Tamanho de Efeito e sua implicação no Cálculo Amostral

Beatriz Regina Brum[1], Isolde Previdelli[2]

[1]Beatriz Regina Brum
[2]Isolde Previdelli

RESUMO

A importância do tamanho amostral no processo de planejamento, em muitos casos tem sido um pouco negligenciado. Este trabalho apresenta uma revisão dos conceitos básicos para delimitar o tamanho amostral em estudos de ensaios clínicos, onde o tamanho do efeito está relacionado bem com obter um poder de teste satisfatório. Um plano amostral na área de odontologia foi avaliado para amostras pareadas, cujos dados seguem uma distribuição t de Student. O ensaio clínico refere-se a avaliação de deformidades dentomaxilofaciais decorrentes de cirurgias ortognáticas. Uma amostra piloto foi considerada para iniciar o processo de determinação do plano amostral e principalmente indicar os possíveis tamanhos de efeitos ao pesquisador da área para garantir uma pesquisa replicável, assunto tão discutido atualmente. O Tamanho do efeito (ES) estabelecido foi o "d" de Cohen. Os resultados indicaram um poder de teste próximo de 85% para um tamanho de efeito maior ou igual 0,2, gerando uma amostra de 116 indivíduos. Portanto, conclui-se que o planejamento amostral deve ser sempre uma prática interdisciplinar entre o bioestatístico e o pesquisador, uma vez que, nem sempre o tamanho de amostra ideal é o mais viável, visto que tamanhos de efeitos pequenos e poder de testes altos, levam a tamanhos de amostras grandes.

Palavras chave: tamanho de efeito, tamanho amostral e poder do teste.

1 INTRODUÇÃO

Determinar o tamanho da amostra é crucial para desenvolver a pesquisa, tanto do ponto de vista de viabilidade como de validade estatística e biológica. Os casos em que isso não ocorre resulta em trabalho caros e muitas vezes com baixa qualidade de informação, bem como poder do teste baixo, que por consequência admitem pouca credibilidade nos resultados. Valores confiáveis dependem diretamente da análise exploratória inicial dos dados, estimativas como média e desvio padrão e distribuição a qual a amostra se enquadra, são ponta pé inicial para o cálculo amostral de alguns tipos de variáveis.

Segundo Dell (2002) aproximadamente três ou quatro fatores devem ser conhecidos para a determinação do tamanho amostral, entre elas estão o tamanho de efeito que pode ser obtido utilizando a média e variância, a distribuição da amostra, o poder do teste e a significância.

De acordo com Cohen (1988), o tamanho do efeito, "effect size" ou (ES) se definido como o grau ou dimensão em que o fenômeno está presente na população. O tamanho do efeito é exclusivo da amostra coletada, representando a mensuração em unidades de desvio-padrão ou a magnitude a qual o pesquisador pretende que o teste de hipótese seja capaz de detectar. A distribuição se refere ao método de análise estatística que se pretende realizar sendo determinado pelo tipo de delineamento experimental. O poder do teste por sua vez se refere a suposição de probabilidade do pesquisador em relação ou poder de detecção do teste em sinalizar a diferença especificada anteriormente pelo tamanho do efeito, o poder do teste em geral é fixado por convenção, na maioria dos trabalhos científicos é igual ou superior a 80%. O risco α também deve ser levado em conta, conhecido como significância expressa a suposição da probabilidade do erro tipo I (rejeição da hipótese nula quando está é verdadeira), o qual também é controlável pelo pesquisador e na maioria de artigos e trabalhos científicos apresenta os seguintes níveis 0,01, 0,05 e 0,10 (Levini, 2005).

São diversos os métodos aplicados para o cálculo do tamanho amostral, tais procedimentos variam de acordo com características específicas dos dados em estudo. Dell (2002) aponta que experimentos em geral envolvem três tipos de variáveis a serem medidas, entre elas estão as variáveis dicotômicas, contínuas e de tempo para a ocorrência de um evento.

O tamanho de amostra para dados dicotômicos não necessitam de média e desvio padrão, pois o objetivo em geral se baseia na comparação de proporções entre grupos, para amostra de dados contínuos os testes buscam detectar as diferenças entre médias, nesse tipo estudo as médias e os desvios devem ser especificados assim como para amostras de período de tempo e ocorrência de evento nos casos em que o tempo é considerada contínuo, quando codificado o tempo para uma variável dicotômica tais estatísticas não são necessários.

Nesse trabalho foi utilizado uma amostra proveniente de dados que segue uma distribuição t de Studente, cujas medidas repetidas foram realizadas em um mesmo indivíduo, o objetivo do pesquisador se concentrava na constatação de diferença entre médias, foram abordos tamanho de efeito utilizando a média como ponto de partida. Existem diferentes coeficientes de medidas de efeito, como o estudo trata da análise de dois grupos torna-se adequada a utilização de médias padronizadas do grupo "d"constituído pelo g de Hedges, Delta de Glass e "d"de Cohen. Métodos computacionais foram abordados por meio de utilizando do software “R”, sendo esse utilizado para o cálculo do tamanho da amostra bem como para o poder do teste, o software pode ser obtido pela internet (www.r-project.org).

2 Metodologia

A amostra piloto utilizada nesse trabalho provem de dados pareados com alta correlação, as variáveis respostas de interesse do pesquisador seguem uma distribuição t utilizada na análise de diferenças entre médias. Para isso realizou-se uma breve revisão dos vários tamanhos de efeito utilizando médias, assim como seu cálculo e casos aos quais os mesmos devem ser utilizados.

Hedges (1985, p. 78) aponta o g como sendo utilizado quando os grupos apresentam tamanhos diferentes de pequena dimensão, sua estimativa é denotada por:

$$\hat{g} = \frac{m_X - m_Y}{\sqrt{\frac{(n_x-1)\sigma_x^2+(n_y-1)\sigma_y^2}{n_x+n_y-2}}}\left(1 - \frac{3}{4(n_x - n_y) - 9}\right) \tag{1}$$

O Delta de Glass utilizado em dados heterogêneos, em geral se tem um grupo experimental e um grupo controle, a variabilidade utilizada se refere ao grupo controle já que o mesmo não é afetado pelo efeito do tratamento, sendo assim o melhor estimador do parâmetro populacional.

$$\Delta = \frac{|m_t - m_c|}{\hat{\sigma}} \tag{2}$$

O “d” de Cohen pode ser utilizado quando o estudo abrange duas amostras que apresentam grupos independentes e de mesmo tamanho. A diferença neste caso é que o desvio padrão utilizado nesse processo é o desvio-padrão agregada das duas amostras.

$$d = \frac{|m_X - m_Y|}{\hat{\sigma}_{agregada}} \tag{3}$$

$$\hat{\sigma}_{agregada} = \sqrt{\frac{S_1^2 + S_2^2}{2}} \tag{4}$$

Em casos de amostras independentes nas quais os tamanhos das amostras são distintos, faz-se uso do $\sigma_{combinado}$ representando a estimativa do desvio padrão combinado das amostras, $n_1 - 1$ e s_1^2 correspondem ao grau de liberdade e a variância do primeiro grupo, $n_2 - 1$ e s_2^2 ambos do segundo grupo.

$$d = \frac{|m_X - m_Y|}{\hat{\sigma}_{combinado}} \tag{5}$$

$$\hat{\sigma}_{combinado} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \tag{6}$$

Dados dependentes apresentam correlação entre as medidas nos diferentes tratamentos ou momentos. Estudos desse caráter apresentam em geral a existência de correlação positiva entre os grupos. Se existe grande associação de modo em que a correlação entre eles seja grande o desvio se altera e o "d" assume a forma seguinte:

$$d_z = \frac{|m_X - m_Y|}{\sqrt{\hat{\sigma}_X^2 - \hat{\sigma}_Y^2 - 2r\hat{\sigma}_X\hat{\sigma}_Y}} \tag{7}$$

As medidas do ES são definidas por Cohen em três níveis: Pequeno se $0,20 < d < 0,50$, médio se $0,50 < d < 0,80$ e $0,80 < d$ grande. Para esse estudo foram utilizados tamanhos mínimos de $d = 0,2$, o R contém o comando que auxilia na determinação do ES, para isto, foi baixado o pacote "effsize"carregado a biblioteca "effsize" e efetue o comando cohen.d. Tamanho de efeito e poder do teste aproximado de 85% fixados, sendo possível determinar o tamanho amostral, o qual foi necessário baixar o pacote 'pwr', carregar a biblioteca 'pwr' e executar a função "pwr.t.test"estimando assim o tamanho da amostra. Os tamanhos de efeito inferiores a 0,2 tornaram-se inviáveis para a pesquisa por apresentarem tamanhos de amostra demasiadamente grandes, devido a isso o menor tamanho de efeito observado foi de 0,26 que gerou um tamanho de amostra aproximado de 116 indivíduos.

3 Conclusão

A importância do tamanho do efeito no cálculo amostral tem ligação direta ao poder do teste, pois quanto maior for a diferença entre as médias dos tratamentos a ser detectada, maior será o tamanho do efeito e quanto maior o tamanho do efeito, menor será o tamanho da amostra. O dimensionamento da amostra leva a pesquisa ao conhecimento do erro do método adotado, tal conhecimento garante a validação da pesquisa, estudos com tamanho amostral planejado e poder do teste bem definidos garantem maior confiabilidade, já que o risco da probabilidade de não-rejeição da hipótese inicial da pesquisa quando ela é falsa se reduz. O conhecimento do tamanho amostral adequado e do poder em uma pesquisa garantem maior segurança às conclusões dos resultados finais da pesquisa, além disso a obtenção desses dois elementos dá origem à primeira fase do processo de pesquisa.

Referências

[1] DELL, RALPH B., Steve Holleran, and Rajasekhar Ramakrishnan. "Sample size determination."Ilar Journal 43.4 (2002): 207-213.

[2] Hedges, Larry V., and Ingram Olkin. Statistical methods for meta-analysis. Academic press, 2014.

[3] J. L. Devore. Probability and Statistics for Engineering and the Sciences. Cengage Learning, 2015.

[4] J. COHEN. Statistical power analysis for the behavior science. Lawrance Eribaum Association, 1988.