APÊNDICE 03
Uma questão de estatística

De onde vêm as expressões para variâncias e desvios padrão?[37]

Conforme já foi antecipado no primeiro texto de apoio da segunda atividade experimental prevista neste livro, há três tipos de desvio padrão: (1) desvio padrão populacional σ_X; (2) desvio padrão amostral S_X; e (3) desvio padrão da média $S_{\bar{X}}$. Tais desvios padrão possuem significados bem diferentes, que serão discutidos cuidadosamente no presente apêndice.

Os três desvios padrão podem ser calculados a partir de um conjunto de dados, envolvendo uma soma de desvios ao quadrado e estão relacionados à variabilidade desses dados. Quanto mais variados forem os dados, maiores serão os desvios padrão.

Outra característica dos desvios padrão é que todos eles envolvem uma raiz quadrada. Com efeito, é justamente em virtude dessa raiz que qualquer desvio padrão possui sempre as mesmas dimensões da grandeza com respeito a qual é calculado e, portanto, é comensurável com a grandeza e com a sua média. Entretanto há outras medidas de dispersão e uma delas é o *desvio padrão ao quadrado*, que é chamada **variância**.

Variância é o desvio padrão ao quadrado.

Assim como os desvios padrão, as variâncias estão relacionadas à variabilidade de um conjunto de dados: quanto mais variados forem os dados, maiores serão as variâncias. Algumas operações importantes são realizadas diretamente com as variâncias mas não com os desvios padrão[38].Há também três tipos de variância: (1) variância populacional σ_X^2; (2) variância amostral S_X^2; e (3) variância da média $S_{\bar{X}}^2$. A cada tipo de variância corresponde um desvio padrão, uma expressão matemática e um significado bem definido segundo o Quadro 1.

Neste apêndice, preferimos tratar a questão da variabilidade dos dados a partir do conceito de variância que a partir do conceito de desvio padrão.

Quadro 1. Variâncias e seus significados[39].

Nome	Definição Matemática	Significado
Variância populacional	$\sigma_X^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$	Determina a variabilidade de uma população de medidas.
Variância amostral	$S_X^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$	Estima a variância populacional a partir de uma quantidade finita de dados experimentais.
Variância da média	$S_{\bar{X}}^2 = \frac{S_X^2}{\mathrm{n}} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n(n-1)}$	Estima a variância da média de uma amostra de medidas.

[37] Discutir a origem e o significado das diferentes expressões para variâncias e desvios padrão é, com muita razão, considerado fora do escopo de uma disciplina introdutória de Física experimental. Por outro lado, para que o presente livro não ficasse incompleto, optou-se por tratar a questão das diferentes variâncias (e desvios padrão) em apêndice. Entretanto, os autores do presente livro não encorajam a discussão deste apêndice com alunos de primeiro semestre por se tratar de uma leitura muito difícil e não-fundamental. Enfim, para citar este apêndice, acrescente às suas referências:
LIMA JUNIOR, P.; SILVA, M.T.X.; SILVEIRA, F.L. De onde vêm as expressões para variâncias e desvios padrão? In: _____. **Mecânica experimental: Subsídios para o laboratório didático**. Porto Alegre: IF-UFRGS, 2011. p. 101-109. Disponível em <www.if.ufrgs.br/fis1258>. Acesso em 26 set. 2011.

[38] Variâncias podem ser somadas diretamente enquanto desvios padrão precisam ser elevados ao quadrado antes de serem somados. Incertezas são interpretadas como desvios padrão e, justamente por isso, elas precisam ser elevadas ao quadrado antes de serem somadas. Para mais informações, veja o Quadro 1 do Roteiro do Professor (Atividade 02).

[39] As equações foram escritas respeitando a seguinte notação: σ_X é o desvio padrão populacional, S_X é o desvio padrão amostral e $S_{\bar{X}}$ é o desvio padrão da média; μ é a média populacional (ou seja, a média de todas as observações que constituem a população de medidas); N é o número de elementos da população; $\bar{X}$ é a média amostral (a média aritmética de todos os elementos que constituem a amostra); e n é o número de elementos da amostra.

Como é possível perceber, a principal diferença entre as expressões das três variâncias está nos seus denominadores, que podem ser iguais a N, $(n-1)$ ou $n(n-1)$. Neste apêndice, deduzimos de maneira mais ou menos rigorosa a expressão para a variância amostral e para a variância da média a partir da primeira definição da variância populacional. Para que essa dedução faça sentido, é preciso conhecer alguns conceitos básicos de estatística.

População, amostra e outros conceitos básicos

Em estatística, chama-se **população** o conjunto de *todos* os elementos que se deseja estudar – por exemplo, as intenções de voto de *todos* os habitantes de uma região, a estatura de *todas* as palmeiras de dada espécie, os resultados de *todas* as observações que virtualmente podemos fazer de uma grandeza física em condições determinadas. Usualmente, as populações envolvem tantos elementos que fica difícil, inviável ou impossível estudá-los por completo caso se desejasse. Por outro lado, o *censo* (investigação completa de uma população) não é necessário quando desejamos inferir *aproximadamente* algumas características da população alvo.

População é o conjunto de todos os elementos que se deseja estudar.

Chama-se **amostra** um subconjunto da população usado para fazer inferências com respeito a essa população. Por exemplo, para conhecer aproximadamente as intenções de voto de todos os habitantes de uma região, não é preciso consultar cada sujeito individualmente. Sabendo as opiniões de alguns, é possível fazer generalizações (válidas dentro de certos intervalos de confiança) para as opiniões da população. Assim, a análise de uma amostra de dados contorna o problema de não ser sempre possível, viável ou desejável colher todos os dados que constituem uma população. É importante destacar que, em princípio, todas as inferências feitas com respeito a uma população a partir de uma amostra são *aproximadas*. Em outras palavras, todo o processo de inferência envolve alguma incerteza.

Amostra é um subconjunto da população usado para fazer inferências.

Em estatística, a palavra **parâmetro** é utilizada para designar alguma característica (quantitativa) de uma população de dados. São exemplos de parâmetros: (1) a estatura ou renda média de todos os habitantes de uma região; ou (2) a média dos resultados de todas as observações que podemos fazer de uma grandeza física em condições determinadas. Por definição, para saber exatamente os valores dos parâmetros de uma população, é preciso conhecê-la por completo[40]. Entretanto, os valores dos parâmetros de uma população podem ser estimados aproximadamente a partir de uma amostra dessa população.

Parâmetros são características das populações.

Chamamos **estimadores** às estatísticas que, calculadas a partir de uma amostra, permitem inferir aproximadamente os valores dos parâmetros da população representada por essa amostra. Por exemplo, calculando a média das notas obtidas pelos estudantes de uma instituição em uma prova do ENADE[41], é possível inferir quantos pontos, em média, todos os estudantes da mesma instituição fariam se a mesma prova fosse aplicada a cada aluno. *Quando tentamos estimar a média de uma população a partir da média de uma amostra, estamos usando a média amostral como um estimador da média populacional*. Essa estimativa é sempre aproximada de tal maneira que a média da amostra, ora subestima, ora superestima a média da população, dificilmente resultando em um valor idêntico à média da população.

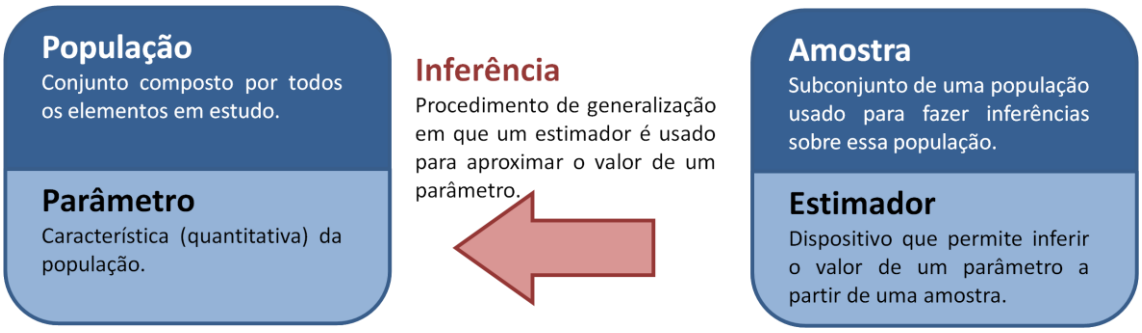
Estimadores são expressões que permitem inferir os valores de parâmetros a partir de amostras.

[40] Em algumas situações se tem conhecimento, baseado em uma *teoria*, sobre os parâmetros populacionais. Por exemplo, para um dado não viciado, a média populacional dos resultados dos lançamentos é 3,5 exatamente.

[41] O ENADE (Exame Nacional do Desempenho de Estudantes) é um dos instrumentos que o Governo Federal utiliza para avaliar os cursos de graduação a partir do rendimento acadêmico dos seus alunos. Como o propósito não é avaliar cada aluno individualmente, a prova do ENADE geralmente não é aplicada a todos os estudantes, mas a uma amostra selecionada aleatoriamente que pode conter tanto alunos ingressantes quanto concluintes. Esse exame tem sido aplicado regularmente em intervalos menores ou iguais a 3 anos desde 2004.

Todos esses conceitos básicos do campo da teoria da amostragem são apresentados resumidamente no Quadro 2.

Quadro 2. Representação esquemática de alguns conceitos básicos de estatística envolvidos na inferência da população a partir de uma amostra.



O viés dos estimadores

Dependendo dos propósitos da análise, alguns estimadores serão mais adequados que outros. A saber, existem vários critérios com respeito aos quais os estimadores podem ser avaliados e classificados. Nesta seção, discutiremos um desses critérios: o viés.

O **viés** de um estimador é definido como a diferença entre o *valor esperado* desse estimador e o *valor do parâmetro* que ele estima. O valor esperado de um estimador, por sua vez, é igual à média de todos os valores que esse estimador pode assumir. Para compreender o que significa o valor esperado de um estimador, é preciso considerar o universo de todas as M amostras com n elementos que se pode obter de uma população com N elementos[42] com $n \ll N$. Ao menos em princípio, conhecendo todos os N elementos da população, seria possível organizá-los em M amostras, calculando o valor do estimador em cada uma dessas amostras. Esse procedimento está representado na Figura 1.

Viés é a diferença entre o valor esperado e o valor do parâmetro.

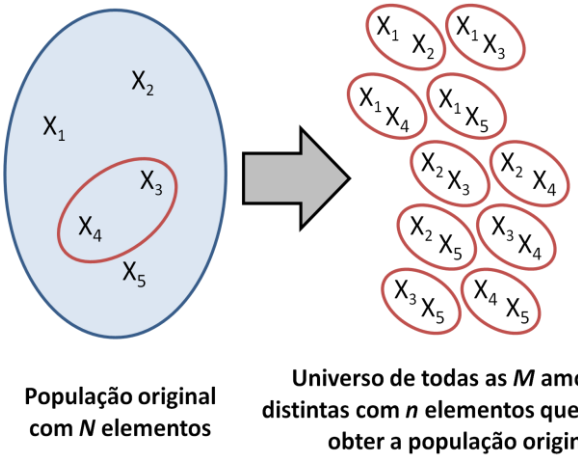


Figura 1. O esquema ao lado representa uma população de *N=5* elementos da qual é possível obter um universo de *M=10* amostras distintas com *n=2* elementos cada. Ao menos em princípio, é possível calcular o valor de um estimador (por exemplo, a média amostral) em cada uma dessas amostras e, a partir daí, avaliar se esse estimador tem viés ou não.

Considere um estimador genérico G do parâmetro populacional γ. Considere também que G_i é o valor assumido pelo estimador G na i-ésima amostra retirada do universo das M amostras. Representando o valor esperado de G por $E(G)$, a definição de valor esperado implica:

$$E(G) = \frac{1}{M}\sum_{i=1}^{M} G_i \ ,$$

[42] A saber, se N e n forem números finitos, a quantidade M de amostras possíveis de se obter deve satisfazer a seguinte equação $M = N!/n!\,(N-n)!$ É importante destacar que, em todas as situações de interesse, M é maior ou muito maior que N. Portanto, se a análise de dados populacionais já era quase sempre inviável ou impossível, a análise de *todas* as amostras que se pode obter de uma população jamais é feita na prática.

pois o valor esperado $E(G)$ é igual à média de todos os valores assumidos por G no universo das amostras. Usando a mesma notação, a definição de viés implica

$$Viés(G) = E(G) - \gamma \ ,$$

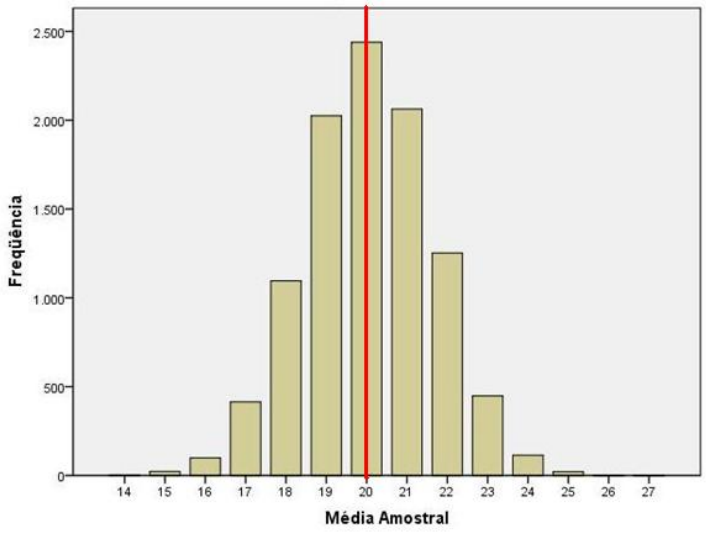
pois viés é a diferença entre o valor do parâmetro e o valor esperado do estimador.

Enfim, um estimador é dito **viesado** quando seu viés é diferente de zero. Igualmente, ele é dito **não-viesado** (ou sem viés) quando seu viés é nulo. Desta forma um estimador é não-viesado quando resulta em valores que não diferem sistematicamente para mais, ou para menos, do valor parâmetro. Ou ainda, os valores do estimador somente diferem do parâmetro por flutuação estatística.

A média amostral: Um exemplo de estimador sem viés

Para ilustrar o significado prático do conceito de viés, simulamos em um programa de computador uma população virtualmente infinita[43] ($N \to \infty$) de números inteiros com média populacional μ igual a 20. Agrupamos esses números em 10.000 amostras distintas com 10 elementos em cada amostra ($n = 10$) e calculamos as médias de cada uma dessas amostras. Tais médias amostrais serão utilizadas para estimar a média populacional μ que, em virtude de estarmos lidando com uma simulação de computador, tem valor conhecido ($\mu = 20$). Essas médias amostrais foram arredondadas para o valor inteiro mais próximo e utilizadas na elaboração do Gráfico 1.

Gráfico 1. Histograma da freqüência das diversas médias amostrais observadas em 10.000 amostras ($n = 10$) tomadas aleatoriamente de uma população com $\mu = 20$.



As barras no histograma[44] no Gráfico 1 representam a freqüência com que cada média amostral foi obtida. A linha vermelha vertical marca a posição da média populacional ($\mu = 20$) que as médias amostrais tentam estimar. A partir da (quase) simetria do gráfico em torno da linha vermelha, é possível perceber que o valor esperado da média amostral (a média das médias amostrais) nesse caso deve ser aproximadamente igual a 20, ou seja, aproximadamente igual à média populacional (parâmetro-alvo). Isso evidencia que a média amostral é um estimador sem viés da média populacional.

[43] Foi utilizado um gerador de números aleatórios com distribuição normal com média igual a 20 e variância igual a 25.

[44] Histograma é como são chamados os gráficos de barras verticais que representam a freqüência com que os valores de uma grandeza foram observados em uma série de muitas observações. Quanto maior for a barra, mais freqüentemente cada valor foi observado.

Uma característica que se pode perceber a partir da simetria[45] do Gráfico 1 é que a média amostral ora subestima, ora superestima a média populacional com a mesma probabilidade. Se o estimador fosse mais propenso a superestimar seu parâmetro-alvo, ele seria um estimador tendencioso. Enfim, estimadores sem viés também podem ser ditos estimadores não tendenciosos e, por essa razão, são utilizados mais freqüentemente em diversas análises estatísticas.

A variância e sua estimativa

Das três expressões de variância apresentadas no Quadro 1, somente a primeira (variância populacional) é, rigorosamente, uma definição. Assim, *define-se* a variância σ_X^2 de uma população de N medidas por:

$$\sigma_X^2 \equiv \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

Variância amostral é um estimador não-viesado da variância populacional.

Nesta expressão, X_i é o *i*-ésimo elemento da população e μ é a média da grandeza X em toda a população de medidas. Com efeito, para saber *exatamente* o valor da variância de uma população de medidas, é preciso conhecer todos os valores que constituem essa população[46]. Entretanto, como foi discutido até agora, deve ser possível construir um estimador (não-viesado) que nos permita conhecer aproximadamente o valor da variância populacional a partir de uma amostra. Tal estimador será chamado **variância amostral**.

Como podemos construir uma expressão para a variância amostral? Bem, podemos tentar uma expressão que nos pareça razoável e testá-la. Se ela for um bom estimador (estimador sem viés), nós a mantemos, se ela for um mau estimador (estimador viesado), nós a corrigimos.

Então vamos propor uma "expressão candidata" à variância amostral (chamada $Cand^2$) que seja o mais semelhante possível à própria definição da variância populacional:

$$Cand^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$$

Nesta expressão, X_i é o *i*-ésimo elemento da amostra, $\bar{X}$ é a média amostral e n é o número de elementos da amostra. Resta agora testar se essa expressão é um estimador não-viesado da variância populacional.

Se o estimador $Cand^2$ possuir viés igual a zero, ele deve satisfazer a equação seguinte:

$$E(Cand^2) = \sigma_X^2$$

Para testar essa afirmação, calculamos o valor esperado do estimador $Cand^2$ conforme segue:

$$E(Cand^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right)$$

Somando e subtraindo a média populacional μ e aplicando o produto notável $(a-b)^2 = a^2 + b^2 - 2ab$, temos:

[45] A simetria do histograma de um estimador não é condição *sine qua non* para que esse estimador seja considerado sem viés. Entretanto, optou-se por analisar um histograma simétrico por motivos didáticos.
[46] A rigor, conhecer a distribuição teórica que, por algum motivo, a população de dados deve satisfazer também permite calcular a variância populacional sem que, para isso, seja necessário conhecer cada elemento da população. Entretanto, essa situação não está sendo levada em consideração no presente texto.

$$E(Cand^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu - \bar{X} + \mu)^2\right)$$

$$E(Cand^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2\right) + E\left(\frac{1}{n}\sum_{i=1}^{n}(\bar{X} - \mu)^2\right) - 2E\left(\frac{1}{n}\sum_{i=1}^{n}(\bar{X} - \mu)(X_i - \mu)\right)$$

Na expressão anterior, considerando que o valor esperado de um estimador qualquer é uma média de todos os valores que esse estimador pode assumir no universo das M amostras possíveis, resulta que:

(1) o primeiro valor esperado é igual à *variância populacional* σ_X^2;
(2) o segundo valor esperado é igual à *variância da população das médias* $\sigma_{\bar{X}}^2$ e, conforme será demonstrado na seção seguinte, pode ser substituído pela variância populacional σ_X^2 dividida pela quantidade de elementos n da amostra;
(3) o terceiro valor esperado, com algumas operações, também resulta na variância da população de médias ($\sigma_{\bar{X}}^2 = \sigma_X^2/n$).

Com isso

$$E(Cand^2) = \sigma_X^2 + \frac{\sigma_X^2}{n} - 2\frac{\sigma_X^2}{n}.$$

Colocando em evidência a variância populacional à direita, temos finalmente que

$$\frac{n}{(n-1)}E(Cand^2) = \sigma_X^2.$$

Com isso, o valor esperado do estimador $Cand^2$ é evidentemente diferente do valor do parâmetro estimado (variância populacional). Portanto, $Cand^2$ é um estimador com viés. Por outro lado, também é possível perceber a partir da expressão acima que, se tomarmos $Cand^2$, multiplicarmos por n e dividirmos por $(n-1)$, obtemos um estimador não-viesado da variância populacional! Enfim, por meio do argumento acima, obtém-se a expressão aceita para variância amostral:

$$S_X^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

Em suma, *a divisão por* $(n-1)$ *na variância amostral* (e, conseqüentemente, em seu respectivo desvio padrão) *resulta de uma correção necessária para que a variância amostral seja efetivamente um estimador sem viés da variância populacional*.

A divisão por (n-1) na variância amostral é uma correção do viés

Definindo a variância da média

Para compreender o significado da variância da média, será preciso considerar, novamente, o universo de todas as M amostras distintas com n elementos que alguém pode obter de uma população de N medidas e calcular a média amostral em cada uma dessas amostras. Devido às flutuações estatísticas inerentes ao processo de amostragem, as médias amostrais terão sempre alguma variabilidade. A **variância da média** é uma medida dessa variabilidade. Ou seja, *a variância da média é uma medida da variabilidade das médias de todas as amostras que se pode obter de uma população*.

A variância da média é uma estimativa da variabilidade das médias de todas as amostras que se pode obter de uma população.

Em geral, quando realizamos muitas observações sob as mesmas condições de um fenômeno natural, consideramos como resultado da

medição a média dos valores observados. Como a variância (e o desvio padrão) da média são estimativas da variabilidade dessa média, é possível usá-los[47] para estimar a incerteza do resultado da medição.

Toda a média aritmética envolve dois procedimentos: (1) em primeiro lugar, toda a média é uma soma de quantidades $(X_1 + X_2 + \cdots + X_n)$; (2) em segundo lugar, toda a média envolve uma multiplicação por uma constante $(1/n)$. Assim, para deduzir a expressão da variância da média, precisamos avaliar primeiro como se comporta a variância de variáveis que resultam de: (1) soma de outras variáveis; (2) multiplicação por uma constante.

Variância da soma e do produto

Considere duas populações de medidas X e Y com N elementos e médias populacionais iguais a μ_X e μ_Y respectivamente. Por definição, as variâncias dessas populações devem ser dadas por:

$$\sigma_X^2 \equiv \frac{\sum_{i=1}^{N}(X_i - \mu_X)^2}{N}$$

$$\sigma_Y^2 \equiv \frac{\sum_{i=1}^{N}(Y_i - \mu_Y)^2}{N}$$

Considere agora que, a partir dessas populações, construímos uma terceira população com N elementos em que o i-ésimo elemento resulta da soma $X_i + Y_i$. Por definição, a variância da variável-soma deve ser dada por:

$$\sigma_{X+Y}^2 = \frac{\sum_{i=1}^{N}(X_i + Y_i - \mu_X - \mu_Y)^2}{N}$$

Aplicando o produto notável $(a+b)^2 = a^2 + b^2 + 2ab$, obtém-se:

$$\sigma_{X+Y}^2 = \frac{\sum_{i=1}^{N}(X_i - \mu_X)^2}{N} + \frac{\sum_{i=1}^{N}(Y_i - \mu_Y)^2}{N} + 2\frac{\sum_{i=1}^{N}(X_i - \mu_X)(Y_i - \mu_Y)}{N}$$

Nessa expressão, as duas primeiras parcelas são quantidades conhecidas (representam as variâncias de X e Y). A terceira parcela, entretanto, representa uma quantidade nova chamada **covariância**.

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2cov(X,Y)$$

A covariância de X e Y é um número positivo ou negativo que representa o quanto as variáveis X e Y estão associadas linearmente. Se uma variação positiva (negativa) de X tende a ser acompanhada por uma variação positiva (negativa) de Y, sua covariância é positiva. Se uma variação positiva (negativa) de X tende a ser acompanhada por uma variação negativa (positiva) de Y, sua covariância é negativa. Se as variações de X e Y não têm nenhuma relação entre si, ou seja, se elas são descorrelacionadas ou estatisticamente independentes, sua covariância deve ser igual a zero.

Considere, por exemplo, dois lançamentos consecutivos de um dado não-viciado de seis faces. Se, no primeiro lançamento, obtemos 5, essa informação não altera a distribuição de probabilidades para o próximo lançamento, isto é, não altera o fato de que no próximo lançamento todos os seis resultados possíveis são equiprováveis. Em conseqüência disso, resultados consecutivos obtidos em um dado são descorrelacionados, ou seja, têm covariância nula.

[47] A rigor, por questões dimensionais, não usamos a variância, mas o desvio padrão da média

Em geral, em uma série de observações da mesma grandeza experimental sob as mesmas condições, os resultados observados também têm covariância nula. Assim, interessa-nos o caso particular em que

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2.$$

Em outras palavras, *a variância de uma variável que resulta da soma de duas variáveis descorrelacionadas é igual à soma das variâncias dessas duas variáveis*[48].

Tendo compreendido o que ocorre com a variância de uma soma de variáveis, passemos agora a analisar a variância de uma variável multiplicada por uma constante. Considere, para isso, somente uma população de medidas X com N elementos e uma constante arbitrária a. Considere que multiplicamos cada elemento de X pela mesma constante positiva a para obter uma nova população de medidas. Por definição, a variância dessa nova população é dada por:

$$\sigma_{aX}^2 = \frac{\sum_{i=1}^{N}(aX_i - a\mu_X)^2}{N}$$

Retirando a constante a dos parênteses, é possível perceber que:

$$\sigma_{aX}^2 = a^2\sigma_X^2$$

Enfim, os dois resultados obtidos até aqui serão utilizados para deduzir a expressão da variância da média: (1) $\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$, quando X e Y são variáveis descorrelacionadas; e (2) $\sigma_{aX}^2 = a^2\sigma_X^2$, em que a é uma constante positiva.

Calculando a variância da média

Por definição, a média $\bar{X}$ de uma amostra com n elementos é dada por

$$\bar{X} = \frac{X_1+X_2+\cdots+X_n}{n}.$$

Sob a hipótese de que as parcelas da soma $X_1 + X_2 + \cdots + X_n$ têm covariância nula e que, pertencendo à mesma população, possuem a mesma variância populacional, resulta

$$\sigma_{X_1+X_2+\cdots+X_n}^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2 + \cdots + \sigma_{X_n}^2 = n\sigma_X^2.$$

Considerando que a média é igual à soma $X_1 + X_2 + \cdots + X_n$ multiplicada por uma constante igual a $1/n$, temos

$$\sigma_{\bar{X}}^2 = \frac{1}{n^2}\sigma_{X_1+X_2+\cdots+X_n}^2 = \frac{1}{n^2}n\sigma_X^2,$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}.$$

Em palavras, a variância da população das médias $\bar{X}$ é igual à variância da população de medidas X dividida pelo número de elementos nas amostras[49]. Acrescentando-se que a variância populacional σ_X^2 pode ser estimada, sem viés, pela variância amostral, chegamos à expressão aceita para a variância da média:

[48] Observe que é exatamente por essa razão que as incertezas são somadas ao quadrado (veja o texto de apoio sobre propagação da incerteza). Por definição, incertezas são quantidades que podem ser interpretadas como desvios padrão. Portanto, dado que a variância de uma soma de variáveis descorrelacionadas é igual à soma das variâncias, resulta que os desvios padrão (e as incertezas) sejam somados uns aos outros sempre elevados ao quadrado!

[49] Observe que essa afirmação foi utilizada anteriormente na dedução da variância amostral neste texto.

$$S_{\bar{X}}^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n(n-1)}$$

Considerações finais

Neste ponto da leitura não deve ser difícil perceber a razão de se abordar a dedução das expressões das variâncias e desvios padrão em apêndice. Trata-se de uma questão muito sofisticada e, embora tenhamos nos esforçado em abordá-la da maneira mais conceitual e agradável possível, a leitura do presente texto ainda deve ser um pouco cansativa para o leitor que não está familiarizado com argumentos do campo da estatística.

Em suma precisamos destacar as seguintes conclusões:

(1) A expressão da **variância populacional** σ_X^2 (e seu respectivo desvio padrão) vale por definição.
(2) A divisão por $(n-1)$ na expressão da **variância amostral** S_X^2 (e do desvio padrão amostral) é necessária para que a variância amostral seja um estimador *não viesado* da variância populacional.
(3) A divisão por $n(n-1)$ na expressão da **variância da média** $S_{\bar{X}}^2$ deve-se, em primeiro lugar, ao fato de que a variância da média é igual à variância populacional σ_X^2 dividida por n, mas, em segundo lugar, deve-se ao fato de que a variância populacional é estimada (sem viés) pela variância amostral.

Essas afirmações são as respostas mais corretas que podemos dar quando nos perguntam por que motivo as expressões dos diferentes desvios padrão são diferentes e com base em que fatos ou argumentos se justificam essas diferenças.