MODELO ESTATÍSTICO PARA DISCRIMINÇÃO LINEAR DE DOIS MÉTODOS DE TREINAMENTO FÍSICO

Maria Emilia Camargo

Departamento de Estatística. Centro de Ciências Naturais e Exatas. UFSM. Santa Maria, RS.

José Elias Rigueira

Departamento de Educação Física. Centro de Ciências Biológicas e da Saúde. UFV. Viçosa, MG.

Gilda B. Sortica

Departamento de Matemática. Centro de Ciências Naturais e Exatas. UFSM. Santa Maria, RS.

RESUMO

Neste trabalho, foi aplicada a análise discriminante para verificar o comportamento da freqüência cardíaca no teste ciclo ergométrico, aplicado a dois grupos de indivíduos, formados de 15 elementos cada um, os quais foram submetidos aos métodos "Circuit-Training" e "Power-Training" de treinamento físico. Determinou-se uma função discriminante com a qual pode-se classificar um novo atleta em um dos grupos de treinamento.

SUMMARY

CAMARGO, M.E.; RIGUEIRA, J.E. and SORTICA, G.B., 1984. Statistic model for linear discrimination of two physical training methods. *Ciência e Natura*, 6:55-58, 1984.

In this paper, the discriminant analysis was applied to verify the behaviour of heart frequency in the ergometric cicle test. It was applied to two groups of subjects, each group containing 15 elements and were submitted to the methods "Circuit-Training" and "Power-Training" of physical training. A discrimination function was arrived at through which a new atheleta may by classified in one of the training groups.

INTRODUÇÃO

A análise discriminante é uma forma muito útil de análise multivariada, cuja principal tarefa é predizer a participação de um determinado elemento em um grupo (2). Neste trabalho, utilizou-se um caso aplicado à medicina do esporte em que um treinador estudou o comportamento da freqüência cardíaca no teste cicloergométrico, aplicado a dois grupos de atletas que foram submetidos a dois diferentes métodos de treinamento físico: "Circuit-Training" e "Power-Training", durante 16 sessões. Através da análise discriminante, buscou-se uma função linear que classifique um novo atleta, em um dos métodos de treinamento para o qual se conhece somente os valores das variáveis: freqüência cardíaca em repouso, freqüência cardíaca após o esforço e a freqüência cardíaca em recuperação.

DESENVOLVIMENTO

Fundamentação Teórica

O modelo estatístico para a discriminação linear, presume que as unidades individuais de duas populações aparecem misturadas e ao acaso. Antes de serem atribuídas é impossível dizer-se a que população uma unidade pertence. Observando-se k características (X_1, ..., X_k) de cada unidade, estas duas populações são medidas k normais multivariadas nas variáveis k (3,4,5).

Admite-se que $\sigma^2_{ii.1}$ seja a variância da i-ésima variável na primeira população; e $\sigma_{ij,1}$ a covariância das i-ésima e j-ésima variáveis na primeira população. Os equivalentes para a segunda população vêm a ser $\sigma^2_{ii.2}$ e $\sigma_{ij.2}$. Sendo que a matriz variância-covariância da primeira população é igual, componente por componente, à matriz variância-covariância da segunda população, ou $[\sigma_{ij.1}]$ = = $[\sigma_{ij.2}]$ para todos os pares (i,j) formados dos números 1,...,k.

Assim, adotou-se a seguinte notação:

σ^2_i = variância da i-ésima variável, em quaisquer das populações;

σ_{ij}= covariância das i-ésima e j-ésima variáveis.

Admite-se que as populações diferem quanto as suas médias, onde, μ_{ij} = média da i-ésima variável (i=1,...,k) na j-ésima população (j=1,2).

Considerando a igualdade das matrizes variância-covariância, as duas populações são idênticas em todos os aspectos, exceto quanto as suas médias. Assim, pode-se constituir duas populações normais univariadas, submetendo cada possível observação k-variadas, nas duas populações, a seguinte transformação linear

$$D = \alpha_1 X_1 + \ldots + \alpha_k X_k \qquad (1)$$

e as populações resultantes terão idênticas variâncias, isto é, $\sigma^2_{D.1} = \sigma^2_{D.2} = \sigma^2_D$, por suas médias, porém, $\mu_{D.1} \neq \mu_{D.2}$.

Os valores numéricos para os α_i são encontrados pela solução do sistema de equações lineares:

$$\begin{array}{l} \alpha_1\sigma_1^2 + \alpha_2\sigma_{12} + \cdots\cdots + \alpha_k\sigma_{1k} = \mu_{11} - \mu_{12} \\ \alpha_1\sigma_{21} + \alpha_2\sigma_2^2 + \cdots\cdots + \alpha_k\sigma_{2k} = \mu_{21} - \mu_{22} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_1\sigma_{k1} + \alpha_2\sigma_{k2} + \cdots\cdots + \alpha_k\sigma_k^2 = \mu_{k1} - \mu_{k2} \end{array} \qquad (2)$$

Assim, encontrando-se α_i desta maneira maximizar-se-á o quadrado da diferença entre as médias das observações transformadas pela unidade de sua variância. Se o quadrado da diferença é um máximo, também o será a diferença por unidade de dispersão. Com o valor de D encontrado em (1), pode-se descobrir o valor crítico C, que é dado pela seguinte expressão:

$$C = (\mu_{D.1} + \mu_{D.2})/2 \tag{3}$$

Com o valor crítico C, pode-se classificar um novo sujeito X em uma das duas populações.

Material

Foram tomadas duas amostras, cada uma formada de 15 atletas que foram submetidos aos métodos "Circuit-Training" e "Power-Training". Em cada elemento de cada amostra, foram medidas as seguintes variáveis: freqüência cardíaca em repouso, freqüência cardíaca após esforço realizado e freqüência de recuperação após 2 minutos, através do teste cicloergométrico.

Os elementos de cada grupo foram submetidos a 16 sessões de treinamento.

RESULTADOS E CONCLUSÃO

Com os dados coletados, encontrou-se a estimativa amostral da função discriminante:

$$D = 0{,}004637\ X_1 - 0{,}001362\ X_2 - 0{,}004781\ X_3 \tag{4}$$

A partir da equação (4), encontrou-se as estimativas das médias das populações "Circuit-Training" e "Power-Training" de valores D. Genericamente,

$$\bar{D}_j = a_1\bar{X}_{1j} + a_2\bar{X}_{2j} + a_3\bar{X}_{3j} \tag{5}$$

determina a estimativa da média da função discriminante para a j-ésima população, encontrando a discriminante das médias $\bar{X}_i$ para a amostra daquele grupo como:

$$\bar{D}_1 = 0{,}004637(73{,}87) - 0{,}001362(140{,}00) - 0{,}004781(112{,}20) = -0{,}384573$$

$$\bar{D}_2 = 0{,}004637(71{,}53) - 0{,}001362(152{,}93) - 0{,}004781(122{,}47) = -0{,}462135$$

Além dos $\bar{D}_j$, necessita-se de uma estimativa do desvio-padrão das duas populações, a fim de determinar as estimativas das probabilidades de classificação correta para um novo elemento, com medidas conhecidas X_1^*, X_2^* e X_3^* em cada um dos grupos estudados.

Pode-se obter tal estimativa substituindo as duas séries de amostras de valores X na equação (4), computando-se assim a variância amostral da função discriminante por:

$$S_D^2 = (\bar{D}_1 - \bar{D}_2)^2/(n_1 + n_2 - 4) \tag{6}$$

em que S_D^2 é o estimador imparcial de σ_D^2. Daí, no caso em estudo tem-se $S_D^2 = 0{,}002983$.

Com os valores de S_D^2, $\overline{D}_1$ e $\overline{D}_2$, pode-se encontrar:

$$\Delta^2 = (\mu_{D1} - \mu_{D2})^2 / (\sigma_D^2) = \frac{((-0,384573)-(-0,462135))^2}{0,002983} = 2,0166$$

Logo, a diferença entre as médias observadas pela unidade de desvio-padrão foi de 1,42. Assim, utilizando a distribuição normal reduzida tem-se a probabilidade de 90% de classificação correta de um novo atleta em um dos grupos de treinamento, quando são conhecidas as variáveis referentes a freqüência cardíaca, ou seja (X_1^*, X_2^*, X_3^*), que constituem o vetor X. Essa técnica pode ser utilizada para outros casos, em que se deseja classificar um indivíduo, por exemplo, num grupo de elementos normais ou em um grupo de hipertensos, através das medidas ecocardiográficas ou eletrocardiográficas.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ARANDA, O.F.J. & MÉNDEZ, I. *Analisis discriminante*. IIMAS, México, 1977.
2. MORRISSON, D.F. *Multivariate Statistical Methods*. New York, McGraw-Hill, 1978.
3. PETERS, W.S. & SUMMERS, G.W. *Análise estatística e processo decisório*. São Paulo, USP, 1973.
4. DIXON, W.J. & MASSEY, F.J. *Introduction to statistical analysis*. New York, Macgraw-Hill, 1960.
5. OSTLE, B. *Estatística aplicada*. México, Limusa-Wiley, 1970.

Recebido em outubro, 1984; aceito em dezembro, 1984.