Our second data set is taken from a study of the effects of a change in environment on blood pressure in Peruvian Indians (Ryan, Joiner, and Ryan 1985, pp. 317–318). Here we regress systolic blood pressure (y) on weight (x_1) and fraction (x_2). (Fraction is defined as years in new environment divided by age.) The claim is made that these data exhibit suppression and analysis yields $\mathrm{SSR}(x_2|x_1) = 2592.01$ and $\mathrm{SSR}(x_2) = 498.06$, which bears out this contention. We also find that θ here is 141.72° and ϕ is 40.78°. For these values S is positive as we can verify by direct calculation and from Figure 3.

[*Received April 1991. Revised September 1991.*]

REFERENCES

Freund, R. J. (1988), "When is $R^2 > r^2_{yx_1} + r^2_{yx_2}$? (Revisited)," *The American Statistician*, 42, 89–90.

Hamilton, D. (1987), "Sometimes $R^2 > r^2_{yx_1} + r^2_{yx_2}$," *The American Statistician*, 41, 129–132.

——— (1988), (Reply to Freund and Mitra), *The American Statistician*, 42, 90–91.

Mitra, S. (1988), "The Relationship Between the Multiple and the Zero-Order Correlation Coefficients," *The American Statistician*, 42, 89.

Narula, S. C., and Wellington, J. F. (1977), "Prediction, Linear Regression and Minimum Sum of Relative Errors," *Technometrics*, 19, 185–190.

Ryan, B. F., Joiner, B. L., and Ryan, T. A. (1985), *Minitab Handbook* (2nd ed.), Boston; Duxbury Press.

A Geometric Interpretation of Partial Correlation Using Spherical Triangles

GUY THOMAS and JOHN O'QUIGLEY*

This article shows how spherical triangles may be helpful in interpreting and visualizing the relations between partial and simple correlations. The formula giving the partial correlation coefficient in terms of the pairwise simple correlations is seen to be identical with the fundamental formula of spherical trigonometry. The spherical representation is applied to illustrate the masking effect of one variable on another in multiple linear regression.

KEY WORDS: Coefficient of determination; Geometry; Masking variable; Spherical trigonometry.

The power and elegance of the geometric approach to statistics were familiar to early authors (see Herr 1980 for a review), and the advantages of presenting the linear model within the setting of Euclidean n-space geometry have been emphasized in the Teacher's Corner section of this journal (Bryant 1984; Margolis 1979; Saville and Wood 1986).

One often has difficulty in gaining an intuitive grasp of the interrelationships between simple, partial, and multiple correlation coefficients, and geometric concepts have proven particularly helpful here (Hamilton 1987). In this article, after briefly reviewing the Euclidean geometric approach to correlation, it is shown that a construction based on spherical triangles can provide further insight into the relationships between the different orders of correlation. The connection between spherical trigonometry and correlation seems to have been first quoted by Good (1979).

1. EUCLIDEAN GEOMETRY AND CORRELATION

The vector geometric approach to correlation is presented by Bryant (1984). Without loss of generality, all variates are assumed to be measured about their means. Given n observations $x_{11}, x_{12}, \ldots, x_{1n}$ on a variable X_1, we can represent this variable as the vector $\mathbf{x}$ in n-space with coordinates $x_{11}, x_{12}, \ldots, x_{1n}$. This representation leads to the following interpretations. The simple correlation coefficient $r_{X_1X_2}$ between two variables X_1 and X_2 is the cosine of the angle between their representative vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ (Fig. 1).

The multiple correlation coefficient R between a variable Y and two explanatory variables X_1 and X_2 is the cosine of the angle between the representative vector $\mathbf{y}$, with coordinates $y_1, y_2, \ldots, y_n$, and its orthogonal projection y' on the plane spanned by the representative vectors $\mathbf{x}_1$ and $\mathbf{x}_2$, with coordinates $x_{11}, x_{12}, \ldots, x_{1n}$ and $x_{21}, x_{22}, \ldots, x_{2n}$ (Fig. 1).

The partial correlation between Y and X_2 given X_1 is the cosine of the angle between the "residual vectors" $\mathbf{y}'$ and $\mathbf{x}_2'$, that is, the components of y and x_2 orthogonal to x_1 (e.g., Draper and Smith 1966, pp. 201–204). It is shown in Kendall and Stuart (1973) how this operation is geometrically equivalent to projecting the angle between y and x_2 on the plane orthogonal to x_1 (Fig. 2.). The corresponding algebraic formula is

$$r_{YX_2;X_1} = \frac{r_{YX_2} - r_{YX_1}r_{X_2X_1}}{\sqrt{(1 - r^2_{YX_1})(1 - r^2_{X_2X_1})}}. \qquad (1)$$

*Guy Thomas is Professor of Toxicology, Department of Biostatistics and Medical Informatics, Paris 7 University, Paris, France. John O'Quigley is Professor, Department of Mathematics, University of California, San Diego, La Jolla, CA 92093.

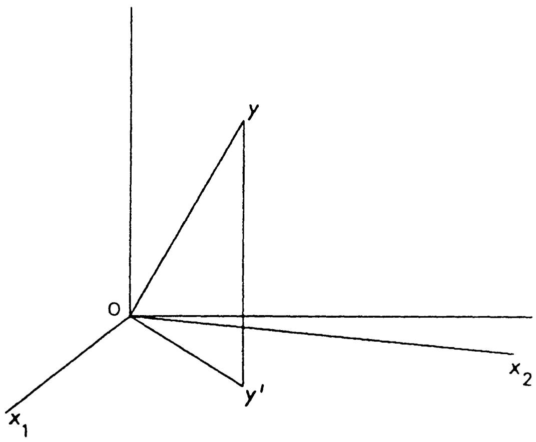

Figure 1. Simple and Multiple Correlations.

2. SPHERICAL TRIANGLES

Firstly we need the concept of a great circle: a great circle on the surface of a sphere is any circle having the same center and same radius as the sphere itself. A spherical triangle is the geometric figure determined in a sphere by three intersecting arcs of great circles (Fig. 3). The sides of the triangle are the arcs of great circles AB, BC, and CA. The vertices are the points of intersection A, B, and C. Angle A of the spherical triangle ABC is defined as the angle between the two tangents at A to the sides of the spherical triangle. As the tangents are perpendicular to OA, this angle is also equal to the dihedral angle between the plane containing O, A, B, and the plane containing O, A, C (e.g., Sommerville 1958). Thus the angle A of the spherical triangle is the projection of the angle between Ox_1 and Ox_2 onto the plane orthogonal to Oy. We denote by a, b, and c the lengths of the three sides BC, AC, and AB.

In the same manner as the sides and the angles of a plane triangle are related to each other, the sides and the angles of a spherical triangle are related by the so-called fundamental formula of spherical trigonometry

$$\cos A = \frac{\cos a - \cos b \cos c}{\sin b \sin c}. \quad (2)$$

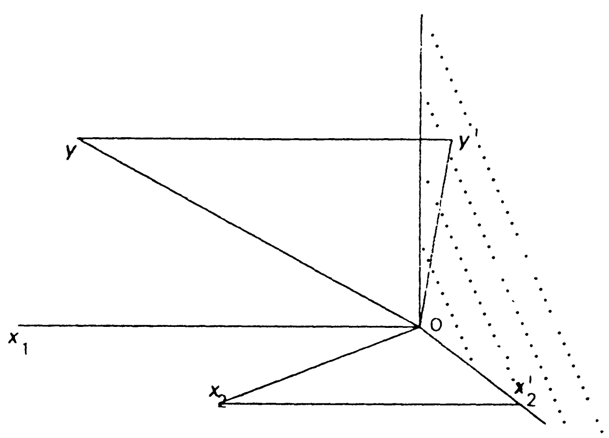

Figure 2. Partial Correlation.

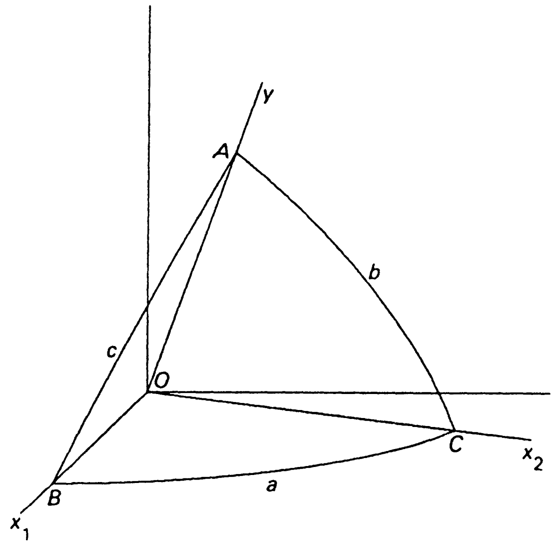

Figure 3. Spherical Triangle.

3. SPHERICAL TRIANGLES AND CORRELATION

Consider three variables Y, X_1, and X_2, and the unit sphere in the three-dimensional subspace spanned by the three variables. The vectors $\mathbf{Oy}$, $\mathbf{Ox}_1$, $\mathbf{Ox}_2$ representing the three variables intersect the surface of the sphere at points A, B, and C, delimiting a spherical triangle ABC with sides a, b, and c (Fig. 3). We agree to take a, b, and c between 0 and π. Now, the radius of the sphere being unity, a, b, and c are also the measures in radians of the angles between Oy, Ox_1, and Ox_2, whence

$$r_{X_1X_2} = \cos a,\ r_{YX_2} = \cos b,\ r_{YX_1} = \cos c.$$

From (1) and (2), it is now clear that:

$$\cos A = r_{X_1X_2;Y},\ \cos B = r_{YX_2;X_1},\ \cos C = r_{YX_1;X_2},$$

so that Figure 3 gives a synthetic pictorial representation of the relations between simple and partial correlations among all three random variables.

4. APPLICATIONS

Using the spherical representation, one may visualize that the partial correlation between X_1 and X_2 given Y can achieve any value between -1 and $+1$, whatever the correlation between X_1 and X_2 provided this is not 1. Indeed partial correlation will be 0 when $A = \pi/2$ (Fig. 4a), and will be ± 1 whenever y lies in the plane Ox_1x_2 spanned by x_1 and x_2 so that $A = 0$ or π (Fig. 4b).

A point that sometimes appears disturbing is the possibility that the coefficient of determination R^2 exceeds the sum of the squares of the simple correlation coefficients (Hamilton 1987). R is geometrically interpreted as the cosine of the angle between Oy and the plane Ox_1x_2. Relationships between R_2 and the elements of the spherical triangle are

$$\begin{aligned} R^2 &= 1 - \sin^2 c \sin^2 B \\ &= 1 - \sin^2 b \sin^2 C \\ &= \frac{\cos^2 b + \cos^2 c - 2 \cos a \cos b \cos c}{\sin^2 a}. \end{aligned}$$

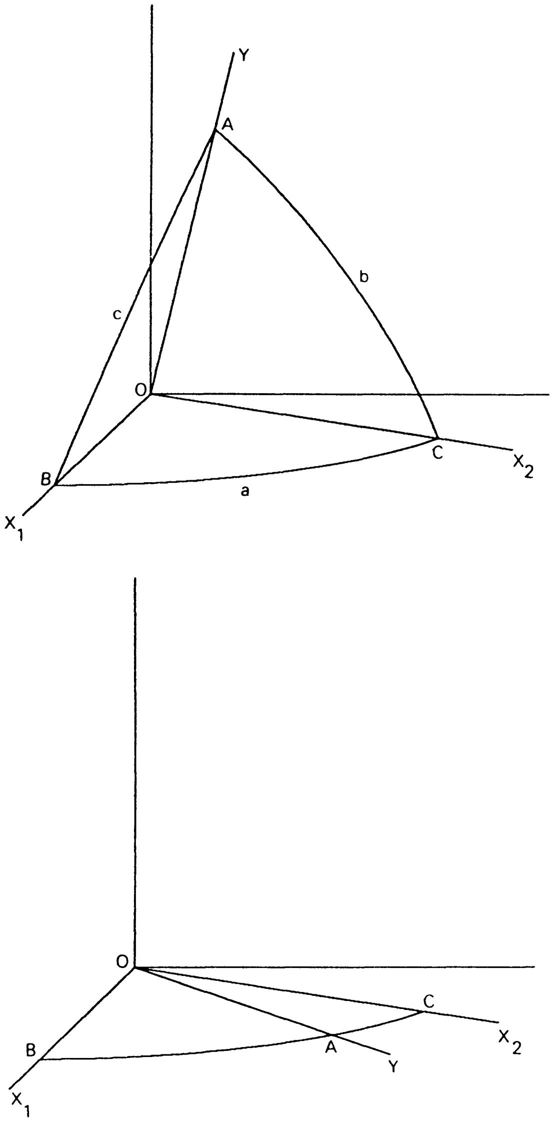

Figure 4. Spherical Representation of Partial Correlation. Partial correlation between X_1 and X_2 given Y is 0 when $A = \pi/2$ (a), and -1 when $A = \pi$ (b).

If y is in the plane spanned by x_1 and x_2, then $R^2 = 1$ always. We can then place Oy at right angles with Ox_1(Y and X_1 uncorrelated), and Ox_2 as close to Ox_1 as we wish (X_1 and X_2 highly correlated), so that the correlation between Y and X_2 is as near to 0 as we want (Fig. 5). However, the partial correlations between Y and X_1 given X_2, or between Y and X_2 given X_1 will be ± 1. This illustrates the masking effect of X_2 on the correlation between Y and X_1 (or the masking effect of X_1 on the correlation between Y and X_2). Obviously, the intensity of the masking effects depends on the intensity of the correlation between X_1 and X_2.

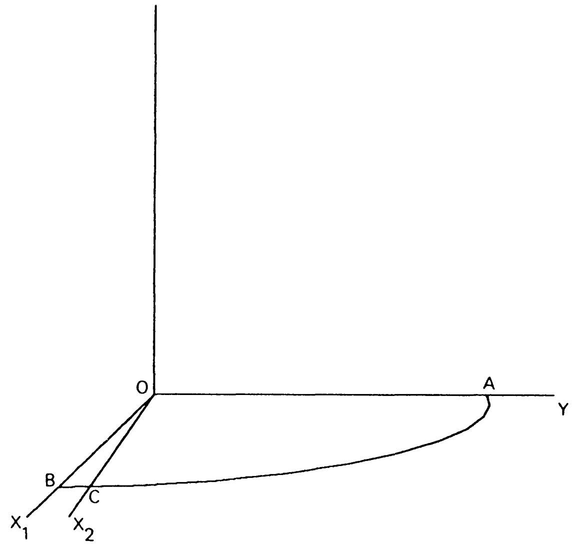

Figure 5. Masking Effect. The multiple correlation coefficient R between Y and the two variables X_1 and X_2 is 1. However, the simple correlation coefficient between Y and X_1 is 0 and the simple correlation coefficient between Y and X_2 is arbitrarily small.

5. CONCLUSION

While not conceptually very different from the classical geometric interpretation of partial correlation, the spherical representation seems easier to handle graphically and more powerful in its ability to illustrate in a single figure all possible relations between three variables.

[*Received October 1991. Revised January 1992.*]

REFERENCES

Bryant, P. (1984), "Geometry, Statistics, Probability: Variations on a Common Theme," *The American Statistician*, 38, 38–48.

Draper, N. R., and Smith, H. (1966), *Applied Regression Analysis*, New York: John Wiley.

Good, I. J. (1979), "Partial Correlation and Spherical Trigonometry," *Journal of Statistical Computing and Simulation*, 9, 243–245.

Hamilton, D. (1987), "Sometimes $R^2 > r^2_{yx1} + r^2_{yx2}$," *The American Statistician*, 41, 129–132.

Herr, D. G. (1980), "On the History of the Use of Geometry in the General Linear Model," *The American Statistician*, 34, 43–47.

Kendall, M. G. (1961), *A Course in the Geometry of n Dimensions*, London: Charles Griffin.

——— (1975), *Rank Correlation Methods*, London: Charles Griffin.

Kendall, M. G., and Stuart, A. (1973), *The Advanced Theory of Statistics* (Vol. 1), London: Charles Griffin.

Loève, M. (1978), *Probability Theory* (Vol. 2, XI, 4th ed.), New York: Springer-Verlag.

Margolis, M. S. (1979), "Perpendicular Projections and Elementary Statistics," *The American Statistician*, 33, 131–135.

Saville, D. J., and Wood, G. R. (1986), "A Method for Teaching Statistics Using N-Dimensional Geometry," *The American Statistician*, 40, 205–214.

Sommerville, D. M. Y. (1958), *An Introduction to the Geometry of N Dimensions*, London: Constable.