Statistics, Energy, and Life-Style

The article "Energy and life-style" by Allan Mazur and Eugene Rosa (15 Nov. 1974, p. 607) contains a number of unfortunate misunderstandings and misuses of statistics. The article is on a topic of major importance—public and private—and its errors mar the continuing discourse about energy.

Perhaps the article's most serious misunderstanding, one that is unhappily widespread, even in the pages of *Science*, is treating null hypotheses not rejected by a significance test as if they were true. In their table 2, Mazur and Rosa leave many blanks where one might expect to see calculated correlation coefficients. They say that only "correlations significant at the .05 level or better are given." In this case, "significant at the .05 level" refers to a conventional significance test procedure for which the null hypothesis is that the true, underlying correlation coefficient is zero.

Yet nonzero hypothetical, and perhaps substantively important, values of the true correlation coefficient may also be at least as compatible with the sample. This is especially noticeable when the sample size is small. For example, with a sample size of 15 (as on the right side of table 2) a sample correlation coefficient of .40 does not differ significantly from zero at the (two-sided) .05 significance level. But neither does .40 differ significantly in the same sense from hypothetical true correlation coefficients between about *minus* .15 and .74.

From a mildly different viewpoint, the above misunderstanding springs from inattention to the power of a significance test.

A second infelicity that arises from significance calculations like those of Mazur and Rosa turns on the assumption of random sampling from a bivariate normal distribution. Of course that assumption need not be exactly true for the calculations to be reasonable, but some kinds of gross deviations from the assumptions could throw the calculations off badly. We might at least be given graphical or other reassurance of approximate bivariate normality.

It is more serious that in fact the Mazur-Rosa numbers are not the realization of proper sampling in any sense, but rather form a complete population of a particular kind. In such a case, the correlation coefficient may provide a useful descriptive summary but conventional significance calculations are thought by many to be jejune, if not downright wrong.

Mazur and Rosa also use the correlation coefficient as a measure of dependence in a casual way. One variable may be an exact function of another and yet the correlation coefficient—which measures only *linear* dependence—may be zero. Again, graphical presentation would mitigate such problems of interpretation.

The misuses and difficulties described here have been pointed out many times in textbooks and journal articles. Readily accessible discussions are in the *International Encyclopedia of the Social Sciences* (*1*), especially in the article by R. F. Tate, "Multivariate analysis, II. Correlation (1)" in volume 10 and in my article "Significance, tests of" in volume 14.

WILLIAM H. KRUSKAL
Department of Statistics, University of Chicago, Chicago, Illinois 60637

References

1. D. L. Sills, Ed., *International Encyclopedia of the Social Sciences* (Macmillan, New York, 1968).

Mazur and Rosa's article contains a glaring example of the misuse of statistical inference and tests of significance. The authors' "sample of 55 countries" is not a random sample but rather the population of all United Nations member nations with a population size of at least 7 million (Communist China, North Korea, and Iran have been omitted from the population because of lack of data, but three smaller countries were included in their place). Yet Mazur and Rosa insist on reporting only those "correlations significant at the .05 level or better." Leaving aside the question of whether looking at such correlations makes any sense to begin with, the use of tests of significance on population data is both inappropriate and misleading. The authors proceed to compound their errors by discarding two extreme "sample" points corresponding to Canada and the United States, and then note that "the majority of correlations drop to insignificance." Why not drop a few other countries and make all the correlations insignificant?

By their misuse of correlation coefficients, the authors have most likely obscured whatever interesting features exist in the data. By their attempt to draw conclusions from cross-national data at one point in time, which can only be based on longitudinal data, they further confuse the unsophisticated reader. By their use of irrelevant tests of significance on correlation coefficients for population data they give a bad name to the rest of us in the social statistics profession.

STEPHEN E. FIENBERG
Department of Applied Statistics, University of Minnesota, St. Paul

The article by Mazur and Rosa gives rise to a number of questions. The authors are dedicated to proving that acceptable life-styles can be maintained in industrial societies without using a tremendous amount of energy, a laudable and worthwhile enterprise. To prove their point, however, they engage in some extremely questionable conceptualization and speculation. The indicators of life-style that they use frequently lack sophistication and often don't measure what the authors believe they do. For example, "general satisfaction" indicators include a disparate collection of variables such as divorces and marriages per capita, population density, sex discrimination in college, and hours worked in manufacturing per week. Obviously divorces per capita is taken to be a sign of "dissatisfaction," but no evidence is offered to make this case. In the absence of convincing evidence it could be argued that divorce is as likely to indicate satisfaction as dissatisfaction; divorced people don't necessarily have to be unhappy over the outcome. Furthermore, there is no theoretical reason to link this indicator to energy usage in any imaginable way.

Similar problems exist with most of the other indicators. Marriages per capita certainly have little, if anything, to do with satisfaction. Particularly in an era of zero population growth, the age of marriage and subsequent reproductive behavior are more closely related to variables other than "satisfaction with life." Similar comments could be made about sex discrimination in high schools and colleges. In these cases satisfaction is obviously in the eye of the beholder. Those discriminated against might be dissatisfied, those doing the discriminating might be very happy indeed. The collected data also indicate that manufacturing work hours per week is negatively related to energy consumption. More properly stated in the reverse—that increased leisure is positively correlated with energy consumption—it would seem that such an indicator should fall under the "economic" heading used by the authors in the last section.

Aside from the faulty use of these and other indicators, the authors are guilty of additional loose reasoning in reaching conclusions that are not warranted from the data. Using a much smaller sample of industrial nations ($N = 15$ to 19), the authors discover that there are fewer significant correlations between energy usage and indicators of life-style for this small sample than there are for all nations. Little wonder that it is difficult to reach a .05 level of significance with such a small sample when all of the previous analysis has been based on an N three times as large. Furthermore, the variance in energy consumption among these highly industrialized countries is not that great. The United States is by far the largest per capita consumer and the rest of these countries cluster remarkably close together, significantly restricting possibilities for significant findings. The authors ignore climatological differences and differing patterns of industrial and agricultural development that could readily account for any variance in energy usage, especially given comparatively small differences within this sample.

In summary, the authors make a point that surely needs to be made, that energy consumption and quality of life are not linearly related, but they make their argument so poorly that it loses its effect. Suffice it to say that they offer no proof that "so long as America's per capita energy consumption does not go below that of other developed nations, we can sustain a reduction in energy use without long-term deterioration of our indicators of health and health care, of education and culture, and of general satisfaction." Making this argument cogently requires much more data than the authors present. Surely social scientists are capable of bringing more exacting scholarship to bear on this very important question.

DENNIS PIRAGES
Program in Science, Technology, and Public Affairs, University of California, San Diego, La Jolla 92037

Kruskal's "most serious" point about the inaccuracy of small-sample estimates of population parameters is certainly correct, but it is irrelevant here. Our "samples" are nearly complete populations of countries, so our "estimates" of correlation coefficients are nearly perfect.

He is again correct in pointing out that correlation coefficients are appropriate only for linear relationships. We assure him that we did do graphical checks for linearity, but it would have been cumbersome to present 224 graphs with the article, as he suggests.

We agree that it is not rigorous to apply significance tests to complete populations, however such tests are commonly used in this manner as a traditional criterion for distinguishing "small" from "large" correlation coefficients. (For sample sizes ranging from 15 to 55, as in our case, the boundary between "small" and "large" correlations ranges from $r = .24$ to .42, but it was usually about .3.) In retrospect, we would have been wiser to avoid significance tests, as they muddied the water, but the issue is a very minor one since all of the correlations deleted as "insignificant" from our table 2 are obviously small. For nations with developed market economies, which were the focus of our analysis, only 3 of the 70 "insignificant" correlation coefficients have an absolute magnitude exceeding .35; about half do not even exceed .20. It must surely be clear that, among developed market economy countries, most noneconomic indicators of life-style show low correlation with energy consumption. That, of course, was our main conclusion.

Fienberg vociferously objects to our treatment of extreme values for Canada and the United States, apparently ignoring our rationale (in our reference 6). He suggests that our treatment improperly reduces the correlation coefficients. In fact, the treatment raises as many correlations as it reduces.

We feel that our original article provides an adequate discussion of additional points raised by Pirages.

ALLAN MAZUR
EUGENE ROSA
Department of Sociology, Syracuse University, Syracuse, New York 13210

Green Thumbs

The hybrid *Chlorella*-animal cell as proposed by N. Nashed (Letters, 27 Dec. 1974, p. 1159) is a very interesting idea, and I hope that someone will be able to pursue this potentially valuable resource. Further information dealing with similar types of naturally occurring relationships has been reviewed for the algae endosymbionts of the Mollusca (*1*). In addition, Nass (*2*) successfully incorporated higher plant chloroplasts into mouse fibroblast cells for short periods of time. Using the information obtained from these studies, I would like to carry the Nashed proposal one step further: eliminate the middleman, that is, the farmer that would grow the hybrid cells, by incorporating chloroplasts or algae cells directly into the skin of humans. This relationship would be much more efficient.

TOD R. FAIRBANKS
Clinical Immunology Lab, Medical Center Hospital of Vermont, Burlington 05401

References

1. L. Muscatine and R. W. Greene, *Int. Rev. Cytol.* **36**, 137 (1973).
2. M. M. K. Nass, *Science* **165**, 1128 (1969).

Hemoglobin in Humans

Thomas H. Maugh II (Research News, 17 Jan., p. 154) indicates that an average-sized human might contain about 1.5 kilograms of hemoglobin. This appears to be in error by a factor of about 2. An average man might contain about 15.5 grams of hemoglobin per 100 milliliters of blood; an average woman has about 12 percent less. It has been known since 1856, when Bischoff (*1*) determined the amount of blood in a condemned criminal, by weighing him before and after decapitation (by guillotine) and washing out residual blood, that total blood mass constitutes approximately 7.7 percent of the total weight. Thus a man weighing 70 kilograms would have about 5.4 liters of blood and 837 grams of circulating hemoglobin. South American miners adapted to living at 5,340 meters (17,500 feet) above sea level may have only 22.9 grams of hemoglobin per 100 milliliters of blood (*2*), which would give 1.24 kilograms (about 20 percent less than the figure given by Maugh), but these miners hardly represent the average. Mountaineers acclimatized to 5,790 meters in the Himalayas contain an average of 20.2 grams of hemoglobin per 100 milliliters of blood (*3*), but only in individuals with pathological polycythemia are values higher than these recorded.

AUSTEN RIGGS
Department of Zoology, University of Texas, Austin 78712

References

1. T. L. W. Bischoff, *Z. Wiss. Zool.* **7**, 336 (1856).
2. D. B. Dill, J. H. Talbot, W. V. Consolazio, *J. Biol. Chem.* **118**, 649 (1937).
3. L. G. C. Pugh, *J. Physiol.* **170**, 344 (1964).