

# ANÁLISE ESTATÍSTICA DE MEDIDAS EM CIÊNCIAS EXATAS

Vito R. Vanin, Philippe Gouffon, Otaviano Helene

Março 2023



## Capítulo 3

# Significado e interpretação da variância

Este capítulo trata de dois assuntos muito comuns na análise de experimentos. O primeiro deles corresponde à estimativa do desvio padrão de uma grandeza que não é medida diretamente, mas sim uma função conhecida de outras grandezas, cujos desvios padrão podem ser estimados. Veremos que estimar corretamente a variância dessa função requer as *covariâncias* entre as grandezas das quais depende. Em seguida, descreveremos como *reduzir* medidas de uma grandeza que tem um valor verdadeiro, supondo os dados independentes estatisticamente e distribuídos conforme a normal, sem realizar qualquer aproximação na avaliação dos níveis de confiança.

### 3.1 Fórmula aproximada do cálculo da variância de uma função de variáveis aleatórias

Considere uma função  $z$  de variáveis aleatórias  $y_1, y_2, \dots, y_\mu$ ,

$$z = z(y_1, y_2, \dots, y_\nu, \dots, y_\mu) \quad . \quad (3.1)$$

Então,  $z$  também é uma variável aleatória. Pode-se calcular a variância de  $z$  quando as variâncias dos  $y_\nu$  são conhecidas,

$$\begin{aligned} \sigma_\nu^2 &= \langle (y_\nu - \langle y_\nu \rangle)^2 \rangle = \\ &= \int \dots \int (y_\nu - \langle y_\nu \rangle)^2 f(y_1, \dots, y_\nu, \dots, y_\mu) dy_1 \dots dy_\mu \quad . \end{aligned}$$

Muitas vezes dá-se o nome de *fórmula de propagação de erros* à fórmula que obteremos para estimar o desvio padrão de  $z$ . Neste livro, porém, estabeleceremos uma distinção entre *erro* e *desvio padrão* (seção 1.3). Para manter a consistência com o que foi discutido lá, não podemos associar o conceito de erro ao desvio padrão, portanto não daremos esse nome, mas sim *fórmula da variância de uma função de variáveis aleatórias*. Como associamos a idéia de incerteza à variância, podemos usar também o nome *fórmula de propagação de incertezas*.

Expandindo  $z(\vec{y})$  em torno do valor médio de  $\vec{y}$ ,  $\vec{y}_0$ , obtém-se

$$z(\vec{y}) = z(\vec{y}_0) + \sum_{\nu=1}^{\mu} \frac{\partial z}{\partial y_{\nu}} \Big|_{\vec{y}=\vec{y}_0} (y_{\nu} - y_{0,\nu}) + O[(y_{\nu} - y_{0,\nu})^2] \quad . \quad (3.2)$$

O valor médio de  $z$  será, então,

$$\langle z(\vec{y}) \rangle = \langle z(\vec{y}_0) \rangle + \sum_{\nu=1}^{\mu} \frac{\partial z}{\partial y_{\nu}} \Big|_{\vec{y}=\vec{y}_0} \langle (y_{\nu} - y_{0,\nu}) \rangle + O[\langle (y_{\nu} - y_{0,\nu})^2 \rangle] \quad ,$$

onde consideramos que os coeficientes da expansão em série são constantes (as derivadas parciais devem ser calculadas em  $\vec{y}_0$ ) para fatorá-los no cálculo do valor médio. Com as definições de média e variância e escrevendo  $\langle z \rangle$  no lugar de  $\langle z(\vec{y}) \rangle$  para uma notação mais compacta, obtem-se

$$\langle z \rangle = z(\vec{y}_0) + O[\sigma_{\nu}^2] \quad ,$$

uma vez que o termo linear da expansão tem média nula porque  $\langle y_{\nu} \rangle = y_{0,\nu}$ . Quando o termo nas variâncias dos  $y_{\nu}$  puder ser ignorado, então

$$\langle z \rangle \cong z(\vec{y}_0) \quad . \quad (3.3)$$

Se as derivadas de  $z$  de ordem superior ou igual à segunda são nulas, ou seja, se  $z$  é uma função linear de  $\vec{y}$ , a expressão (3.3) acima é exata e não aproximada. Caso  $\langle z(\vec{y}) \rangle$  não seja linear, mas tiver um comportamento bastante linear em uma região de alguns desvios-padrão  $\sigma_{\nu}$  em torno dos valores  $y_{0,\nu}$ , essa aproximação será boa.

Com a aproximação do valor médio da fórmula (3.3) e a expansão em série da relação (3.2), novamente na mesma aproximação de ignorar termos de ordem mais alta do que a primeira, obtemos

$$\sigma_z^2 \cong \langle [z(\vec{y}) - z(\vec{y}_0)]^2 \rangle \cong \left\langle \left( \sum_{\nu=1}^{\mu} \frac{\partial z}{\partial y_{\nu}} \Big|_{\vec{y}=\vec{y}_0} (y_{\nu} - y_{0,\nu}) \right)^2 \right\rangle ,$$

$$\sigma_z^2 \cong \sum_{\nu=1}^{\mu} \frac{\partial z}{\partial y_{\nu}} \sum_{\lambda=1}^{\mu} \frac{\partial z}{\partial y_{\lambda}} \langle (y_{\nu} - y_{0,\nu})(y_{\lambda} - y_{0,\lambda}) \rangle .$$

Na fórmula acima, subentendemos que as derivadas devem ser calculadas para  $\vec{y} = \vec{y}_0$ . Assim, a variância de uma função depende de todas as *covariâncias* entre os  $y_{\nu}$ , uma vez que o valor médio

$$\langle (y_{\nu} - y_{0,\nu})(y_{\lambda} - y_{0,\lambda}) \rangle = \text{cov}(y_{\nu}, y_{\lambda}) = \int \dots \int (y_{\nu} - \langle y_{\nu} \rangle)(y_{\lambda} - \langle y_{\lambda} \rangle) f(y_1, \dots, y_{\mu}) dy_1 \dots dy_{\mu} , \quad (3.4)$$

não é nulo em geral, como vimos na f.d.p. multinormal da seção 2.4. Quando  $f(y_1, \dots, y_{\mu})$  pode ser fatorada de maneira que a dependência em  $y_{\nu}$  ou em  $y_{\lambda}$  fica isolada em um único fator, dizemos que  $y_{\nu}$  é independente de  $y_{\lambda}$  e a covariância entre essas duas variáveis aleatórias é nula. Além disso, a variância de uma grandeza pode ser entendida como a covariância dela com ela mesma,

$$\sigma_{\nu}^2 = \text{cov}(y_{\nu}, y_{\nu}) = \langle (y_{\nu} - y_{0,\nu})(y_{\nu} - y_{0,\nu}) \rangle = \text{var}(y_{\nu}) .$$

A fórmula da variância reduz-se, então, a

$$\sigma_z^2 \cong \sum_{\nu=1}^{\mu} \frac{\partial z}{\partial y_{\nu}} \Big|_{\vec{y}=\vec{y}_0} \sum_{\lambda=1}^{\mu} \frac{\partial z}{\partial y_{\lambda}} \Big|_{\vec{y}=\vec{y}_0} \text{cov}(y_{\nu}, y_{\lambda}) . \quad (3.5)$$

A fórmula fica mais compacta quando representamos o conjunto das variâncias e covariâncias da mesma maneira que fizemos ao apresentar a multinormal (veja equação 2.15), usando uma matriz  $\mathbf{V}$ , cujos elementos são

$$V_{\nu\lambda} = \text{cov}(y_{\nu}, y_{\lambda}) .$$

Representamos as derivadas parciais por um vetor coluna, cujo elemento  $\nu$  é dado por

$$\left( \frac{\partial z}{\partial y} \right)_{\nu} = \frac{\partial z}{\partial y_{\nu}} \Big|_{\vec{y}=\vec{y}_0} ,$$

ou seja,

$$\left. \frac{\partial \vec{z}}{\partial \vec{y}} \right|_{\vec{y}_0} = \begin{pmatrix} \frac{\partial z}{\partial y_1} \\ \frac{\partial z}{\partial y_2} \\ \vdots \\ \frac{\partial z}{\partial y_\mu} \end{pmatrix},$$

onde todas as derivadas parciais são calculadas no ponto  $\vec{y}_0$ . Esse vetor é simplesmente o gradiente de  $z(\vec{y})$  nas coordenadas  $y_\nu$  e calculado no ponto  $\vec{y}_0$ . Na notação acima, a fórmula (3.5) pode ser reescrita de maneira compacta e muito adequada para o caso em que há muitas variáveis envolvidas,

$$\sigma_z^2 \cong \sum_{\nu=1}^{\mu} \frac{\partial z}{\partial y_\nu} \Big|_{\vec{y}_0} \sum_{\lambda=1}^{\mu} \frac{\partial z}{\partial y_\lambda} \Big|_{\vec{y}_0} V_{\nu\lambda} = \frac{\partial \vec{z}}{\partial \vec{y}} \Big|_{\vec{y}_0}^t \mathbf{V} \frac{\partial \vec{z}}{\partial \vec{y}} \Big|_{\vec{y}_0}. \quad (3.6)$$

A seção seguinte fornece um exemplo simples de uma situação em que as variáveis aleatórias de interesse não são independentes. As covariâncias não são nulas também entre os parâmetros de uma função, ajustados a um conjunto de dados experimentais. Poderíamos citar muitos outros exemplos e, de fato, não há como escapar das covariâncias, elas são mais comuns do que parece à primeira vista.

---

### Exemplo 3.1

---

Considere a função de variáveis aleatórias  $z = 3y_1 + 2y_2$  com  $y_{0,1} = 4, 0$ ,  $y_{0,2} = 5, 0$  e

$$\mathbf{V}_y = \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 0,5 \end{pmatrix}$$

é a matriz de variâncias das grandezas  $y_1$  e  $y_2$ . Nesse caso,

$$\left. \frac{\partial \vec{z}}{\partial \vec{y}} \right|_{\vec{y}_0} = \begin{pmatrix} \frac{\partial z}{\partial y_1} \\ \frac{\partial z}{\partial y_2} \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Portanto,  $\sigma_z^2$  (equação (3.6)) é

$$\sigma_z^2 = \begin{pmatrix} 3 & 2 \end{pmatrix} \begin{pmatrix} 1,0 & 0,5 \\ 0,5 & 0,5 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 17.$$

Com este resultado mais a relação (3.3), temos as estimativas  $\hat{z} = 22$  e  $\sigma_z = 4$ .

---

Considere um experimento de medida da energia de partículas carregadas por deflexão no ímã de um espectrômetro magnético. A deflexão é proporcional ao inverso da quantidade de movimento e o erro na energia calculada provirá de duas fontes: *i*) a medida da deflexão do feixe de partículas e *ii*) a calibração do espectrômetro. Na comparação de dados obtidos com o mesmo aparelho, os erros de calibração afetam feixes diferentes da mesma maneira ou, pelo menos, de maneiras não independentes. Assim, duas energias medidas pelo mesmo espectrômetro terão sempre uma covariância estatística (dizemos também que são *correlacionadas*). Já na comparação entre energias medidas em espectrômetros diferentes e calibrados independentemente, os erros de calibração originam variâncias que devem ser adicionadas às variâncias das medidas de deflexão do feixe.

O mesmo método usado no cálculo da variância de uma função pode ser empregado no cálculo da covariância entre duas funções. Considere que, além da função  $z_1(\vec{y})$ , definimos outra função de variáveis aleatórias,  $z_2(\vec{y})$ ,

$$\begin{aligned} z_1 &= z_1(y_1, y_2, \dots, y_\nu, \dots, y_\mu) \\ z_2 &= z_2(y_1, y_2, \dots, y_\nu, \dots, y_\mu) \quad . \end{aligned}$$

A covariância entre  $z_1$  e  $z_2$  é

$$\begin{aligned} \text{cov}(z_1, z_2) &\cong \langle [z_1(\vec{y}) - z_1(\vec{y}_0)] \cdot [z_2(\vec{y}) - z_2(\vec{y}_0)] \rangle \\ &\cong \sum_{\nu=1}^{\mu} \frac{\partial z_1}{\partial y_\nu} \sum_{\lambda=1}^{\mu} \frac{\partial z_2}{\partial y_\lambda} \text{cov}(y_\nu, y_\lambda) \quad . \end{aligned}$$

Em notação matricial, fica

$$\text{cov}(z_1, z_2) \cong \frac{\overrightarrow{\partial z_1}^t}{\partial y} \mathbf{V}_y \frac{\overrightarrow{\partial z_2}}{\partial y} \quad . \quad (3.7)$$

**Q3.1** *Obtenha a expressão (3.7) acima. Mostre que a expressão (3.6) pode ser entendida como uma caso particular da expressão (3.7).*

As expressões (3.6) e (3.7) são exatas quando  $z_1$  e  $z_2$  são funções lineares em  $\vec{y}$ , porque todos os termos de ordem igual ou superior a 2 na expansão (3.2) são nulos. Além disso, a derivada primeira é constante e pode ser calculada

mesmo sem o conhecimento de  $\vec{y}_0$ , o que possibilita aplicar estas fórmulas na prática.

Quando as funções envolvidas no cálculo das fórmulas (3.6) e (3.7) não forem lineares nas variáveis aleatórias  $\vec{y}$ , será impossível calcular exatamente o vetor das derivadas parciais no ponto  $\vec{y}_0$ , simplesmente porque este ponto é o valor verdadeiro de  $\vec{y}$ , geralmente desconhecido e impossível de determinar. Tudo o que poderemos obter será uma aproximação, usando as fórmulas (3.6) e (3.7) com o vetor das derivadas calculado no ponto  $\hat{\vec{y}}$ , o acento circunflexo significando *valor estimado*, ou seja,  $\hat{\vec{y}}$  é a *estimativa* de  $\vec{y}_0$ ,

$$\left( \frac{\partial z}{\partial y} \right)_{\nu} = \frac{\partial z}{\partial y_{\nu}} \Big|_{\vec{y}=\hat{\vec{y}}} . \quad (3.8)$$

Daí o título da seção ser *fórmula aproximada*...

A variância de  $z$  pode ser calculada a partir da f.d.p. de  $z$ ,  $g(z)$ , como

$$\sigma_z^2 = \langle (z - \langle z \rangle)^2 \rangle = \int (z - \langle z \rangle)^2 g(z) dz ,$$

e nesta seção calculamos a variância sem, aparentemente, conhecer  $g(z)$ . Esta f.d.p., porém, pode ser calculada a partir das informações disponíveis, como discutiremos a seguir.

Se conhecemos a f.d.p. das variáveis aleatórias  $y_{\nu}$ ,  $f(y_1, \dots, y_{\nu}, \dots, y_{\mu})$ , e a expressão de  $z$  em função dos  $y_i$ , podemos calcular a f.d.p. da variável  $z$  para qualquer  $z'$  a partir de  $f(y_1, \dots, y_{\nu}, \dots, y_{\mu})$ , integrando  $f$  em todas as variáveis  $y_1, \dots, y_{\mu}$  de maneira que o vínculo explicitado por  $z(\vec{y}) = z'$  seja respeitado,

$$g(z') = \int \dots \int \delta(z(\vec{y}) - z') f(y_1, \dots, y_{\nu}, \dots, y_{\mu}) dy_1 \dots dy_{\mu} ,$$

onde o vínculo foi imposto pela função  $\delta$ . Substituindo na expressão anterior,

$$\sigma_z^2 = \int (z' - \langle z \rangle)^2 \int \dots \int \delta(z(\vec{y}) - z') f(\vec{y}) dy_1 \dots dy_{\mu} dz' ,$$

e integrando primeiro em  $dz'$ , obtem-se

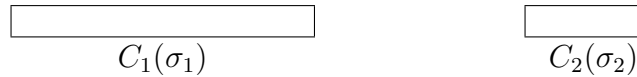
$$\sigma_z^2 = \int (z(\vec{y}) - \langle z \rangle)^2 f(y_1, \dots, y_{\mu}) dy_1 \dots dy_{\mu} .$$

A expressão acima corresponde exatamente à calculada nesta seção, explicando como calculamos a variância de uma variável  $z$  sem aparentemente conhecer sua f.d.p..



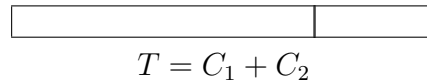
## 3.2 Covariância – um exemplo

Este exemplo vem de um trabalho clássico de W. Mannhart [Mannhart] sobre a origem das covariâncias em medidas físicas. Imagine duas barras de comprimentos  $C_1$  e  $C_2$  conhecidos com desvios padrão  $\sigma_1$  e  $\sigma_2$ , respectivamente, e covariância nula entre eles.



O comprimento total das duas barras,  $T$ , será

$$T = C_1 + C_2 \quad \text{com} \quad \sigma^2 = \sigma_1^2 + \sigma_2^2$$



Deve ser possível recalcular  $C_2$  a partir da diferença entre  $T$  e  $C_1$ . Assim, calculamos  $C'_2$  como

$$C'_2 = T - C_1$$



Se fizermos o cálculo da variância de  $C'_2$  omitindo a covariância entre  $T$  e  $C_1$ , obtemos

$$\sigma_2'^2 = \sigma^2 + \sigma_1^2 = 2\sigma_1^2 + \sigma_2^2 \quad .$$

Este resultado está incorreto, pois  $C'_2$  é idêntico a  $C_2$ , de maneira que suas variâncias têm que ser iguais. O cálculo deu errado porque omitimos o termo da covariância entre  $T$  e  $C_1$ . Na medida de  $T$  efetuada como descrito acima (em duas parcelas, medindo  $C_1$  e  $C_2$ ), parte do erro provém da medida de  $C_1$ . É a existência de um erro comum em duas medidas (ou de erros dependentes) que determina a existência de uma covariância entre elas. Em particular, se o valor  $C_1$  observado for menor que o comprimento verdadeiro, então  $T$  tende a ser menor que seu valor verdadeiro — a covariância entre  $T$  e  $C_1$  será positiva porque ambas as variações tenderão a ocorrer no mesmo sentido. Podemos calcular a covariância entre  $T$  e  $C_1$  por meio da fórmula (3.7), sendo que para aplicá-la é preciso calcular as derivadas

$$\frac{\partial T}{\partial C_1} = 1, \quad \frac{\partial T}{\partial C_2} = 1, \quad \frac{\partial C_1}{\partial C_1} = 1 \quad \text{e} \quad \frac{\partial C_1}{\partial C_2} = 0 \quad ,$$

com as quais montamos os vetores

$$\frac{\partial \vec{T}}{\partial C} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{e} \quad \frac{\partial \vec{C}_1}{\partial C} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} .$$

Precisamos também montar a matriz das variâncias das variáveis aleatórias  $C_1$  e  $C_2$ ,

$$\mathbf{V} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} .$$

A covariância entre  $T$  e  $C_1$ , calculada pela expressão (3.7), fica

$$\text{cov}(T, C_1) = (1 \ 1) \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \sigma_1^2 ,$$

simplesmente a parcela comum entre as variâncias de  $T$ ,  $\sigma_1^2 + \sigma_2^2$ , e de  $C_1$ ,  $\sigma_1^2$ .

Ao incluir essa covariância no cálculo da variância de  $C'_2$ , obtemos

$$C'_2 = T - C_1 \quad \text{com}$$

$$\sigma_2'^2 = \sigma^2 + \sigma_1^2 - 2 \text{cov}(T, C_1) = 2\sigma_1^2 + \sigma_2^2 - 2\sigma_1^2 = \sigma_2^2 ,$$

onde o sinal  $(-)$  que multiplica a covariância vem da derivada de  $C'_2$  em relação a  $C_1$  e calcula-se a variância de  $C'_2$  com a fórmula (3.5) ou com a expressão matricial (3.6), da variância de uma função de variáveis aleatórias.

---

### Exemplo 3.2

---

O exercício 3.5 proposto no final deste capítulo é um exemplo de introdução de covariâncias entre resultados por um processo de calibração, em que um espectrômetro de partículas carregadas permite observar apenas uma fração das partículas emitidas pela fonte. A eficiência de detecção,  $\epsilon$ , é definida simplesmente como a razão entre o número de partículas detetadas,  $n_d$ , e o número de partículas emitidas,  $n_e$ , ou seja,

$$\epsilon = \frac{\text{número de partículas detetadas}}{\text{número de partículas emitidas}} = \frac{n_d}{n_e} .$$

A eficiência é, frequentemente, uma função complicada da energia da partícula,  $E$ , e passaremos a representar a eficiência pela função  $\epsilon(E)$ . No exercício 3.5, supõe-se que

$$\epsilon = \frac{\alpha}{E} ,$$

uma forma simples que evitará complicar os cálculos e permitirá realçar o efeito de introdução da covariância que pretendemos mostrar.

O parâmetro  $\alpha$  da curva de eficiência dificilmente pode ser obtido com precisão a partir do conhecimento do espectrômetro, sendo normalmente medido em um experimento. Nesse experimento de *calibração*, usam-se fontes que emitem um número conhecido de partículas por segundo (*fontes calibradas*), a fim de determinar um conjunto de dados  $(E_i; n_{e,i}; n_{d,i})$ , onde o índice  $i$  identifica a fonte. Além dos valores observados das grandezas, determinam-se os seus desvios-padrão. A esse conjunto de dados, ajusta-se o parâmetro  $\alpha$  da função de calibração,

$$n_{d,i} = \epsilon(E_i) \cdot n_{e,i} = \alpha \frac{n_{e,i}}{E_i} \quad .$$

No caso do exercício 3.5, obteve-se  $\alpha = 1,00(4) \cdot 10^{-2}$  MeV.

A tabela do exercício 3.5 mostra que, para uma fonte desconhecida, observaram-se 1025(40) partículas de 4,000 MeV e 800(30) partículas de 10,000 MeV. Com os valores dados pela eficiência calibrada,

$$\epsilon(4,000) = 2,50(10) \cdot 10^{-3}$$

$$\epsilon(10,000) = 1,00(4) \cdot 10^{-3} \quad ,$$

deduz-se que os números de partículas emitidas são

$$n_{e,4} = \frac{n_{d,4}}{\epsilon(4)} = \frac{1025(40)}{2,50(10) \cdot 10^{-3}} = 410(23) \cdot 10^3 \text{ partículas} \quad ,$$

$$n_{e,10} = \frac{n_{d,10}}{\epsilon(10)} = \frac{800(30)}{1,00(4) \cdot 10^{-3}} = 800(44) \cdot 10^3 \text{ partículas} \quad .$$

Estes dois resultados, porém, não são estatisticamente independentes. Caso o parâmetro  $\alpha$  obtido tenha sido, por exemplo, subestimado e o valor verdadeiro seja, digamos,  $1,05 \cdot 10^{-2}$  MeV, *ambos* os resultados estarão superestimados, sendo os valores calculados de acordo com esta última hipótese  $n_{e,4} = 390 \cdot 10^3$  e  $n_{e,10} = 762 \cdot 10^3$  partículas. Isto mostra que os resultados  $n_{e,4}$  e  $n_{e,10}$  não são independentes e que a covariância entre eles é *positiva*, desde que ambos alteram-se no mesmo sentido.

Podemos calcular a covariância entre  $n_{e,4}$  e  $n_{e,10}$  com a fórmula (3.7). As variáveis aleatórias envolvidas na determinação de  $n_{e,4}$  são  $n_{d,4}$  e  $\alpha$ , enquanto na determinação de  $n_{e,10}$ , são  $n_{d,10}$  e  $\alpha$ . Assim, o vetor de variáveis aleatórias a considerar é  $\vec{y} = (n_{d,4}; n_{d,10}; \alpha)$ , sendo todas *independentes*.

A matriz de covariância das variáveis aleatórias é, portanto, diagonal e vale

$$\mathbf{V}_y = \begin{pmatrix} \sigma_{n_{d,4}}^2 & 0 & 0 \\ 0 & \sigma_{n_{d,10}}^2 & 0 \\ 0 & 0 & \sigma_\alpha^2 \end{pmatrix} .$$

Os vetores das derivadas parciais são

$$\frac{\overrightarrow{\partial n_{e,4}}}{\partial y} = \begin{pmatrix} 1/\epsilon(4) \\ 0 \\ -n_{e,4}/\alpha \end{pmatrix} \quad \mathbf{e} \quad \frac{\overrightarrow{\partial n_{e,10}}}{\partial y} = \begin{pmatrix} 0 \\ 1/\epsilon(10) \\ -n_{e,10}/\alpha \end{pmatrix} .$$

Efetuada o cálculo da fórmula (3.7), obtem-se

$$\text{cov}(n_{e,4}, n_{e,10}) = n_{e,4} \cdot n_{e,10} \cdot \frac{\sigma_\alpha^2}{\alpha^2} = 5,2 \cdot 10^8 .$$

A melhor maneira de avaliar a relevância de uma covariância é usar o *coeficiente de correlação*,  $\rho$ , que é uma grandeza adimensional definida como

$$\rho = \frac{\text{cov}(n_{e,4}, n_{e,10})}{\sigma_{n_{e,4}} \sigma_{n_{e,10}}} = 0,5 .$$

Veremos adiante que o coeficiente de correlação é um número de módulo menor ou igual a 1, sendo que o valor 0,5 corresponde a uma correlação (ou dependência estatística) moderada.

### 3.3 Cálculo de variâncias e covariâncias: formalismo matricial

No final da seção 3.1, discutimos as estimativas das variâncias e covariâncias de duas funções,  $z_1$  e  $z_2$ , de variáveis aleatórias  $y_1, y_2, \dots, y_\mu$ , a partir do conhecimento das variâncias e covariâncias destas variáveis,  $\mathbf{V}$ , equações (3.6) e (3.7).

Podemos generalizar aqueles resultados no caso de  $\rho$  funções  $z_1, z_2, \dots, z_\rho$  de  $\mu$  variáveis aleatórias  $\vec{y}$ . Nesse caso, a matriz de covariância de  $\vec{z}$ ,  $\mathbf{V}_z$ , é dada por

$$\mathbf{V}_z \cong \mathbf{C} \cdot \mathbf{V} \cdot \mathbf{C}^t , \quad (3.9)$$

onde  $\mathbf{C}$  é a matriz das derivadas parciais de  $\vec{z}$  em relação a  $\vec{y}$ ,

$$C_{\nu\lambda} = \left. \frac{\partial z_\nu}{\partial y_\lambda} \right|_{\vec{y}=\vec{y}_0} , \quad (3.10)$$

### 3.3. CÁLCULO DE VARIÂNCIAS E COVARIÂNCIAS: FORMALISMO MATRICIAL83

$$\nu = 1, 2, \dots, \rho \quad \text{e} \quad \lambda = 1, 2, \dots, \mu.$$

Como discutido anteriormente, se  $\vec{z}$  for uma função linear de  $\vec{y}$ , então a equação (3.9) torna-se uma igualdade; caso contrário, as derivadas parciais da equação (3.10) devem ser calculadas nos pontos experimentais de  $\vec{y}$ ,  $\hat{\vec{y}}$ , como discutido no contexto da equação (3.8).

---

#### Exemplo 3.3

---

Para se determinar as massas de duas amostras líquidas, elas foram medidas dentro de um recipiente, obtendo-se os valores (em g)  $M_1 = 25,1 \pm 3,0$  e  $M_2 = 34,7 \pm 4,0$ . Posteriormente, a massa do recipiente foi medida e obteve-se o valor  $M_3 = 12,5 \pm 2,0$ . Vamos supor que as covariâncias entre esses resultados sejam nulas, de modo que a matriz de covariância de  $M_1$ ,  $M_2$  e  $M_3$  é

$$\mathbf{V}_M = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 4 \end{pmatrix} .$$

Os valores a serem adotados para as massas das duas amostras são, respectivamente,

$$\begin{aligned} m_1 &= M_1 - M_3 = 12,6 \text{ g} \quad , \\ m_2 &= M_2 - M_3 = 22,2 \text{ g} \quad . \end{aligned}$$

A matriz de covariância entre  $m_1$  e  $m_2$  é dada pela equação (3.9) com

$$C_{\nu\lambda} = \frac{\partial m_\nu}{\partial M_\lambda}$$

ou seja,

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} .$$

Aplicando a equação (3.9), com  $\mathbf{C}$  e  $\mathbf{V}_M$  acima, tem-se

$$\mathbf{V}_m = \mathbf{C}\mathbf{V}_M\mathbf{C}^t = \begin{pmatrix} 13 & 4 \\ 4 & 20 \end{pmatrix} .$$

Note que a covariância entre  $m_1$  e  $m_2$  não é nula pela dependência em comum de ambas em relação a  $M_3$ .

---

### 3.4 Qui-quadrado quando as variáveis são correlacionadas

No capítulo 2, vimos que, a partir de um conjunto de dados  $\{y_i, i = 1..N\}$ , pode-se definir a variável aleatória  $\chi^2$  pela fórmula (2.35)

$$\chi_N^2 = \sum_{i=1}^N \frac{(y_i - y_{0,i})^2}{\sigma_{0,i}^2} \quad ,$$

quando a variância de  $y_i$  é  $\sigma_{0,i}^2$  e seu valor verdadeiro é  $y_{0,i}$ . No entanto, na seção precedente, vimos que, quando os dados  $y_i$  são correlacionados, necessitamos de toda sua matriz de covariância  $\mathbf{V}$  com

$$V_{ij} = \text{cov}(y_i, y_j) \quad ,$$

de modo que a definição de  $\chi^2$  precisa ser modificada para lidar com esses casos. Para isso, definimos o vetor de dados  $\vec{y}$  como a matriz linha

$$\vec{y} = (y_1 \ y_2 \ \cdots \ y_i \ \cdots \ y_N) \quad (3.11)$$

e o vetor de valores verdadeiros  $\vec{y}_0$  como outra matriz linha

$$\vec{y}_0 = (y_{0,1} \ y_{0,2} \ \cdots \ y_{0,i} \ \cdots \ y_{0,N}) \quad . \quad (3.12)$$

Com essa notação, define-se  $\chi_N^2$  pela expressão

$$\chi_N^2 = (\vec{y} - \vec{y}_0) \mathbf{V}^{-1} (\vec{y} - \vec{y}_0)^t \quad , \quad (3.13)$$

que tem a f.d.p. da fórmula (2.45), desde que a f.d.p dos dados  $y_i$  seja a *multinormal*<sup>1</sup>. Além desta exigência da normalidade da função de probabilidade para que essa variável seja qui-quadrado, vale ressaltar que a distribuição de  $\chi_N^2$  depende do número  $N$  de elementos do vetor de dados  $\vec{y}$ , da mesma forma que a variável estudada na seção 2.8.

**Q3.2** Mostre que a definição de  $\chi_N^2$  da fórmula (3.13) acima reduz-se à expressão (2.35) quando os dados não são correlacionados, ou seja, quando a matriz de covariâncias  $\mathbf{V}$  é diagonal.

<sup>1</sup>Estritamente falando, a variável calculada pela expressão 3.13 recebe o nome de qui-quadrado somente quando a f.d.p dos dados  $y_i$  for a multinormal.

### 3.5 A f.d.p. de $t$ de Student

Em geral, na análise de um conjunto de dados  $\{x_i, i = 1, \dots, N\}$ , calcula-se a média

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

a estimativa não-tendenciosa da variância

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

e o desvio padrão da média

$$\sigma_m = \frac{\sigma}{\sqrt{N}} \quad .$$

Na seção 1.10, usamos essas grandezas para determinar intervalos de confiança para a grandeza medida, ou seja, intervalos que contenham o valor verdadeiro com uma probabilidade conhecida, dados pelas expressões (1.42) e (1.43), que, porém, são relações aproximadas. O que acontece quando analisamos um experimento é que, normalmente, estima-se o desvio-padrão  $\sigma_m$  a partir dos dados experimentais obtidos, de modo que é necessário considerar tanto a flutuação da média  $\bar{x}$  (já levada em conta em 1.42 e 1.43) quanto a do seu desvio-padrão,  $\sigma_m$ , a fim de determinar corretamente a probabilidade associada ao intervalo de confiança.

Começamos, portanto, com escrever as f.d.p.s da média e do seu desvio-padrão. Quando a f.d.p. dos dados é a normal e os dados são independentes, deduz-se que a média tem a f.d.p. (seção 2.6)

$$f(\bar{x}) = \left( \frac{N}{2\pi\sigma_0^2} \right)^{1/2} \exp \left\{ -\frac{N(\bar{x} - x_0)^2}{2\sigma_0^2} \right\} \quad (3.14)$$

e que a variância estimada se distribui como a variável  $\chi^2$  com  $N - 1$  graus de liberdade, sendo necessária alguma transformação de variável (seção 2.9). Nesta seção, usaremos a seguir a f.d.p. da variável  $s$ , cujo quadrado é

$$s^2 = \frac{1}{N} \sum_i^N (x_i - \bar{x})^2 = \frac{N-1}{N} \sigma^2 \quad . \quad (3.15)$$

A f.d.p. de  $s$  fica

$$g(s) = \frac{1}{2\Gamma\left(\frac{N-1}{2}\right)} \left(\frac{Ns^2}{2\sigma_0^2}\right)^{\frac{N-3}{2}} \exp\left\{-\frac{Ns^2}{2\sigma_0^2}\right\} \frac{2Ns}{\sigma_0^2} . \quad (3.16)$$

As estatísticas  $\bar{x}$  e  $s$  são independentes, isto é, a f.d.p. conjunta de  $\bar{x}$  e  $s$  é o produto das f.d.p.s de cada uma delas. Pode-se demonstrar esse fato rigorosamente (veja, por exemplo, [Kendall, capítulo 11]), mas aqui nos contentaremos em fornecer alguns argumentos de plausibilidade. Na seção 2.7, exemplo 2.4 e questão Q2.7, demonstramos esta independência quando os parâmetros eram determinados para medidas com 2 dados. Note também que a média dá a localização da f.d.p. e o desvio padrão mede a dispersão dos dados, características da f.d.p. de naturezas distintas, de maneira que não é impensável que as estatísticas associadas à média e ao desvio padrão sejam independentes.

A partir das f.d.p.s daquelas estatísticas, podemos calcular a probabilidade  $\alpha$  de certos intervalos conterem o valor verdadeiro  $x_0$ ,

$$P(\bar{x} - t_\alpha \sigma_m \leq x_0 \leq \bar{x} + t_\alpha \sigma_m) = \alpha . \quad (3.17)$$

Quando o número de dados é grande ( $N > 10$  é suficiente em muitos casos), o desvio padrão pode ser estimado com boa precisão e pode-se estimar as probabilidades  $\alpha$  da relação (3.17) usando simplesmente a f.d.p. da média, que é normal. A tabela 3.1 abaixo apresenta a relação entre  $\alpha$  e  $t_\alpha$ , nesse caso.

Tabela 3.1: Probabilidade do valor verdadeiro estar contido em um intervalo de largura  $t_\alpha$  desvios padrão da média em torno do valor médio, quando o número dos dados na medida é grande.

$t_\alpha$	$\alpha$
1	68%
2	95%
3	~99%

Se o número de dados é pequeno ou se necessitamos de precisão na determinação do intervalo com probabilidade próxima a 99%, devemos efetuar o cálculo exato. Reescrevendo o intervalo da relação (3.17) de maneira a isolar  $t_\alpha$ ,

$$-t_\alpha \leq \frac{x_0 - \bar{x}}{\sigma_m} \leq t_\alpha ,$$

vemos que a variável aleatória da qual precisamos conhecer a f.d.p. é a variável  $t$  definida como

$$t = \frac{\bar{x} - x_0}{\sigma_m} , \quad (3.18)$$



referida na literatura como  $t$  de Student. A expressão (3.17) consiste, então, no cálculo da probabilidade  $\alpha$  usando a f.d.p. de  $t$ ,  $\psi(t)$ ,

$$P(|t| \leq t_\alpha) = \alpha = \int_{-t_\alpha}^{t_\alpha} \psi(t) dt \quad . \quad (3.19)$$

Tanto o numerador quanto o denominador da função que define  $t$  na expressão (3.18) são variáveis aleatórias, o numerador com f.d.p. gaussiana e o denominador com uma f.d.p. parecida com aquela descrita em (3.16). O cálculo da f.d.p. de  $t$  é bastante complexo e, para não cairmos em expressões algébricas difíceis de lidar, precisaremos calcular intermediariamente a f.d.p. da variável  $T$  definida por

$$T = \frac{\bar{x} - x_0}{s} \quad , \quad (3.20)$$

com  $s = \sqrt{\frac{N-1}{N}}\sigma$ , definido na fórmula (3.15).

A transformação de variáveis a efetuar corresponde a

$$\left. \begin{array}{l} \bar{x} \\ s \end{array} \right\} \rightarrow \left\{ \begin{array}{l} T = (\bar{x} - x_0)/s \\ u = s \end{array} \right. \quad . \quad (3.21)$$

Calcularemos a f.d.p.  $h(T, u)$  a partir da f.d.p. conjunta da média e da variância, depois integremos para todo  $u$  de maneira a achar a f.d.p. de  $T$ .

**Q3.3** Mostre que o Jacobiano da transformação  $(\bar{x}, s) \rightarrow (T, u)$  vale

$$J = \frac{\partial(T, u)}{\partial(\bar{x}, s)} = \frac{1}{s} = \frac{1}{u} \quad , \quad (3.22)$$

o que significa que  $d\bar{x} ds = s dT du = u dT du$ .

**Q3.4** Deduza a fórmula (3.16), usando os resultados das seções 2.8 e 2.9.

Com as fórmulas (3.14) e (3.16), o Jacobiano da transformação (3.22) e a regra de transformação da variáveis da seção 2.7, podemos calcular a f.d.p. conjunta de  $T$  e  $u$ ,

$$h(T, u) = \left( \frac{N}{2\pi\sigma_0^2} \right)^{1/2} \frac{1}{2\Gamma\left(\frac{N-1}{2}\right)} \cdot \exp\left\{-\frac{N(\bar{x} - x_0)^2}{2\sigma_0^2}\right\} \left(\frac{Ns^2}{2\sigma_0^2}\right)^{\frac{N-3}{2}} \exp\left\{-\frac{Ns^2}{2\sigma_0^2}\right\} \frac{2Ns^2}{\sigma_0^2} \quad ,$$

em que, trocando  $\bar{x}$  e  $s$  por  $T$  e  $u$  conforme a relação (3.21), obtemos

$$h(T, u) = \left( \frac{N}{2\pi\sigma_0^2} \right)^{1/2} \frac{4}{2\Gamma\left(\frac{N-1}{2}\right)} \cdot \exp\left\{-\frac{Nu^2}{2\sigma_0^2}(T^2 + 1)\right\} \left(\frac{Nu^2}{2\sigma_0^2}\right)^{\frac{N-1}{2}} .$$

A f.d.p. de  $T$ ,  $\phi(T)$ , é simplesmente a integral de  $h$  para todo  $u$  (lembre,  $u = s$ ), ou seja

$$\phi(T) = \int_0^\infty h(T, u) du = \left( \frac{N}{2\pi\sigma_0^2} \right)^{1/2} \frac{4}{2\Gamma\left(\frac{N-1}{2}\right)} \cdot \int_0^\infty \exp\left\{-\frac{Nu^2}{2\sigma_0^2}(T^2 + 1)\right\} \left(\frac{Nu^2}{2\sigma_0^2}\right)^{\frac{N-1}{2}} du .$$

Para calcular a integral, mudamos de variável

$$v^2 = \frac{Nu^2}{2\sigma_0^2}(T^2 + 1)$$

quando a integral transforma-se em

$$\int_0^\infty e^{-v^2}(T^2 + 1)^{-(N-1)/2} v^{N-1} dv \left( \frac{2\sigma_0^2}{N(T^2 + 1)} \right)^{1/2} .$$

Uma consulta a [Arfken, cap. 10.1] permite verificar que

$$\Gamma(z) = 2 \int_0^\infty \exp(-v^2) v^{2z-1} dv \quad \text{quando} \quad \mathcal{R}(z) > 0 .$$

A função  $\Gamma(z)$  é a função fatorial, cuja definição exata para números inteiros e semi-inteiros foi dada na fórmula (2.43). Podemos então reunir o resultado

$$\phi(T) = \left(\frac{1}{\pi}\right)^{1/2} \frac{1}{\Gamma\left(\frac{N-1}{2}\right)} \cdot \left(\frac{1}{(T^2 + 1)}\right)^{\frac{N}{2}} \Gamma(N/2) .$$

Agora, a mudança de variável é mais simples, basta verificar que

$$t = \sqrt{N-1} T ,$$

o que fornece

$$\psi(t) = \left( \frac{1}{(N-1)\pi} \right)^{1/2} \frac{\Gamma(N/2)}{\Gamma\left(\frac{N-1}{2}\right)} \cdot \left( \frac{1}{\frac{t^2}{N-1} + 1} \right)^{\frac{N}{2}} . \quad (3.23)$$

Chamando o número de graus de liberdade de  $\nu$ ,

$$\nu = N - 1 \quad , \quad (3.24)$$

correspondente aos  $N$  dados menos o vínculo representado pela média, podemos reescrever  $\psi(t)$  como

$$\psi_\nu(t) = \left( \frac{1}{\nu\pi} \right)^{1/2} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \left( \frac{1}{\frac{t^2}{\nu} + 1} \right)^{\frac{\nu+1}{2}} . \quad (3.25)$$

Identifica-se a f.d.p. de  $t$  pelo subscrito  $\nu$  para recordar a dependência com o número de graus de liberdade.  $\psi_\nu(t)$  é a f.d.p. da variável aleatória  $t$  de Student para  $\nu$  graus de liberdade.

A f.d.p. de  $t$  tem valor máximo em  $t = 0$  e decresce na medida em que  $|t|$  cresce. Sua média é nula para  $\nu > 1$  grau de liberdade<sup>2</sup> e a variância é

$$\text{var}(t) = \sigma_t^2 = \frac{\nu}{\nu - 2}, \quad \text{para } \nu > 2 \quad . \quad (3.26)$$

Pode-se mostrar que

$$\psi_\nu(t) \xrightarrow{\nu \rightarrow \infty} N(t; 0, 1) \quad .$$

## 3.6 Intervalos de confiança

Na hipótese dos dados  $x_i$  de uma medição terem uma f.d.p. normal, é possível calcular exatamente o nível de confiança associado ao intervalo  $[\bar{x} - \sigma_m, \bar{x} + \sigma_m]$  com a f.d.p.  $\psi(t)$  da fórmula (3.25), em que a variável aleatória  $t$  de Student está definida na fórmula (3.18). Calcula-se

$$P(\bar{x} - \sigma_m \leq x_0 \leq \bar{x} + \sigma_m) = \int_{-1}^1 \psi_\nu(t) dt = \alpha \quad .$$

Essa probabilidade é uma função do número  $N = \nu + 1$  de dados. A tabela 3.2 exhibe o resultado para alguns valores de  $N$ .

<sup>2</sup>O valor médio de  $t$  quando  $\nu = 1$  é indefinido, conforme será discutido no exemplo 6.1.

Tabela 3.2: Probabilidade do intervalo  $[\bar{x} - \sigma_m, \bar{x} + \sigma_m]$  conter o valor verdadeiro em função do número  $N$  de dados da medida, quando esses dados têm distribuição normal.

$N$	$P(\bar{x} - \sigma_m \leq x_0 \leq \bar{x} + \sigma_m) = \alpha$
2	50%
3	58%
4	61%
5	63%
10	65%
20	67%
$\infty$	68,3%

Podemos, também, fixar o nível de confiança e calcular a largura do intervalo correspondente, ou seja, podemos calcular quanto vale o número  $t_I$  tal que

$$P(\bar{x} - t_I \sigma_m \leq x_0 \leq \bar{x} + t_I \sigma_m) = \int_{-t_I}^{t_I} \psi_\nu(t) dt = 68,3\% \quad ,$$

onde a probabilidade de 68,3% foi escolhida para que  $t_I \xrightarrow{N \rightarrow \infty} 1$ . Podemos definir de maneira análoga o valor  $t_{II}$  com a propriedade que  $t_{II} \sigma_m$  corresponda ao mesmo nível de confiança que o intervalo de mais ou menos dois desvios-padrão em torno da média de uma gaussiana, ou seja, 95,5%:

$$P(\bar{x} - t_{II} \sigma_m \leq x_0 \leq \bar{x} + t_{II} \sigma_m) = \int_{-t_{II}}^{t_{II}} \psi_\nu(t) dt = 95,5\%$$

com  $t_{II} \xrightarrow{N \rightarrow \infty} 2$ . A tabela 3.3 apresenta os valores de  $t_I$  e  $t_{II}$  para alguns valores do número de dados,  $N$ ; uma versão mais completa está no apêndice (Tabela ??).

**Q3.5** Verifique os valores das tabelas 3.2 e 3.3 usando as tabelas ?? e ?? dos apêndices.

Note a propriedade  $t_{II} \neq 2t_I$ , embora  $t_{II} \xrightarrow{N \rightarrow \infty} 2t_I$ , ou seja, a menos que o número  $N$  de dados seja muito grande, “dois desvios-padrão efetivos” não iguala duas vezes o “desvio-padrão efetivo”.

Tabela 3.3: Valores limites de  $t$  que definem os intervalos de confiança de 68,3% e 95,5%, para diversos valores do número de observações na medida.

$N$	$t_I(\alpha = 68,3\%)$	$t_{II}(\alpha = 95,5\%)$
2	1,84	14,0
3	1,32	4,53
4	1,20	3,31
5	1,14	2,87
10	1,06	2,32
20	1,03	2,14
$\infty$	1,00	2,00

Neste capítulo, nos restringimos a discutir a f.d.p. de  $t$  na definição de intervalo de confiança. No capítulo 5 voltaremos a usar a distribuição de  $t$  em testes de hipótese.

### 3.7 Critérios de arredondamento

A dúvida acerca do número de algarismos significativos com que devemos representar a resposta final de uma medida pode ser resolvida com o seguinte critério: arredonda-se o número que representa a grandeza de forma que o último algarismo apresentado corresponda ao último algarismo apresentado no seu desvio padrão, sendo que arredondamos o desvio padrão ao algarismo que represente um valor menor que o desvio padrão do desvio padrão, que calculamos na seção 2.9 para o caso da distribuição normal.

A tabela 3.4 mostra a razão  $\sigma_{\sigma_m}/\sigma_m$ , para diversos valores de  $N$ , número de dados, quando os dados se distribuem como a normal.

Esta tabela explica porque raramente faz sentido apresentarmos dois algarismos significativos para o desvio padrão da média quando o primeiro algarismo é 8 ou 9 — precisaríamos de mais de uma centena de dados para que o desvio padrão fosse conhecido com duas casas! Já quando o primeiro algarismo é 1 ou 2, estamos perdendo informação quando representamos o desvio padrão com um único algarismo significativo mesmo que o número de dados seja menor que 5.

Tabela 3.4: A coluna da direita apresenta a precisão relativa do desvio padrão da média, em função do número de dados na medida, que está na coluna da esquerda. A mesma tabela vale para a relação  $\sigma_\sigma/\sigma$ .

$N$	$\sigma_{\sigma_m}/\sigma_m$
5	0,35
10	0,24
100	0,07
1000	0,02

---

### Exemplo 3.4

---

#### EXEMPLO NUMÉRICO

Na medida da espessura de uma folha de Au por espalhamento de partículas  $\alpha$ , uma experimentadora obteve os seguintes dados:

$$\{ 1,92; 1,93; 2,26; 1,68; 1,70 \} \text{ em } \mu\text{m} \quad .$$

A hipótese dos dados terem f.d.p. normal é razoável neste experimento. O valor verdadeiro será estimado por

$$\bar{x} = \frac{1}{N} \sum_i^N x_i = 1,898 \mu\text{m}$$

e a variância por

$$\sigma^2 = \frac{1}{N-1} \sum_i^N (x_i - \bar{x})^2 = 0,0584 \Rightarrow \sigma = 0,234 = 0,23 \quad ,$$

onde representamos  $\sigma$  com dois algarismos porque  $\sigma_\sigma \sim 0,35 \cdot \sigma = 0,08$ , neste exemplo (um pouco no limite, poderíamos também ficar com um único algarismo). O desvio padrão da média é

$$\sigma_m = \sigma/\sqrt{N} = 0,234/\sqrt{5} = 0,1047 = 0,10 \mu\text{m} \quad ,$$

devendo ser representado com dois algarismos, porque  $\sigma_{\sigma_m} \sim 0,04$ . Consultando a tabela 3.3 (ou a ?? do apêndice), construímos o intervalo que contém o valor verdadeiro com 68% de probabilidade,

$$[\bar{x} - t_I \sigma_m; \bar{x} + t_I \sigma_m] = [1,78; 2,02] = 1,90(12) = 1,90 \pm 0,12 \quad .$$

O intervalo que contém o valor verdadeiro com 95% de probabilidade,

$$[\bar{x} - t_{II}\sigma_m; \bar{x} + t_{II}\sigma_m] = [1,60; 2,20] \quad ,$$

que também pode ser escrito

$$1,9(3) \text{ ou } 1,9 \pm 0,3, \text{ com nível de confiança de } 95\% \quad .$$

Para o resultado com nível de confiança de 95%, o fator  $t$  modificou bastante o intervalo, em relação àquele que seria obtido ignorando a flutuação estatística da estimativa da variância.

**Q3.6** Defina as regras de arredondamento, em função do valor do primeiro dígito decimal do desvio padrão da média, quando o número de dados é  $N = 5$ ;  $N = 100$ ;  $N = 10000$ , quando os dados se distribuírem de acordo com a normal. Convença-se que algumas constantes físicas muito medidas podem chegar a ser apresentadas com um  $\sigma_m$  escrito com 3 casas decimais, especialmente se acontecer da primeira conter o dígito 1.

## EXERCÍCIOS

3.1. Um experimentador obteve a grandeza  $A$  a partir da relação

$$A = x - y$$

tendo obtido para  $x$  e  $y$  os resultados

$$x = 13,4(5)$$

$$y = 10,4(3)$$

$$\text{cov}(x, y) = 0,12 \quad .$$

- (a) Determine o valor da grandeza  $A$  e do seu desvio padrão.
- (b) Determine o valor da grandeza  $B = x + y$  e do seu desvio-padrão.
- (c) Determine a covariância entre  $A$  e  $B$ .

3.2. Uma grandeza cujo valor verdadeiro é aproximadamente 20 é observada com um instrumento que permite obter dados com desvio padrão,  $\sigma$ , aproximadamente igual a 1,5.

Suponha, porém, que  $\sigma$  é estimado dos dados obtidos e não da informação privilegiada do enunciado do problema.

- (a) Se você toma 3 dados, com quantas casas você provavelmente representaria a grandeza? O intervalo de confiança de 68% de probabilidade tem largura  $2\sigma_m$ , ou é preciso calcular o intervalo de confiança usando o valor  $t$  (de Student) adequado?
- (b) Idem item a), para medidas com 20 dados e para medidas com 200 dados.
- (c) Quantos dados precisariam ser tomados para que houvessem 3 algarismos significativos após a vírgula na representação decimal de  $\bar{x}$ ?

3.3. A tabela abaixo apresenta os dados de 20 medidas, cada uma delas com 3 dados.

Tabela. As colunas marcadas “ordem” contêm os números que identificam a medida, cujos dados estão relacionados sob o título “conjunto dos dados”.

ordem	conjunto dos dados	ordem	conjunto dos dados
1	{ 9,8; 7,1; 10,3 }	11	{ 9,8; 10,4; 8,4 }
2	{ 12,9; 12,5; 12,9 }	12	{ 11,3; 10,2; 8,8 }
3	{ 9,1; 8,4; 10,2 }	13	{ 11,1; 12,5; 11,6 }
4	{ 8,6; 10,3; 9,7 }	14	{ 11,7; 7,14; 11,5 }
5	{ 10,4; 11,8; 10,1 }	15	{ 10,6; 11,2; 13,3 }
6	{ 9,5; 9,4; 12,4 }	16	{ 9,6; 6,7; 9,0 }
7	{ 9,5; 8,5; 7,0 }	17	{ 9,7; 11,2; 10,3 }
8	{ 9,8; 9,7; 9,3 }	18	{ 8,0; 10,4; 6,9 }
9	{ 9,8; 7,0; 7,5 }	19	{ 8,6; 9,4; 9,1 }
10	{ 11,3; 10,4; 9,8 }	20	{ 7,1; 13,3; 11,1 }

- (a) Determine para cada medida  $i$  a média,  $\bar{x}_i$ , o desvio padrão,  $\sigma_i$ , e o desvio padrão da média,  $\sigma_{m_i}$ .
- (b) Os dados da tabela foram obtidos por simulação, por meio de um gerador de números aleatórios com distribuição gaussiana. Determine em quantos casos o valor verdadeiro  $x_0 = 10,0$  está no intervalo  $[\bar{x}_i - \sigma_{m_i}, \bar{x}_i + \sigma_{m_i}]$ . Compare com o esperado pela integral da f.d.p. da estatística  $t$  de Student. (Você precisará da função distribuição cumulativa de  $t$  de Student, tabela ?? do apêndice.)



- (c) Calcule o intervalo de confiança de 68% para os grupos de dados (de novo, usando a função distribuição cumulativa de  $t$ ). Determine a fração de grupos em que  $x_0$  está no intervalo  $[\bar{x}_i - t_I \sigma_{m,i}, \bar{x}_i + t_I \sigma_{m,i}]$ .
- (d) Na simulação, usou-se  $\sigma_o = 1,5$ . Determine a fração de medidas em que  $x_0$  está no intervalo  $[\bar{x}_i - \sigma_o/\sqrt{3}, \bar{x}_i + \sigma_o/\sqrt{3}]$ .
- 3.4. Considere o conjunto de observações da intensidade de muons por segundo em um grande detetor de raios cósmicos,

$$\{ 8,8; 7,4; 10,4; 9,9; 10,0 \} \text{ em s}^{-1}$$

- (a) Calcule  $\bar{x}$ ,  $\sigma$ ,  $\sigma_m$ ,  $\sigma_\sigma$  e  $\sigma_{\sigma_m}$ .
- (b) Mostre, em função dos valores obtidos, que você deve considerar significativa apenas uma casa depois da vírgula na representação decimal de  $\bar{x}$ .
- (c) Calcule a probabilidade do valor verdadeiro estar no intervalo  $[\bar{x} - \sigma_m; \bar{x} + \sigma_m]$ .
- (d) Calcule intervalos de valores da intensidade de muons com níveis de confiança iguais a 68,3% e 95,5%.
- 3.5. Um espectrômetro deteta partículas com uma eficiência,  $\epsilon$ , variável com a energia,  $E$ , de maneira que  $\epsilon = \alpha/E$ . Quando  $E$  é dado em MeV,  $\alpha = 1,00(4) \cdot 10^{-2}$  MeV (veja a discussão ao final da seção 3.2). A tabela abaixo dá o número detetado de partículas de cada energia, bem como suas incertezas; ignore erros nas medidas das energias. Nessas condições:

E(MeV)	$\epsilon$	número detetado	$I = \text{número}/\epsilon$	$I/I(4 \text{ MeV})$
2,500		370(25)		1
3,333		870(35)		
4,000		1025(40)		
5,000		210(20)		
7,500		480(28)		
10,000		800(30)		

- (a) Calcule o número de partículas *emitidas* de cada energia,  $I$ .
- (b) Calcule a intensidade relativa às partículas de energia igual a 4,000 MeV,  $I_{\text{relativa}} = I/I(4\text{MeV})$ .

Sugestão: O item b) exige que você considere que as intensidades dadas na coluna marcada  $I$  são covariantes. Você pode, por isso, evitar usar essas intensidades, calculando diretamente a razão

$$\frac{I_1}{I_2} = \frac{n_1/\epsilon_1}{n_2/\epsilon_2} = \frac{\frac{n_1 E_1}{\alpha}}{\frac{n_2 E_2}{\alpha}} = \frac{n_1 E_1}{n_2 E_2} \quad ,$$

o que elimina a dependência comum de  $I_1$  e  $I_2$  em  $\alpha$ , que gera as covariâncias entre as intensidades  $I$ .

- 3.6. Para medir a carga de um capacitor, uma experimentadora mediu a capacitância  $C$  e a voltagem  $V$ , com os seguintes resultados:

$$C = \{0, 95; 0, 93; 1, 24; 0, 95; 0, 85\}, \text{ dados em } \mu\text{F e}$$

$$V = \{288; 303; 320; 291; 316; 296; 287; 298; 299; 309\}, \text{ em volt.}$$

- (a) Determine a carga do capacitor a partir desses dados. Não esqueça de estimar o desvio-padrão.
- (b) Muitas medições semelhantes depois dessa determinaram que os desvios-padrão das medições de  $C$  e de  $V$  são  $0,20 \mu\text{F}$  e  $15 \text{ V}$ , respectivamente. Com essa informação, determine os números de medições de  $C$  e de  $V$ ,  $n_C$  e  $n_V$ , respectivamente, que darão a melhor precisão na carga do capacitor quando se tomam  $n_C + n_V = 15$  dados no total e os valores médios de  $C$  e  $V$  são semelhantes aos dos conjuntos de dados acima. *Adote os valores fornecidos dos desvios-padrão de  $C$  e de  $V$  como verdadeiros para o cálculo deste item.*

- 3.7. Mede-se o perímetro,  $p$ , de um retângulo *perfeito* observando-se independentemente um lado menor e um lado maior, obtendo-se  $a(\sigma_a)$  e  $b(\sigma_b)$ , respectivamente. Primeiro, o experimentador calcula o perímetro pela fórmula

$$p = 2a + 2b \quad ,$$

obtendo a variância de  $p$  como

$$\sigma_p^2 = 4\sigma_a^2 + 4\sigma_b^2 \quad .$$

Depois, resolve calcular o perímetro como

$$p = a + a + b + b \quad ,$$

e obtém a variância de  $p$  como

$$\sigma_p^2 = \sigma_a^2 + \sigma_a^2 + \sigma_b^2 + \sigma_b^2 = 2\sigma_a^2 + 2\sigma_b^2 \quad .$$

Qual dos resultados está correto? Onde um dos cálculos está errado? Como corrigi-lo?

- 3.8. A atividade de uma amostra radioativa foi medida na presença de um ruído de fundo, obtendo-se o resultado  $R_1 \pm \sigma_1$ . Posteriormente, o ruído de fundo foi medido obtendo-se o resultado  $R_2 \pm \sigma_2$ . A covariância entre  $R_1$  e  $R_2$  é nula.

- (a) Determine a matriz de covariância da atividade da amostra (estimada como  $A = R_1 - R_2$ ) e do ruído de fundo  $F = R_2$ .
- (b) Quando a covariância entre  $R_1$  e  $R_2$  for nula, a covariância entre  $A$  e  $F$  será *negativa*. Interprete qualitativamente esse sinal.

- 3.9. Considere  $N$  dados experimentais  $\{x_1, x_2, \dots, x_N\}$  com matriz de covariância

$$\mathbf{V} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{pmatrix} \quad .$$

- (a) Determine a variância da média desses dados
- (b) Mostre que, quando  $N \rightarrow \infty$ , a variância da média tende a  $\rho\sigma^2$ .
- 3.10. O ajuste dos parâmetros  $a$  e  $b$  de uma função  $y = a + bx$  aos pontos experimentais dá origem a uma matriz de covariância entre as estimativas  $\hat{a}$  e  $\hat{b}$  obtidas com a forma

$$\mathbf{V}_{ab} = \begin{pmatrix} \sigma_a^2 & \rho\sigma_a\sigma_b \\ \rho\sigma_a\sigma_b & \sigma_b^2 \end{pmatrix} \quad .$$

(No próximo capítulo discutiremos o ajuste de parâmetros; para este problema, tudo que se necessita é a matriz  $\mathbf{V}_{ab}$ .)

- (a) Determine a matriz de covariância entre dois valores  $y_1$  e  $y_2$  calculados em  $x_1$  e  $x_2$ , respectivamente, com os parâmetros  $\hat{a}$  e  $\hat{b}$  obtidos.
- (b) Verifique que a correlação entre  $y_1$  e  $y_2$  tende a 1 quando  $x_1$  tende a  $x_2$ .
- (c) Mostre que o ponto interpolado com menor variância corresponde a  $x_m = -\rho\sigma_a/\sigma_b$  e sua variância é dada por  $\sigma_m^2 = \sigma_a^2(1 - \rho^2)$ .

# Bibliografia

- [Arfken] Mathematical Methods for Physicists, G.Arfken & H.Weber, Academic Press, 4ª edição (1995)
- [Bard] Nonlinear Parameter Estimation, Yonathan Bard, Academic Press (1974)
- [Benzécri] Histoire et Préhistoire de l'Analyse des Données, J.P.Benzécri, Ed. Bordas, Paris 1982
- [Bevington] Data Reduction and Error Analysis for the Physical Sciences, P.Bevington, McGraw-Hill, 1969
- [Birge] The calculation of errors by the method of least squares, Raymond T. Birge, Phys Rev *40 (1932) 207-227*
- [Conover] Practical Nonparametric Statistics, W.J.Conover, John Wiley & Sons Inc. 1971
- [CRC] Handbook of Tables for Probability and Statistics, CRC
- [Eadie] Statistical Methods for Physicists, W.T.Eadie et al., North Holland Pub.Co. 1971
- [Escoubes] Experimental Signs Pointing to a Bayesian Instead of a Classical Approach for Experiments with Small Number of Events, B.Escoubes, S.De Unamuno e O. Helene, Nuclear Instruments and Methods A257(1987)346
- [Feller] Feller, An Introduction to Probability Theory and its Applications, John Wiley, 2ª Ed. (1957)
- [Feynman] Lectures on Physics Vol.I, Chap.6, Feynman Leighton & Sands

- [Firestone] Analysis of  $\alpha$ ,  $\beta$ , and  $\gamma$  ray emission probabilities, R.B. Firestone, Nuclear Instruments and Methods A286(1990)584
- [Forbes] Forbes, Eric G., *Gauss and the Discovery of Ceres*. Journal for the History of Astronomy. 2 (1971) 195-199.
- [Frieden] Fisher's Information as the basis for the Schrödinger wave equation, B. Roy Frieden, Am. J.Phys. 57(1989)11
- [Geraldo] L.P. Geraldo e D.L Smith, Nuclear Instruments and Methods A290(1990)499
- [Grosser] Morton Grosser, The Discovery of Neptune, Harvard University Press, Cambridge, Massachusetts (1962)
- [Gray] C.G.Gray, Am. J.Phys. 59(1991)282
- [Guimarães-Filho] Z.O. Guimarães-Filho e O. Helene, One Step Self-Calibration Procedure in Gamma-Ray Energy Measurements. Brazilian Journal of Physics, v. 33, n.2, (2003) 280-281.
- [Lyons] How to combine correlated estimates of a single physical quantity, L.Lyons, D.Gibaut e P. Clifford, Nuclear Instruments and Methods A270(1988)110
- [Helene] Tratamento Estatístico de dados em Física Experimental, O.Helene, V. R. Vanin, Ed. Edgard Blücher, 2ª Ed., 1991
- [Helene 83] Upper Limit of Peak Area, O.Helene, Nuclear Instruments and Methods 212(1983)319
- [Helene 84] Errors in Experiments with Small Number of Events, O.Helene, Nuclear Instruments and Methods 228(1984)120
- [Helene 91b] Determination of the Upper Limit of Peak Area, O.Helene, Nuclear Instruments and Methods A300(1991)132
- [Helene 91] O que é uma medida?, O. Helene, Shan.P.Tsai, R.P.Teixeira, preprint IFUSP/P-854 (1990) e Revista de Ensino de Física, Vol.13 p.12, SBF (1991).

- [Helene 93] O.Helene and V.R.Vanin, Nuclear Instruments and Methods A335(1993)227
- [Helene 2013] O. Helene, Método dos Mínimos Quadrados com formalismo matricial, Editora Livraria da Física, São Paulo, 2ª edição (2013).
- [James] A review of pseudorandom number generators, F.James, Computer Physics Communications 60(1990)329-344
- [Kendall] The Advanced Theory of Statistics, M.Kendall, A.Stuart & J.K.Ord, Charles Griffin & Company Limited, London
- [Magalhães] Noções de Probabilidade e Estatística, Marcos N. Magalhães e Antonio Carlos P. Lima, Editora da Universidade de São Paulo - EDUSP, 2011
- [Mannhart] A Small Guide to Generating Covariances of Experimental Data, Report PTB-FMRD 84, Berlin, 1981. ISSN 0341-6666
- [Marquardt] An Algorithm for Least-Squares Estimation of Nonlinear Parameters, D. Marquardt, SIAM J. Appl. Math. 11, 431-441, 1963
- [Merzbacher] Quantum Mechanics, E.Merzbacher, John Wiley & sons, New York 1961
- [Mises] Probability, Statistics and Truth, R.von Mises, Dover, 1955
- [Moralles] M.Morales, P.R.Pascholati, V.R.Vanin and O.Helene, Applied Radiation and Isotopes 46-2(1995)133
- [Mucciolo] E.R.Mucciolo and O.Helene, Nuclear Instruments and Methods A256(1987)153
- [Noether] Introdução à Estatística – Uma abordagem não paramétrica, G.E.Noether, Guanabara Dois, 1983
- [Smith] D.L. Smith, Nuclear Instruments and Methods A257(1987)361
- [Stigler] *Gauss and the Invention of Least Squares*. Stephen M. Stigler, Annals of Statistics, 9 (1981) 465-474 - doi:10.1214/aos/1176345451
- [Vanin 1989] V.R.Vanin e M.Aiche, Nuclear Instruments and Methods A284(1989)452

- [Vanin 1997] V.R.Vanin, G.Kenchian, M.Morales, O.Helene e P.R. Pascholati, Nuclear Instruments and Methods A391(1997)338
- [Vuolo] Fundamentos da Teoria de Erros, J.H.Vuolo, Ed. Edgard Blücher, 1992
- [Youden] Statistical Methods for Chemists, W.J.Youden, John Wiley 1951
- [Zar] J.H. Zar, Appl. Statist. 27(1978)n.3, 280-290