Inferences about correlations¹

LEE ROY BEACH, UNIVERSITY OF WASHINGTON
THOMAS S. SCOPP, UNIVERSITY OF MICHIGAN

The study examined intuitive inferences about population correlations on the basis of samples of data from the populations. The Ss saw samples of bivariate observations, inferred the signs of the population correlations, and stated their confidence in the inferences. Optimality of inferences and of confidence both increased as the magnitude of the sample correlations increased.

Correlated events are an important part of everyday experience and, therefore, Ss' abilities to detect and to learn correlations have received a good deal of experimental attention. The usual method is to take successive random samples from a multivariate population, tell S the sampled values of $n-1$ of the variables (cues) in each sample, and require prediction of the value for n th variable (criterion). The correspondence between his trial-by-trial predictions and those that would be made using multiple regression permits evaluation of the optimality of S's performance.

For both the multiple regression model and for S, adequate prediction depends upon prior inferences about the correlations between the cue and criterion dimensions. However, the procedure of requiring trial-by-trial predictions inextricably mixes S's inferences about correlation with his ability to use the inferred correlations to make predictions.² Separation of intuitive inferences about correlations from predictive performance demands a different experimental procedure and a different optimal model.

It would be convenient to obtain intuitive inferences of correlation coefficients, but correlation is a sophisticated concept and is difficult to explain to naive Ss. Therefore, this experiment used a simpler procedure. Samples from bivariate populations were presented and Ss inferred the sign of the population correlation and then stated their confidence that the inference was true.

Bayesian statistics prescribes optimal performance in this task (Lindley, 1965). Before the sample is observed the probabilities are equal that the population correlation is positive or that it is negative. The evidence provided by the sample leads to revision of these (prior) probabilities and to a new (posterior) probability distribution. The larger the magnitude of the sample correlation the more the posterior probabilities favor sign of that correlation. Assuming that Ss follow a similar procedure, their optimal behavior is to infer that the sign of the population correlation is the one favored by their posterior subjective probability distribution and to state the associated subjective probability as their confidence.

Method

Apparatus. The stimuli ostensibly were samples of $n=20$ index cards from each of 10 population decks of

$N=50$. However, to insure that all Ss saw the same samples, each deck actually consisted of the 20 sample cards on top of 30 blank cards. There was a black and a green number written on each card. The numbers ranged from 1-10 for both variables, no number was ever paired with itself, and each of the 10 pairs appeared twice to make $n=20$. The Pearson product-moment correlations between the variables were $\pm.14$, $\pm.15$, $\pm.30$, $\pm.50$, and $\pm.85$ for the 10 samples. To produce negative correlations, one of the variables was reversed in the positively correlated decks, i.e., 10 became 1, 9 became 2, etc.

Answer booklets contained 100-point response scales that had .50/.50 in the center and numbers increasing in units of .05 to 1.00 at both ends. The left side of the scale was labeled Negative and the right side was labeled Positive. Ss inferred the signs of the correlations by selecting the appropriate side of the scale. Then they marked the scale to indicate their confidence; .50/.50 for complete uncertainty, 1.00 for certainty, and intermediate values for intermediate confidence. Confidence in an incorrect inference was scored as confidence of less than .50 for the correct sign.

Procedure. Instructions stated that Ss would see the sequential presentation of the first 20 pairs of numbers from shuffled decks of 50. Then they were to infer whether there was a positive or negative correlation between the numbers in the deck of 50, and state their confidence. Stress was put on the necessity to estimate confidence rather than degree of correlation and upon using intuition rather than computations. While recording their responses, Ss covered their previous answers with a blank paper. There was no feedback about either the sign or the magnitude of the correlations.

Correlation was explained as the degree to which one set of numbers corresponds to another: "positive... means that high numbers of one set generally are paired with high numbers of the other set and vice versa; negative... means that high numbers of one set generally are paired with low numbers of the other set and vice versa." This was illustrated with correlations of ± 1.00 .

The 10 decks were presented twice in random order and the 20 sample cards were shuffled behind a screen before each presentation. Presentation rate was one

Table 1.

Proportion of optimal inferences for each of the 10 correlations.

Sign of the sample correlation	Magnitude of the sample correlation					Over-all + and - responses
	$\pm.14$	$\pm.15$	$\pm.30$	$\pm.50$	$\pm.85$	
Positive	.54	.59	.64	.77	.91	.473
Negative	.69	.62	.73	.78	.90	.527

Fig. 1.

card every 2-3 sec. with 5-6 sec. between decks for recording responses.

Subjects. Thirty-two University of Michigan students served as paid Ss.

Results

Evaluation of the optimality of Ss' inferences and confidence requires comparisons with Bayesian inferences and posterior probabilities. If Ss' behavior is similar to the Bayesian model they should always infer that the population correlation has the same sign as the sample correlation. Of course, Ss cannot compute the sample correlations and, as a result, the probability that they will make optimal inferences should increase as the sample correlations become more discernable, i.e., as the magnitudes increase. Table 1 contains the proportion of Ss who made optimal inferences about the sign of each correlation. Clearly, they did better as the magnitudes of the sample correlations increased, although there was a tendency to call all low correlations negative and thus there was a slightly higher chance of being correct when a low negative correlation occurred.

Figure 1 shows the increase in the median of Ss' confidence responses as sample correlations increased in magnitude. The smooth curve represents the Bayesian posterior probabilities corresponding to the coefficients. While Ss' confidence increases, it falls short of optimal confidence. This suggests that Ss' .50-.50 prior subjective probabilities were less modified by the sample data than the Bayesian model prescribed and, as a result, that Ss' posterior subjective probabilities were conservative compared to the Bayesian optimal probabilities. This result is congruent with those obtained in a variety of other tasks in which Ss modify their subjective probabilities in light of data; the modifications are in the direction prescribed by the optimal model but the magnitude is less than the prescribed amount (e.g., Peterson & Miller, 1965; Phillips, Hays, & Edwards, 1966).

In short, the optimality of Ss' inferences about population correlations and the degree of confidence they have in their inferences both increase as the magnitude of the sample correlations increase. However, both inferences and confidence are suboptimal when compared to the behavior prescribed by the optimal Bayesian statistical model.

References

- Lindley, D. V. *Introduction to probability and statistics from a Bayesian viewpoint: Part 2, inference*. Cambridge: Cambridge University Press, 1965.
- Peterson, C. R., & Miller, A. J. Sensitivity of subjective probability revision. *J. exp. Psychol.*, 1965, 70, 117-121.
- Phillips, L. D., Hays, W. L., & Edwards, W. Conservatism in complex probabilistic inference. *IEEE Trans. Hum. Factors Electronics*, 1966, 7, 7-18.

Notes

1. The research reported here was undertaken in the Engineering Psychology Laboratory, Institute of Science and Technology, The University of Michigan. The research was supported by USPHS Fellowship MF-12,744, USPHS Grant NIH-MH-11496 and by the Air Force Office of Scientific Research under Contract AF 49(638) 1731.
2. This same criticism is appropriate to probability learning studies.