Determinação da média aritmética e desvio padrão de quocientes de séries independentes e dependentes

J. T. A. GURGEL, FREDERICO PIMENTEL GOMES e
A. P. TRIVELIN

E. S. A. "Luiz de Queiroz" U. S. P. — Piracicaba

ÍNDICE

1. INTRODUÇÃO

Embora tenha sido ultimamente pouco discutido, o cálculo da média aritmética e do desvio padrão de produtos e quocientes apresenta grande interêsse estatístico, apesar da nova técnica desenvolvida por GEARY (1930) para o estudo de quocientes. Sôbre o assunto podem ser consultados DAHLBERG (1940), JOHNSON (1951) e YULE e KENDALL (1937).

Neste trabalho apresentamos aos estudiosos a dedução das fórmulas utilizadas no cálculo da média e do desvio padrão de quocientes, e também exemplificamos a aplicação dessas fórmulas. Os quocientes apresentados provêm de medições efetuadas em cavalos da raça Mangalarga.

2. MATERIAL E MÉTODO

Os dados apresentados no Quadro 1 são os que serviram de base para elaboração do presente trabalho. Foram obtidos em cavalos de 2,5 a 3 anos de idade, da raça Mangalarga, com auxílio do bastão de Lydtin, exceptuando-se o comprimento do pescoço que foi tomado com uma fita métrica. O critério que adotamos para mensurações foi o descrito por JARDIM (1952), com ligeiras modificações, conforme as ressalvas mencionadas.

Altura

(a) — Na cernelha — o ponto mais alto desta.
(b) — Na garupa — o ponto mais elevado da garupa.
(c) — Do costado — do ponto mais alto da cernelha ao externo, considerando a mesma vertical em que foi tomada a altura na cernelha (4 e 5).

Comprimento

(a) — Da cabeça — do vértice da cabeça à ponta do focinho.
(b) — Do corpo — da ponta da espádua à ponta do focinho.
(c) — Do pescoço — da nuca ao limite do pescoço com a cernelha.
(d) — Da espádua — da ponta ao alto da espádua.
(e) — Do braço — da ponta da espádua à ponta do codilho.
(f) — Da garupa — da ponta da garupa à das nádegas.

Agradecemos ao prof. F. G. Brieger as sugestões apresentadas durante a elaboração dêste trabalho.

(g) — Da soldra-anca — da ponta da anca à da rótula.
(h) — Da soldra-nádega — da ponta da rótula à ponta da nádega.

Segundo o sistema Eclético de proporções elaborado por LESBRE (1930), para cavalos de sela do tipo médiolíneo, as dimensões das diferentes regiões do corpo animal devem apresentar uma certa proporção (ou quociente) em relação ao comprimento da cabeça que é tomado como base.

Os quocientes observados entre as dimensões das regiões consideradas, e o comprimento da cabeça, estão inclusos no Quadro 2.

3. FÓRMULAS UTILIZADAS

Quando temos uma série de quocientes e desejamos a média aritmética e o desvio padrão, podemos seguir dois caminhos:
a) Fazendo-se prèviamente o cálculo de todos os quocientes individuais $q = \frac{x}{y}$ e, encarando-os como simples variáveis, determinar as estatísticas com as fórmulas conhecidas.

$$(2.1) \qquad \bar{q} = \frac{\Sigma q}{n}$$

$$(2.2) \qquad s(q) = \sqrt{\frac{\Sigma q^2 - \frac{(\Sigma q)^2}{n}}{n-1}}$$

b) Utilizando-se fórmulas especiais, em que figuram as médias de x e de y ($\bar{x}$, $\bar{y}$) e os desvios padrões respectivos s (x) e s (y).

Na ausência de correlação as fórmulas são:

$$(2.3) \qquad \bar{q} = \frac{\bar{x}}{\bar{y}} \left[1 + \frac{s^2(y)}{\bar{y}^2} \right]$$

$$(2.4) \qquad s(q) = \frac{\bar{x}}{\bar{y}} \sqrt{\frac{s^2(x)}{\bar{x}^2} + \frac{s^2(y)}{\bar{y}^2} + 3\,\frac{s^2(x) \cdot s^2(y)}{\bar{x}^2\,\bar{y}^2}}$$

Quando há correlação, as fórmulas são :

$$(2.5) \qquad \bar{q} = \frac{\bar{x}}{\bar{y}} \left[1 + \frac{s^2(y)}{\bar{y}^2} - \frac{Cov(x,y)}{\bar{x}\,\bar{y}} \right]$$

$$(2.6) \qquad s(q) = \frac{\bar{x}}{\bar{y}} \sqrt{\frac{s^2(x)}{\bar{x}^2} + \frac{s^2(y)}{\bar{y}^2} - 2\frac{Cov(x,y)}{\bar{x}\,\bar{y}}}$$

Estas fórmulas são apenas aproximadas, mas dão bom resultado quando é baixo o coeficiente de variação de y (inferior a 30%), e melhor ainda, quando também é baixo o coeficiente de variação de x.

4. DEDUÇÃO DAS FÓRMULAS

Suponhamos que a variável x tem média α e variância σ^2 (x), ao passo que y tem média β e variância σ^2 (y). Seja ainda :

$$q = \frac{\alpha + dx}{\beta + dy} = \frac{\alpha}{\beta} \left(1 + \frac{dx}{\alpha}\right) \left(1 + \frac{dy}{\beta}\right)^{-1}$$

Façamos $u = \dfrac{dx}{\alpha}$, $v = \dfrac{dy}{\beta}$ e ficará:

$$q = \frac{\alpha}{\beta} \left(1 + u\right) \left(1 + v\right)^{-1}$$

Se fôr pequeno o coeficiente de variação de y, v será pequeno, o que justificará que, no desenvolvimento em série de (1 + y) desprezamos as potências de v superiores a segunda.

Obteremos então :

$$q = \frac{\alpha}{\beta}\left(1 + u\right)\left(1 - v + v^2\right) =$$

$$(3.1) \qquad = \frac{\alpha}{\beta}\left(1 + u - v + v^2 - uv + uv^2\right)$$

Se o coeficiente de variação de x também não fôr grande, os valores de u serão pequenos e poderemos desprezar o têrmo uv^2 da última equação, ficando, pois,

$$(3.2) \qquad q = \frac{\alpha}{\beta}\left(1 + u - v + v^2 - uv\right).$$

A média μ de q e, por definição, E(q), onde o símbolo E indica esperança matemática. Logo :

$$(3.3) \qquad \mu = E(q) = \frac{\alpha}{\beta}\left[1 + \frac{\sigma^2(y)}{\beta^2} - \frac{Cov(x,y)}{\alpha\beta}\right]$$

Pois $\quad E(u) = E\left(\frac{dx}{\alpha}\right) = 0$

$$E(v) = E\ \frac{dy}{\beta} = 0$$

$$E(v^2) = E\left(\frac{dy^2}{\beta^2}\right) = \frac{1}{\beta^2}\,E\left(dy^2\right) = \frac{\sigma^2(y)}{\beta^2}$$

$$E(uv) = E\left(\frac{dx\,dy}{\alpha\beta}\right) = \frac{Cov(x,y)}{\alpha\beta}$$

Se substituirmos em (3.3) os parâmetros μ, α, β, σ^2 (y) pelas estimativas respectivas $\bar{q}$, $\bar{x}$, $\bar{y}$, $s^2(y)$ obteremos a fórmula (2.5). E se as variáveis $\bar{x}$ e $\bar{y}$ forem independentes, sua covariância será nula e teremos a fórmula (2.3).

De (3.1) obteremos ainda

$$q^2 = \frac{\alpha^2}{\beta^2}\left(1 + 2u - 2v + u^2 + 3v^2 - 4uv\right)$$

desde que se desprezem os termos de grau superior ao segundo. Daí obteremos então

$$(3.4) \qquad E(q^2) = \frac{\alpha^2}{\beta^2}\left[1 + \frac{\sigma^2(x)}{\alpha^2} + 3\frac{\sigma^2(y)}{\beta^2} - 4\frac{\mathrm{Cov}(x,y)}{\alpha\,\beta}\right]$$

Utilizamos aí a relação não dada acima

$$E(u^2) = E\left(\frac{dx^2}{\alpha^2}\right) = \frac{\sigma^2(x)}{\alpha^2}$$

Agora, sabe-se que

$$\sigma^2(q) = E(q^2) - \left[E(q)\right]^2$$

logo, de (3.3) e (3.4) obtemos

$$\sigma^2(q) = \frac{\alpha^2}{\beta^2}\left[\frac{\sigma^2(x)}{\alpha^2} + \frac{\sigma^2(y)}{\beta^2} - 2\frac{\mathrm{Cov}(x,y)}{\alpha\beta}\right]$$

de onde deduziremos a fórmula (2.6) mediante substituição dos parâmetros pelas estimativas respectivas e abandono dos têrmos de grau mais elevado. Se, a seguir, admitirmos que x e y são independentes, chegaremos à fórmula

$$s(q) = \frac{\bar{x}}{\bar{y}} \sqrt{\frac{s^2(x)}{\bar{x}^2} + \frac{s^2(y)}{\bar{y}^2}},$$

que difere de (2.4). Para obter a fórmula (2.4), mais aproximada, é preciso considerar mais termos no desenvolvimento de q^2, isto é, tomar

$$q^2 = \frac{\alpha^2}{\beta^2} (1+u)^2 (1+v)^{-2} =$$

$$= \frac{\alpha^2}{\beta^2} (1 + 2u + u^2)(1 - 2v + 3v^2) =$$

$$= \frac{\alpha^2}{\beta^2} (1 + 2u - 2v + u^2 - 4uv + 3v^2 -$$
$$- 2u^2v + 6uv^2 + 3u^2v^2)$$

Mas, como u e v agora são independentes,

$$E(uv) = E(u)\,E(v) = 0$$

$$E(u^2v) = E(u^2)\,E(v) = \frac{\sigma^2(x)}{\alpha} \cdot 0 = 0$$

$$E(uv^2) = E(u)\,E(v^2) = 0$$

$$E(u^2v^2) = E(u^2)\,E(v^2) = \frac{\sigma^2(x)\,\sigma^2(y)}{\alpha^2\beta^2}$$

logo,

$$E\,(q^2) = \frac{\alpha^2}{\beta^2}\left[1 + \frac{\sigma^2(x)}{\alpha^2} + 3\,\frac{\sigma^2(y)}{\beta^2} + 3\,\frac{\sigma^2(x)\;\sigma^2(y)}{\alpha^2\,\beta^2}\right]$$

de onde se conclui que

$$\sigma^2(q) = E\,(q^2) - \left[E\,(q)\right]^2 =$$

$$= \frac{\alpha^2}{\beta^2}\left[\frac{\sigma^2\,(x)}{\alpha^2} + \frac{\sigma^2\,(y)}{\beta^2} + 3\,\frac{\sigma^2(x)\;\sigma^2(y)}{\alpha^2\,\beta^2}\right]$$

Se aí substituirmos os parâmetros α, β, σ^2 (x), σ^2 (y) pelas estimativas respectivas $\bar{x}$, $\bar{y}$, $s^2(x)$, $s^2(y)$, teremos obtido (2.4).

5. APLICAÇÃO DAS FÓRMULAS

Conforme tivemos oportunidade de frisar acima, iremos aqui aplicar as fórmulas dadas para uma série de quocientes, obtidos com dados em cavalos da raça Mangalarga, entre algumas medidas consideradas e o comprimento da cabeça.

Os quocientes obtidos encontram-se no quadro 2.

Os valores determinados para as estatísticas, aplicando-se as fórmulas (2.1) e (2.2) e aqueles obtidos pelas fórmulas (2.3) a (2.6) são dados no quadro 3; como aliás já dissemos, para a aplicação das duas últimas fórmulas necessitamos prèviamente conhecer se há ou não correlação entre as séries consideradas. Para sanar esta dificuldade, fizemos também o cálculo da correlação e estudamos a sua significação estatística.

6. RESULTADOS OBTIDOS

Conforme depreendemos pelo quadro 3, os resultados obtidos na determinação da média aritmética e do desvio padrão de quocientes, mostram que os dois processos utilizados oferecem pràticamente o mesmo resultado.

Portanto, não há razão para aplicarmos o processo mencionado em (a), em que os quocientes são considerados separadamente e as estatísticas determinadas segundo as fórmulas (2.1) e (2.2); assim, podemos sem relutância utilizar o processo citado em (b) no qual a média e o desvio padrão dos quocientes são baseados unicamente no valor dessas mesmas estatísticas, porém, de cada uma das séries separadamente. Devemos ainda, atentar para o caso em que as séries sejam independentes ou dependentes entre si.

7. ABSTRACT

The authors prove some approximate formulas for the computation of the mean and the standard error of quotients of two variates, correlated or uncorrelated, with not too high coefficient of variation. The formulas obtained are subsequently applied to some date on mensuration of horses of the Brazilian breed Mangalarga, by the eclectic system of LESBRE. The relsults obtained directly by the actual computation of the quotients as well as by means of the formulas with the aid of statistics of the numerators and the denominators are given in table 3, showing excellent agreement.

8. BIBLIOGRAFIA

DAHLBERG, G., 1940 — Statistical Methods for Medical and Biological Students, 1a. edição, 232 pags. George Allen & Unwin Ltd. Londres.

GEARY, R. C., 1930 — The Frequency Distribution of the Quotient of two Normal Variates. Jour. Roy. Statist. Soc, 93: 442-446.

JARDIM, W. R., 1952 — Exterior e julgamento dos equídeos. 156 pags. 2a. Edição. Piracicaba.

JOHNSON, N. L., 1951 — Statistics, an Intermediate Textbook. 1a. edição, vol. 1, 294 pgs. Cambridge University Press. Cambridge.

LESBRE, F. X., 1930 — Précis d'Exterieur du Cheval. 3a. edição, 631 pags. Vigot Frères, Éditeurs, Paris.

PACI, CORRADO, 1947 — Zoognostica. 867 pags. Instituto Editoriale Cisalpino, Milano.

YULE, G. U. and M. G. KENDALL, 1937 — An introduction to the theory of Statistics. 11a. edição. Charles Griffin and Co. Londres.

QUADRO 1

Dados obtidos das principais medidas de cavalos da raça Mangalarga, com idades variando de 2,5 a 3,5 anos

ALTURA			COMPRIMENTO							
na cernelha	do costado	na garupa	da cabeça	do pescoço	do corpo	da espádua	do braço	da garupa	soldra-anca	soldra-nádega
150	64	148	55	66	147	61	36	50	50	51
147	64	148	59	67	152	60	36	48	52	50
149	68	148	60	64	149	66	36	49	53	50
144	64	143	55	63	144	58	35	50	48	48
151	68	150	59	—	151	66	37	53	45	46
152	67	149	58	63	152	60	36	50	46	46
144	68	143	59	68	151	60	33	50	51	49
150	71	149	60	63	151	64	39	53	48	48
154	70	154	63	65	157	67	38	53	47	48
154	71	153	62	69	152	63	38	54	50	48
145	66	147	61	65	146	58	37	52	49	49
149	68	149	60	66	150	62	36	50	49	46
150	69	149	58	60	152	63	38	52	52	47

QUADRO 2

Quocientes entre as principais medidas consideradas e o comprimento da cabeça Dados obtidos com equinos de 2,5 a 3,5 anos, pertencentes à Raça Mangalarga

ALTURA			COMPRIMENTO						
cernelha	garupa	costado	corpo	pescoço	espádua	**braço**	garupa	soldra-anca	soldra-nádega
2,73	**2,69**	1,16	2,67	1,20	1,11	0,65	0,91	0,91	0,93
2,49	**2,51**	1,08	2,58	1,14	1,02	0,61	0,81	0,88	0,85
2,48	**2,47**	1,13	2,48	1,07	1,10	0,60	0,82	0,88	0,83
2,62	**2,60**	1,16	2,62	1,15	1,05	0,64	0,91	0,87	0,87
2,56	**2,54**	1,15	2,56	—	1,12	0,63	0,90	0,76	0,78
2,62	**2,57**	1,16	2,62	1,09	1,03	0,62	0,86	0,79	0,79
2,44	**2,42**	1,15	2,56	1,15	1,02	0,56	0,85	0,86	0,83
2,50	**2,48**	1,18	2,52	1,05	1,07	0,65	0,88	0,80	0,80
2,44	**2,44**	1,11	2,49	1,03	1,06	0,60	0,84	0,75	0,76
2,48	**2,47**	1,15	2,45	1,11	1,02	0,61	0,87	0,81	0,77
2,38	**2,41**	1,08	2,39	1,07	0,95	0,61	0,85	0,80	0,80
2,48	**2,48**	1,13	2,50	1,10	1,03	0,60	0,83	0,82	0,77
2.72	**2,57**	1,19	2,62	1,03	1,09	0,66	0.90	0,90	0 81

QUADRO 3

Determinação da média e do desvio padrão do quociente das principais medidas consideradas e o comprimento da cabeça, em equinos de 2,5 a 3,5 anos, pertencentes à Raça Mangalarga.

Processo utilizado	Estatísticas	ALTURA			COMPRIMENTO						
		Cernelha	Garupa	Costado	Corpo	Pescoço	Espádua	Braço	Garupa	Soldra anca	Soldra nádega
Cálculo das estatísticas pelas fórmulas (2.1) e (2.2)	$\overline{x}$	2,53	2,51	1,14	2,54	1,10	1,05	0,62	0,86	0,83	0,81
	S	± 0,11	± 0,08	± 0,04	± 0,08	± 0,05	± 0,05	± 0,02	± 0,04	± 0,05	± 0,05
	c. v.	4,4%	3,2%	3.1%	3,1%	4,8%	4,8%	3,9%	4,1%	6,4%	5.9%
Coeficiente de correlação linear	r	0,42	0,64*	0,62*	0,62*	0,29	0,49	0,45	0,50	— 0,03	— 0,16
Cálculo das estatísticas pelas fórmulas (2.3) a (2.6)	$\overline{x}$	2,53	2,51	1,14	2,54	1,11	1,05	0,62	0,86	0,83	0,82
	S	± 0,11	± 0,08	± 0,04	± 0,09	± 0,07	± 0,07	± 0,04	± 0,05	± 0,05	± 0,04
	c. v.	4,6%	3,1%	3.2%	3,6%	6,6%	6,3%	5,9%	5,4%	6,3%	5,2%