

Fisioter Bras 2019;20(4):460-1  
<https://doi.org/10.33233/fb.v20i4.3217>

## EDITORIAL

### Você entende o valor de P?

Antônio Marcos Andrade da Costa, M.Sc.\*, Alice Miranda de Oliveira\*\*, Jefferson Petto, D.Sc.\*\*\*

*\*Professor do Curso de Educação Física da Faculdade Social, Salvador/BA, \*\*Curso de Fisioterapia da Faculdade Social, Salvador/BA, \*\*\*Professor Adjunto da Escola Bahiana de Medicina e Saúde Pública, Salvador/BA, Professor da Faculdade Adventista da Bahia, Cachoeira/BA, Professor da Faculdade Social, Salvador/BA*

**Correspondência:** Jefferson Petto, Escola Bahiana de Medicina e Saúde Pública (EBMSP), Av. Dom João VI, 275 Brotas 44657-086 Salvador BA, E-mail: gfpecba@bol.com.br; Antônio Marcos Andrade da Costa: antoniomarcoshand@gmail.com; Alice Miranda de Oliveira: licemoliveira@hotmail.com

É perceptível o quanto de dúvidas muitos estudantes e profissionais de saúde têm na sessão análise estatística, ao lerem artigos científicos. Não somente surgem dúvidas, como por vezes sua interpretação é equivocada e não crítica. Entre o que existe de necessário ao entendimento dessa sessão está o que a estatística determina como valor de probabilidade ou valor de P. Diante dessa carência, objetivamos neste texto discorrer de forma breve sobre o valor de probabilidade (valor de P), atizando um pouco mais de luz a esse tema, que é tão dificilmente entendido.

A história relatada no livro “Uma senhora toma chá...” afirma que a ideia de valor de P, nasceu quando em uma tarde de verão em Cambridge, na Inglaterra, em 1920, um grupo de professores universitários, entre eles Sr Ronald Fisher, foram surpreendidos enquanto tomavam seu chá da tarde. Duas senhoras discutiam, e uma delas afirmava ser capaz de identificar se em uma xícara o leite foi servido antes ou depois de colocado o chá. Entusiasmado com o problema das senhoras, um dos professores propôs um experimento: colocar duas xícaras com chá e leite e solicitar a essa senhora dizer se o leite foi colocado na xícara antes ou depois do chá. O professor Fisher não ficou convencido com o experimento do colega, pois, a probabilidade da senhora acertar era a mesma dela errar, ou seja, uma probabilidade de 50%. Intrigado, elaborou um segundo experimento: colocar seis xícaras distribuídas aleatoriamente, em três, o leite seria colocado antes do chá e, nas outras três, o leite seria colocado depois. Nesse momento, a probabilidade da senhora acertar em apenas uma tentativa passou a ser 1/20, ou seja, uma probabilidade de acerto 0,05 [1]. Muitos especulam que surgiu desse acontecido a ideia do nível de significância de P menor que 0,05, que aparece na maioria dos artigos científicos. Mesmo que a autenticidade dessa história nunca tenha sido demonstrada, foi assim que o fundamento probabilístico científico se desenvolveu.

De forma pragmática, o valor de P representa a probabilidade da hipótese nula ser verdadeira, ou seja, que a diferença encontrada, entre o resultado de duas ou mais amostras, seja devido ao acaso, e não devido a fatores que estão sendo estudados. Portanto, o que buscamos ao testarmos hipóteses, é falseá-las e não confirmá-las. Embora a princípio, esse raciocínio pareça contra intuitivo, essa ideia fica mais clara a luz do que Karl Popper [2] postulou. Segundo ele, no modelo científico, primeiro buscamos refutar a hipótese de que existe diferença entre os valores das amostras estudadas, se não conseguirmos confirmar a não diferença (hipótese nula), deduzimos que a diferença existe (hipótese alternativa). Para tanto, utilizamos o valor de P, que nos aponta a probabilidade de estarmos rejeitando a hipótese nula sendo ela verdadeira. De forma universal e convencional, aceitamos que a hipótese nula seja rejeitada quando a probabilidade da diferença encontrada entre os resultados for menor ou igual a 0,05% [3]. Observe que a possibilidade do acaso é diminuída, mas não é neutralizada. Ao contrário do que muitos acreditam, o valor de P é uma forma de se minimizar a incerteza. Vejamos um exemplo.

Vamos imaginar que temos uma suspeita baseada no que estudamos e/ou vivenciamos na prática. Suspeitamos que o exercício que chamaremos de A, aumente mais a força muscular que outro, que chamaremos de B. Para que essa hipótese seja testada de forma confiável e

sistematizada (método científico) construiremos um experimento que evite erros aleatórios. Diante disso, formulamos duas hipóteses:

H0 ou Hipótese Nula (refuta a diferença entre os dois exercícios): Realizar o exercício A não gera maior incremento de força muscular que realizar o exercício B (valor de  $P > 0,05$ ). Pensando de forma prática, a suspeita baseada no que estudamos e/ou vivenciamos não se evidenciou ao se expor a prova científica.

H1 ou Hipótese Alternativa (aceita que a diferença entre os dois exercícios não foi ao acaso): Realizar o exercício A promove maior ganho de força muscular que realizar o exercício B (valor de  $P \leq 0,05$ ). Pensando de forma prática, a suspeita baseada no que estudamos e/ou vivenciamos não ocorre por acaso.

Dessa maneira, se torna evidente que o valor de P é um instrumento útil na análise dos dados. Porém, quando avaliado de maneira isolada pode induzir a conclusões equivocadas sobre o verdadeiro comportamento do fenômeno testado. Segundo Fisher [4], o teste de significância só faz sentido se for realizado no contexto de uma sequência de experimentos. Logo, não faz sentido coletarmos vários dados sem que se tenha objetivos estabelecidos a priori.

O valor de P deve ser considerado dentro de um contexto estatístico e metodológico específico, no qual presumimos que a escolha do teste foi correta, o cálculo de suficiência amostral para a principal variável do estudo foi estipulado de forma adequada, a diferença observada entre as amostras atingiu um poder estatístico de pelo menos 80%, e que essa diferença quando atingida na práxis seja significativa. Ademais, é importante avaliar o valor de P estipulando o intervalo de confiança dos resultados e calculando o tamanho do efeito da intervenção, quando possível. Aventar a veracidade da plausibilidade ou das evidências encontradas antes de se testar a hipótese, estimando minimizar o risco de viés metodológico na execução do trabalho científico, também é um elemento fundamental da análise crítica da significância estatística (valor de P) encontrada em uma pesquisa.

Finalmente, à revelia de todas essas questões, é essencial ratificarmos que sempre existirá a possibilidade de o valor de P não refletir a realidade. Isso porque, como dissemos anteriormente, a estatística não elimina a incerteza totalmente, mas apenas a reduz em níveis consensualmente aceitáveis. Assim, mesmo diante de um experimento bem projetado e executado sempre existirá a possibilidade da hipótese nula ser rejeitada (erro Tipo 1) ou aceita (erro Tipo 2) de forma equivocada. Isto, nos permite refletir um pouco sobre a vida, pois, como na estatística, na vida, nem tudo se resume a um valor de P e invariavelmente em nossa jornada, cometeremos erros Tipo I ou Tipo II.

## Referências

1. Salsburg DS. Uma senhora toma chá... como a estatística revolucionou a ciência no século XX. Rio de Janeiro: Zahar; 2009. p. 17-22.
2. Popper KR. A lógica da pesquisa científica. São Paulo: Cultrix; 2004. p.41-3.
3. Fisher RA. The design of experiments. 8 ed. Edinburgh: Oliver and Boyd; 1966. p.13.
4. Fisher RA. The statistical method in psychical research. Proceedings of the Society for Psychical Research 1929;39:189-92.