

Equações de regressão para predição de diversas variáveis da aptidão física relacionada à saúde.

Na maioria das rotinas de avaliação física realizadas em academias, o principal objetivo do avaliador é caracterizar o estado de aptidão física de seus alunos. Com isso, é possível conhecer fatores de risco, aspectos que devem ser enfatizados e, em casos de reavaliações, o progresso dos programas de atividade física. Geralmente, as variáveis de aptidão física de interesse são aquelas relacionadas à saúde, tais como: flexibilidade, composição corporal, força, resistência e aptidão cardiorrespiratória ($VO_{2\text{ máx}}$).

Em muitos casos, a avaliação direta dessas variáveis pode requerer testes muito sofisticados, com elevado custo e tempo de aplicação. Esses testes são também conhecidos como “padrão-ouro” e têm fundamental importância na área da pesquisa científica. Entretanto, é inviável aplicá-los em rotinas de academias ou sempre que houver demanda por testes rápidos e de baixo custo.

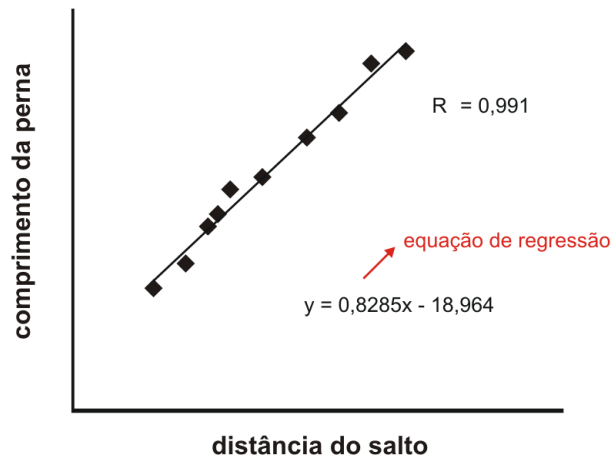
Existem, contudo, alternativas bastante válidas e muito mais viáveis para os testes padrão-ouro. Essas alternativas envolvem testes mais simples, baratos e, via de regra, equações de regressão construídas com base nos resultados dos testes padrão-ouro.

As equações de regressão são construídas para predição de determinada variável com base em outras variáveis conhecidas. Elas são passíveis de serem construídas sempre que houver boa correlação entre duas ou mais variáveis. Para facilitar a compreensão, tomemos o seguinte exemplo hipotético:

- 1) partimos da hipótese de que quanto maior o comprimento das pernas, maior a distância do salto horizontal;
- 2) testamos nossa hipótese em 10 pessoas e obtivemos o seguinte resultado:

sujeito	perna (cm)	salto (cm)
1	55	26
2	58	28
3	60	31
4	62	34
5	61	32
6	65	35
7	69	38
8	72	40
9	75	44
10	78	45

- 3) como pode-se observar na tabela acima, quanto maior o comprimento da perna, maior a distância do salto. Isso significa que há uma **correlação** entre essas duas variáveis. É possível exprimir essa correlação de forma gráfica e matemática, como ilustra a figura abaixo:



Ao construirmos um gráfico plotando em um dos eixos a variável “comprimento da perna” em função da “distância do salto”, no outro eixo, teremos um gráfico de dispersão, no qual é possível visualizar o grau de associação entre duas variáveis quaisquer. Quanto mais horizontal a linha de tendência, menor a correlação; quanto mais próximo da inclinação de 45°, maior a correlação. A correlação ainda pode ser negativa, ou seja, quanto maior uma variável, menor a outra (nesses casos, a inclinação da linha de tendência seria decrescente).

A partir de um conjunto de dados como exemplificado na tabela, pode-se calcular o coeficiente de correlação entre as variáveis. Esse cálculo pode ser facilmente realizado em programas matemáticos e estatísticos. No exemplo acima, usando o coeficiente de correlação de Pearson, obtivemos o resultado mostrado na parte superior do gráfico ($r = 0,991$). O valor de “r”, ou coeficiente de correlação, é a expressão matemática do grau de associação entre as duas variáveis. Esse valor sempre estará em -1 e 1. Quanto mais próximo de 1, maior é a correlação. Quanto mais próximo de 0, menor é a correlação. Quanto mais próximo de -1, maior é a correlação, porém esta será negativa. Ainda utilizando programas matemáticos ou estatísticos, pode-se facilmente calcular a equação de regressão para aquele conjunto de dados, conforme demonstra o gráfico acima. Essa equação é a expressão matemática de predição de uma das variáveis a partir da outra variável. Nesse exemplo, usamos apenas correlação simples, entre duas variáveis. Pode-se também construir regressões a partir de múltiplas variáveis (ex.: peso, idade, estatura e etc.).

Por incrível que pareça, esse conceito é muito importante em rotinas de avaliação física e é também freqüentemente utilizado por avaliadores. Conforme mencionado anteriormente, quase nunca é possível utilizar testes diretos em avaliações físicas de academias como, por exemplo, pesagem hidrostática para determinar o percentual de gordura ou medida de gases expirados para determinar o consumo máximo de oxigênio (VO_{2max}). Entretanto, podemos utilizar somatório de dobras cutâneas e resposta da freqüência cardíaca (FC) a um determinado esforço. Sabendo que existe uma boa **correlação** entre percentual de gordura e somatório de dobras, assim como existe boa correlação entre VO_{2max} e FC, pode-se utilizar **equações de regressão** para, a partir de uma variável conhecida (FC e somatório de dobras cutâneas), **estimar** ou **predizer** variáveis desconhecidas (percentual de gordura e VO_{2max}). Pode-se dizer, portanto, que algumas das mais importantes variáveis utilizadas em rotinas de avaliação física não são medidas efetivamente, mas **estimadas**.

Para complicar um pouco mais, devemos sempre nos atentar à validade das equações de regressão. Isso porque elas somente serão válidas quando empregadas respeitando **todos** os pressupostos necessários, os quais podem variar, dependendo da equação. De forma geral, o mais importante é respeitar as características de população de origem e a forma de aquisição dos dados que deram origem à equação. Exemplificando, as equações de predição do percentual de gordura devem ser aplicadas somente para pessoas da mesma faixa etária, sexo, etnia e índice de adiposidade da população que originou a equação, da mesma forma que as dobras cutâneas devem ser coletadas exatamente da mesma maneira como foram coletadas no trabalho original.

Vale dizer que nenhuma estimativa com base em equações de regressão está isenta de erro. Os estudos de validação demonstram qual o grau de acurácia das equações, ou seja, qual é o erro de estimativa esperado para uma determinada equação de predição. Equações com erro de estimativa muito elevado não possuem acurácia desejável e, portanto, não produzem resultados confiáveis. Caso haja violação dos pressupostos de uma equação, o erro de estimativa aumentará demasiadamente, de modo que aquela estimativa passa a não mais ter validade.

Com base nisso, esteja ciente de que muitas das variáveis que você utiliza não são medidas diretamente, mas preditas com base em equações de regressão. Tais equações sempre trazem um determinado erro, o qual deve ser baixo para que a predição possa ter alguma validade. Logo, opte sempre pelas equações de menor erro de predição. Por fim, respeite sempre todos os pressupostos de validade das equações que você usa, para que possa confiar nos resultados obtidos em sua avaliação física.

Guilherme G. Artioli

Mestrando em Educação Física pela EEFE-USP

Pesquisador do Laboratório de Nutrição e Metabolismo da EEFE-USP

Bacharel em Educação Física pela USP

Bacharelado em Ciências da Saúde pelo ICB-USP