

Por Dentro da Estatística

O uso de métodos estatísticos vem crescendo vigorosamente em pesquisas da área médica. Com frequência, médicos e profissionais da Saúde são expostos a informações provenientes de análises de dados, nem sempre claras e de fácil interpretação. Esta seção visa familiarizar pesquisadores com conceitos e termos estatísticos comumente presentes em artigos científicos. Com ênfase na discussão conceitual em detrimento a fórmulas matemáticas, o objetivo é esclarecer algumas dúvidas frequentes e contribuir com o desenvolvimento do senso crítico na hora de analisar, descrever e interpretar dados.

Ângela Tavares Paes

Editora da seção

Desvio padrão ou erro padrão: qual utilizar?

Standard deviation or standard error: which one to choose?

Ângela Tavares Paes*

* Doutora em Estatística, Centro de Pesquisa Clínica do Instituto Israelita de Ensino e Pesquisa do Hospital Israelita Albert Einstein – IIEPAE/HIAE, São Paulo (SP), Brasil.

Ao descrever dados quantitativos, é comum o pesquisador ficar em dúvida sobre qual medida de variabilidade utilizar: desvio padrão (dp) ou erro padrão. Em textos introdutórios sobre Estatística⁽¹⁻⁴⁾, lê-se que o dp é a medida de variabilidade dos dados individuais em torno da média amostral. No entanto, médias de diferentes amostras também têm sua variabilidade. Geralmente não estamos interessados na média de uma amostra em particular, mas sim na média da população pela qual a amostra foi selecionada⁽²⁾. Cada amostra produz média diferente, sendo interessante saber quanto a média amostral se distancia da média populacional. Pode-se mostrar matematicamente que a variabilidade das médias de amostras de tamanho n é dada pela razão entre o dp e a raiz quadrada de n . Esta quantidade, dp/\sqrt{n} , é denominada erro padrão (EP) e representa a precisão ou incerteza da média de uma única amostra como uma estimativa da média da população. Portanto, dp é a medida de variabilidade individual e EP é a medida de variabilidade associada à média amostral.

Em um artigo científico, obviamente, não é necessário fornecer as duas medidas, pois quando o tamanho da amostra é conhecido, basta saber o dp para calcular

o EP e vice-versa. Qualquer um deles pode ser utilizado e a escolha entre um ou outro deve levar em conta o enfoque com que se pretende analisar os resultados. Se o objetivo é descrever a casuística, o dp parece mais adequado, pois avalia as distâncias dos indivíduos em relação à média do grupo. Assim, pode-se saber se uma amostra de pacientes é mais homogênea ou heterogênea no que se refere à medida de interesse. Caso o objetivo seja fazer inferências sobre as médias, como por exemplo, comparar médias de diferentes populações, o EP ou intervalos de confiança (IC) são considerados mais adequados.

Esclarecida essa questão, é conveniente discutir sobre a forma de apresentação das medidas de variabilidade. É comum, em artigos da área médica, serem apresentados dados quantitativos em médias \pm dp ou EP. Porém, essa representação pode implicar em erros⁽³⁾, pois induz o leitor a somar e subtrair os valores após o símbolo \pm e interpretar os valores calculados como limites de um IC. O IC fornece um intervalo de valores plausíveis para a média populacional e é calculado a partir da média amostral $\pm c \times EP$, no qual c é um coeficiente associado a uma distribuição teórica (normal ou t de Student). Na maioria dos estudos, o valor de c adotado é aproximadamente igual a dois, o que corresponde a um grau de confiança de 95% para o intervalo. Também é comum utilizar o intervalo com média $\pm 2 \times dp$, conhecido como intervalo de referência⁽⁴⁾. Supondo que a distribuição dos dados seja normal, espera-se que 95% dos valores individuais estejam dentro dos limites do intervalo de referência. Nesse contexto, a média $\pm dp$ corresponde a um intervalo de referência de 68%.

Considere, por exemplo, um estudo com 590 indivíduos em que foi observada média de colesterol total de 190 mg/dl e dp de 44 mg/dl. A representação média \pm dp corresponde ao intervalo [146;234], para média \pm EP temos o intervalo [188;192], enquanto que um IC de 95% é igual a [186;194]. Então, se utilizarmos o dp e a amostra for relativamente grande teremos um intervalo muito maior que um IC. Por outro lado, se utilizarmos o EP, a amplitude do intervalo obtido é aproximadamente a metade da amplitude de um IC com o nível usual de 95%. Para evitar erros de interpretação, uma sugestão é apresentar o dp ou EP entre parênteses ao invés do símbolo \pm ⁽⁵⁾.

Em resumo, para descrever medidas de variabilidade não existe certo ou errado, mas é fundamental que o pesquisador saiba identificar qual o propósito

da análise (descritiva ou inferencial) e assim escolher a forma mais adequada para apresentação dos resultados.

REFERÊNCIAS

1. Altman DG. Practical Statistics for Medical Research. London: Chapman & Hall; 1991.
2. Altman DG, Bland JM. Standard deviations and standard errors. *BMJ*. 2005;331(7521):903.
3. Altman DG. Statistics in medical journals: developments in the 1980s. *Stat Med*. 1991;10(12):1897-913.
4. Kirkwood BR, Sterne JAC. Essential medical statistics. 2nd ed. Malden: Blackwell Science, 2003.
5. Strasak AM, Zaman Q, Pfeiffer KP, Göbel G, Ulmer H. Statistical errors in medical research – a review of common pitfalls. *Swiss Med Wkly*. 2007;137(3-4):44-9.