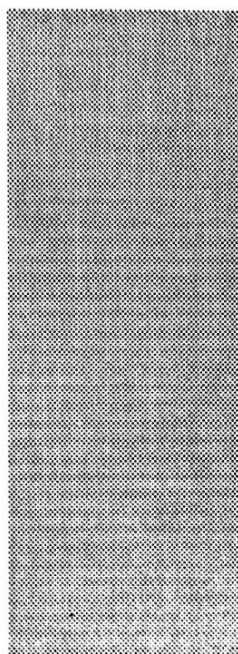


Prof. Valério Gozzano ()*

Uma análise estatística de co-variança

(*) Professor de Estatística na Universidade
de Sorocaba - UNISO



RESUMO

Nem todos os problemas pedagógicos podem ser analisados e, eventualmente, resolvidos somente com boas intenções e propósitos subjetivos. A Estatística permite a verificação e comprovação objetivas que proporcionarão a certeza da existência e extensão do problema e as indicações concretas que ensejam as soluções.

ABSTRACT

Not all pedagogical problems can be analysed and eventually solved only with good intention and subjective purposes. Statistics grants objective checking and confirmation which will lead to the assurance of the existence and extension of the problem and to factual indications to make solutions possible.

Introdução

Várias e variadas situações pedagógicas só merecem investigação se forem suficientemente comprovadas. Este fundamento objetivo é o primeiro passo da pesquisa - a qual será também objetiva, embora com outros critérios.

De 3 classes do 2^a ano do Curso de Administração, na UNISO, o consenso geral indicava: uma, de aproveitamento normal; outra, de aproveitamento insatisfatório e a terceira de aproveitamento muito satisfatório, dispostas no quadro com as indicações III, I, II. As notas, obtidas aleatoriamente, são notas de Direito Comercial, representadas por x e notas de Estatística Aplicada, y .

A opção por disciplinas com poucos pontos de afinidade (princípios filosóficos) e muitos discrepantes (natureza, métodos), mostrou-se providencial, pelo fato da comprovada aleatoriedade das classes em Direito Comercial, ainda que o professor da disciplina considerasse as classe como diversas, provavelmente pelas atitudes perante a matéria.

A confirmação da hipótese se dá já numa primeira aproximação, sem se considerar a possível correlação entre as notas. Com esta consideração, não somente as classes se revelaram claramente distintas, mas permite-se comprovar (apesar das baixas correlações, uma delas inversa), que as classes diferem duas a duas, de modo altissimamente relevante a 99% (I e II; I e III); e a 95% (II e III).

I		II		III	
x1	y1	x2	y2	x3	y3
70	45	85	75	90	55
60	50	80	95	55	60
60	45	20	95	30	60
20	30	50	100	80	95
70	95	60	70	75	95
75	65	85	85	60	90
55	15	85	95	50	75
75	40	80	100	60	75
85	15	70	70	80	75
55	45	50	95	60	65
90	80	40	75	55	75
80	35	65	70	80	85
45	50	95	95	60	25
75	40	50	60	50	95
50	75	90	45	45	85
70	90	50	70	50	80
80	90	80	85	70	65
80	75	65	75	60	45
60	80	80	100	90	95
95	65	85	75	30	50
100	100	65	100	65	70
30	45	45	35	90	100
100	60	60	95	50	75
100	75	40	60	85	70
90	100	95	80	65	45
$\Sigma=1770$	$\Sigma=1505$	$\Sigma=1670$	$\Sigma=2000$	$\Sigma=1585$	$\Sigma=1805$

$$Mx_1 = 70,8$$

$$Mx_2 = 66,8$$

$$Mx_3 = 63,4$$

$$My_1 = 60,2$$

$$My_2 = 80,0$$

$$My_3 = 72,2$$

$$\Sigma(x1+x2+x3) = \Sigma x = 5025$$

$$\Sigma(y1+y2+y3) = \Sigma y = 5310$$

$$\Sigma(x1y1+x2y2+x3y3) = \Sigma xy = 364600$$

$$\Sigma(x1^2+x2^2+x3^2) = \Sigma x^2 = 365125$$

$$\Sigma(y1^2+y2^2+y3^2) = \Sigma y^2 = 411300$$

Correções

$$Cx = 5025^2/75 = 336675$$

$$Cy = 5310^2/75 = 375948$$

$$Cxy = (5025)(5310)/75 = 355770$$

$$x : 365125 - 336675 = 28450 *$$

$$y : 411300 - 375948 = 35352 **$$

$$xy : 364600 - 355770 = 8830 ***$$

$$[(1770^2+1670^2+1582^2)/25]-336625 = 689$$

$$[(1505^2+2000^2+1805^2)/25]-375948 = 4974$$

$$[(1770 \times 1505 + 1670 \times 2000 + 1582 \times 1805)/25] - 355770 = -819$$

$$28450 - 689 = 27761$$

$$35352 - 4974 = 30378$$

$$8830 - (-819) = 9649$$

df ₁ = 2	689	4974	689/2=344,5	4974/2=2487
df ₂ = 72	<u>27761</u>	<u>30378</u>	27761/72=385,6	30778/72=421,9
	28450 *	35352 **	F _x =344,5/385,6	F _y =2487/421,9
			F _x = 0,89	F _y = 5,89

Tabela F (varianças): coluna 2 linha 70: F_{0,05} = 3,13; F_{0,01} = 4,92

0,89 < 3,13 - amostras bastante aleatórias quanto às notas de Direito; 5,89 > 4,92 - amostras não aleatórias quanto às notas de Estatística. Portanto, mesmo não considerando a eventual correlação entre as notas, há uma comprovada diferença entre os procedimentos das classes no que diz respeito à Estatística.

$$35352 - (8230)^2/28450 = 32971,1$$

$$30378 - (9642)^2/22671 = 27024,4$$

$$32971,2 - 270224,4 = 5946,8$$

df ₁ = 2	689	4974	- 819	5542,8	5542,8/2=2771,4
df ₂ = 71	<u>27761</u>	<u>30378</u>	<u>9649</u>	27428,4	27024,4/71=386,3
	28450 *	35352 **	8830 ***		F _{xy} =2771,4/386,3

$$F_{xy} = 7,17 ; F_{0,05} = 3,13 ; F_{0,01} = 4,92$$

Sendo 7,71 > 4,92 a relevância, após considerar a eventual correlação, é bastante superior a 99%.

$$r_{total} = 8830 / \div (28450 \times 35352) = 0,28 \quad b_{total} = 8830/28450 = 0,31$$

$$r_{médias} = -819 / \div (689 \times 4974) = -0,44 \quad b_{médias} = -819/689 = -1,19$$

$$r_{intra} = 9649 / \div (30378 \times 27761) = 0,31 \quad b_{intra} = 9649/30378 = 0,32$$

As correlações resultaram pouco expressivas, sendo uma delas inversa.

$$I : 60,2 - 0,32(70,8-67,0) = 59,0$$

$$II : 80,0 - 0,32(66,8-67,0) = 80,1$$

$$III : 72,2 - 0,32(63,4-67,0) = 75,8$$

	N	Mx	My	Mz
I	25	70,8	60,2	59,0
II	25	66,8	80,0	80,1
III	25	63,4	72,2	75,8

$$E = \sqrt{380,6} \rightarrow E = 19,51$$

$$E_M = 19,51/\sqrt{25} \rightarrow E_M = 3,90$$

$$E_{DIF} = 3,90/\sqrt{2/25} \rightarrow E_{DIF} = 1,10$$

Sendo as amostras pequenas, consultemos Student:

$$df = 71 \rightarrow \text{todos}_{0,05} \approx 2,0; t_{0,01} \approx 2,2$$

$$\begin{array}{ll} \text{Portanto} & D_{0,05} = 2,0 \times 1,10 \quad \rightarrow \quad D_{0,05} = 2,2 \\ & D_{0,01} = 2,2 \times 1,10 \quad \rightarrow \quad D_{0,01} = 3,0 \end{array}$$

I e II: |59,0-80,1| = 21,1 relevante a 99%

I e III: |59,0-75,8| = 16,8 relevante a 99%

II e III: |80,1-73,4| = 4,3 relevante a 95%

Duas das diferenças são altamente relevantes ao nível 99% (0,01), particularmente a das classes extremas de aproveitamento insatisfatório e bastante satisfatório; a outra relevante a 95%.

Comprovadas as relevâncias, o problema pedagógico pode ser investigado: porque a discrepância entre as classes? Um fato meramente ocasional (que seria pouco provável)? A classe I sentia-se censurada e a classe II elogiada (fatos naturais no caso)? Teríamos que reconsiderar a formação das classes, segundo critérios mais pedagógicos? E outras tantas.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ARMITAGE, P. **Statistical Methods in Medical Research**. Oxford: Blackwell, 1971.
2. COSTA NETO, P. L. O. **Estatística**. 14. ed. S. Paulo: Edgard Blücher Ltda, s.d.
3. DIXON, J. W. e MASSEY JR., F. J. **Introduction to Statistical Analysis**. 4. ed. New York : McGraw Hill, 1985.
4. GARRETT, H. E. **A Estatística na Psicologia e Educação**. Trad. M. E. Melo e Cunha e R. Rocha. Rio de Janeiro: Fundo de Cultura, 1962.
5. FONSECA, J. S. e MARTINS, G. A. **Curso de Estatística**. 5ª ed. S. Paulo : Atlas, 1995.
6. VIEIRA, S., HOFFMAN, R. **Estatística Experimental**. São Paulo : Atlas, 1989.