$

$$\sigma^2 = \bar{x}(1 - p) \qquad \text{(Eq. 19)}$$

- O desvio padrão, é dado pela raiz da variância, ou seja:

$$\sigma = \sqrt{\sigma^2} \qquad \text{(Eq. 20)}$$

O desvio padrão nos fornece a diferença entre o valor medido e a média verdadeira.

3.2 Distribuição de Poisson

Como citado anteriormente, esta distribuição se adequa àqueles processos binários onde a probabilidade de obter sucesso é pequena, ou seja $p \ll 1$, e não varia com o tempo. Portanto, pode ser aplicada a maioria das contagens nucleares onde o número de núcleos é grande e o tempo de observação é muito menor que o tempo de meia vida do elemento radioativo.

A aproximação da distribuição binomial para se adequar a estes casos é dada por:

$$P(x) = \frac{(\bar{x})^x e^{-\bar{x}}}{x!} \qquad \text{(Eq. 21)}$$

onde $\bar{x} = pn$. Esta simplificação é util quando se conhece apenas a média da distribuição e não se sabe o tamanho e as probabilidades individuais da amostra.

- A distribuição de Poisson é normalizada, de modo que:

$$\sum_{x=0}^{n} P(x) = 1 \qquad \text{(Eq. 22)}$$

- A média da distribuição de Poisson é dada por:

$$\bar{x} = \sum_{x=0}^{n} x \cdot P(x) \qquad \text{(Eq. 23)}$$

- A variância da distribuição de Poisson é dada por:

$$\sigma^2 \equiv \sum_{x=0}^{n} (x - \bar{x})^2 \cdot P(x) = pn \qquad \text{(Eq. 24)}$$

ou seja,

$$\sigma^2 = \bar{x} \qquad \text{(Eq. 25)}$$

- O desvio padrão é dado por:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\bar{x}} \qquad \text{(Eq. 26)}$$

Esta distribuição é aproximadamente centrada em torno da média embora seja consideravelmente assimétrica devido aos valores pequenos da média.

3.3 Distribuição Normal (Gaussiana)

A Distribuição de Poisson fornece uma aproximação matemática da distribuição binomial para os casos em que $p \ll 1$. A distribuição normal faz a aproximação da distribuição de Poisson para os casos em que o valor médio da distribuição é grande (maior que 25 ou 30). A Gaussiana é dada por:

$$P(x) = \frac{1}{\sqrt{2\pi x}} e^{-\frac{(x-\bar{x})^2}{2\bar{x}}} \qquad \text{(Eq. 27)}$$

As propriedades da Distribuição Normal são:

- É normalizada, ou seja:

$$\sum_{x=0}^{n} P(x) = 1 \qquad \text{(Eq. 28)}$$

- A média da distribuição Normal é dada por:

$$\bar{x} = \sum_{x=0}^{n} x \cdot P(x) \qquad \text{(Eq. 29)}$$

- A variância da distribuição Normal é dada por:

$$\sigma^2 \equiv \sum_{x=0}^{n} (x - \bar{x})^2 \cdot P(x) = pn \qquad \text{(Eq. 30)}$$

ou seja,

$$\sigma^2 = \bar{x} \qquad \text{(Eq. 31)}$$

□ O desvio padrão é dado por:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\bar{x}} \qquad \text{(Eq. 32)}$$

Os valores de desvio padrão podem ser utilizados como limites superiores e inferiores da margem de erro de forma que dependendo da quantidade de desvios padrão utilizados para estabelecer os limites há um percentual de chance de o valor verdadeiro para a média estar dentro do intervalo. A tabela 1 mostra os limites do intervalo para uma medida única com valor igual a 100, onde a confiança representa o percentual de chance do valor verdadeiro da média estar dentro daquele intervalo, e o desvio padrão é dado por $\sigma = \sqrt{100} = 10$ (Para uma contagem única x, $\sigma = \sqrt{x}$).

Tabela 1: *Exemplo de intervalos de para uma medida única $x = 100$*

Desvios Padrão	Intervalo	Confiança
$x \pm 0.67\sigma$	93.3 - 106.7	50%
$x \pm \sigma$	90.0 - 110.0	68%
$x \pm 1.64\sigma$	83.6 - 116.4	90%
$x \pm 1.96\sigma$	80.4 - 119.6	95%
$x \pm 2.58\sigma$	74.2 - 125.8	99%

O desvio padrão relativo é definido como

$$\sigma/x$$

4 Propagação de Erros

Dada uma medida u, obtida através das medidas x, y, z,... ou seja, u = u(x, y, z, ...), cada uma com seus respectivos desvios padrões σ_x, σ_y, σ_z, ...; O desvio padrão da medida u é dada por:

$$\sigma_u^2 = \left(\frac{\partial u}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial u}{\partial y}\right)^2 \sigma_y^2 + \left(\frac{\partial u}{\partial z}\right)^2 \sigma_z^2 + ... \qquad \text{(Eq. 33)}$$

4.1 Soma ou Subtração de Contagens

Dado

$$u = x + y \qquad ou \qquad u = x - y$$

O desvio padrão será:

$$\sigma_u = \sqrt{\sigma_x^2 + \sigma_y^2} \qquad \text{(Eq. 34)}$$

4.2 Multiplicação por Constante

Dado

$$u = A \cdot x$$

, onde A é uma constante, o desvio padrão de u será:

$$\sigma_u = A\sigma_x \quad \text{(Eq. 35)}$$

4.3 Divisão Por Constante

Dado

$$u = \frac{1}{B} \cdot x$$

, onde B é uma constante, o desvio padrão de u será:

$$\sigma_u = \frac{1}{B}\sigma_x \quad \text{(Eq. 36)}$$

4.4 Multiplicação ou Divisão De Contagens

Dado

$$u = x \cdot y$$

ou

$$u = \frac{x}{y}$$

onde x e y são contagens, o desvio padrão de u será:

$$\sigma_u = u\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2} \quad \text{(Eq. 37)}$$

4.5 Valor Médio de Múltiplas Contagens Independentes

Dado

$$\mathcal{S} = x_1 + x_2 + ... + x_N$$

onde x_i foram medidas tomadas da mesma variável no mesmo intervalo de tempo. Dado N medições, a média das medidas é dada por:

$$\bar{x} = \frac{\mathcal{S}}{N}$$

Portanto, o desvio padrão para a média é dado por:

$$\sigma_{\bar{x}} = \sqrt{\frac{\bar{x}}{N}} \quad \text{(Eq. 38)}$$

Referências

[1] Frank Herbert Attix. *Introduction to radiological physics and radiation dosimetry.* John Wiley & Sons, 2008.

[2] Glenn F Knoll. *Radiation detection and measurement.* John Wiley & Sons, 2010.