

A Practical Way for Computing Approximate Lower and Upper Correlation Bounds

Hakan Demirtas & Donald Hedeker

To cite this article: Hakan Demirtas & Donald Hedeker (2011) A Practical Way for Computing Approximate Lower and Upper Correlation Bounds, The American Statistician, 65:2, 104-109, DOI: [10.1198/tast.2011.10090](https://doi.org/10.1198/tast.2011.10090)

To link to this article: <http://dx.doi.org/10.1198/tast.2011.10090>



Published online: 24 Jan 2012.



Submit your article to this journal [↗](#)



Article views: 240



View related articles [↗](#)



Citing articles: 11 View citing articles [↗](#)

A Practical Way for Computing Approximate Lower and Upper Correlation Bounds

Hakan DEMIRTAS and Donald HEDEKER

Correlations among variables are typically not free to vary between -1 and 1 , with bounds determined by the marginal distributions. Computing upper and lower limits of correlations given the marginal characteristics often raises theoretical and computational challenges. We propose a simple sorting technique that is predicated upon a little-known consequence of a well-established fact from statistical distribution theory to obtain approximate correlation bounds. This approach works regardless of the data type or distribution. We believe that it has practical value in appropriately specifying the correlation structure in simulation studies.

KEY WORDS: Correlation limits; Simulation; Sorting.

1. INTRODUCTION

It is well known that correlations are not bounded between -1 and $+1$ in most bivariate settings as different upper and/or lower bounds may be imposed by the marginal distributions (Hoeffding 1940; Frechet 1951). These restrictions apply to discrete variables as well as continuous ones. Although the Frechet–Hoeffding bounds were established years ago, the maximal correlation and anticorrelation aspect is not always recognized within the statistics community. In particular, we feel its utility has not been fully realized in the context of teaching statistics, and more broadly in simulation specifications. In our teaching of graduate-level courses involving computing at the University of Illinois at Chicago in the Division of Epidemiology and Biostatistics, the design and execution of simulation studies is an essential component of the coursework. The steps of model building, estimation, and testing typically require simulation to assess the validity, reliability, and plausibility of the inferential techniques, to evaluate how well the implemented statistical models capture the specified true population values, and how reasonably these models respond to departures from underlying assumptions, among other things. In such simulation studies, the specification of the correlation structure among

variables is inevitably required. Students often are unaware of or have difficulty in deriving the correlation bounds (which are subsequently needed in the simulation design) imposed by the marginal features of the variables. This motivates the current work. In Section 4, we propose a potentially useful and practical sorting idea to address this problem.

In our experience, students often specify the correlations using the broadest possible limits (i.e., -1 to $+1$) in designing simulations. It is not always straightforward to realize that some correlation specifications are simply infeasible. Failure to define the association parameters within a plausible range can easily result in suboptimal, inconsistent, and/or incorrect conclusions. Furthermore, it is a common tendency among simulation designers to try extreme situations to gauge the limits of the statistical procedures or models under consideration. For this reason, these extremes need to be carefully considered and defined. The purpose of this work is to suggest a very simple sorting approach that gives approximate lower and upper bounds for correlation. Again, the theoretical rationale of this technique, which we describe in Section 2, is well-established; however, its practical merits may not be fully appreciated in teaching and designing of simulation studies.

2. SOME THEORY

Let $\Pi(F, G)$ be the set of all cumulative distribution functions (cdf's) H on R^2 having marginal cdf's F and G . Hoeffding (1940) and Frechet (1951) proved that in $\Pi(F, G)$, there exist cdf's H_L and H_U , called the lower and upper bounds, having minimum and maximum correlation. For all $(x, y) \in R^2$, $H_L(x, y) = \max[F(x) + G(y) - 1, 0]$ and $H_U(x, y) = \min[F(x), G(y)]$. For any $H \in \Pi(F, G)$ and all $(x, y) \in R^2$, $H_L(x, y) \leq H(x, y) \leq H_U(x, y)$. If δ_L , δ_U , and δ denote the Pearson correlation coefficients for H_L , H_U , and H , respectively, then $\delta_L \leq \delta \leq \delta_U$. One can infer that if V is uniform in $[0, 1]$, then $F^{-1}(V)$ and $G^{-1}(V)$ are maximally correlated; and $F^{-1}(V)$ and $G^{-1}(1 - V)$ are maximally anticorrelated. This can be proven by the Wasserstein distance between two probability measures, which is a natural way of comparing the probability distributions of two variables (Dobrushin 1970). Heuristically, with a slight abuse of notation, the Wasserstein distance $W(F, G)$ is defined as $W(F, G) = \int |x - y|^2 dH(x, y) = E[(X - Y)^2] = E(X^2) + E(Y^2) - 2E(XY)$. Since X and Y have fixed marginals F and G , respectively, the quantity $E(X^2) + E(Y^2)$ remains constant for $H \in \Pi(F, G)$, so $W(F, G)$ is minimized when the covariance term, $E(XY)$, is maximized. The choice $(X, Y) = (F^{-1}(V), G^{-1}(V))$ does

Hakan Demirtas (E-mail: demirtas@uic.edu) is Associate Professor and Donald Hedeker is Professor of Biostatistics, Division of Epidemiology and Biostatistics (MC923), University of Illinois at Chicago, 1603 West Taylor Street, Chicago, IL 60612. The authors are grateful to Luc Devroye for helpful discussions, and to an associate editor and a referee for suggestions that considerably improved the article.

this as it ensures that X and Y covary in the same direction to the greatest possible extent. Similarly, replacing Y by $-Y$, we see that $E(XY)$ is minimized by taking $X = F^{-1}(V)$ and $G^{-1}(1 - V)$. Under this choice, the minimal covariation (and the maximal Wasserstein distance) between X and Y is reached; that is, X and Y covary in the opposite direction to the greatest possible extent. For a mathematically rigorous proof, see the book by Rachev and Ruschendorf (1998, chap. 3).

3. SOME CASES ENCOUNTERED IN TEACHING

Random number generation is often one of the core topics in computation-oriented courses. The specification of the correlation structure in multivariate settings needs to be done meticulously so that the prescribed correlations do not go beyond the feasible correlation range. If the computing environment alerts simulation designers of this problem (e.g., by warning or error messages), then remedial action to avoid the discrepancies between expectation and reality that arise from correlation range violations is straightforward. However, depending on the nature of the computing platforms, programs, linkers, compilers, debuggers, and/or idiosyncratic aspects of what is to be simulated, students and other users may not even realize that the specified values are infeasible. More importantly, correlation bounds are typically unavailable in closed form, hence incalculable in most cases.

Below, we provide a few scenarios for the bivariate case from our teaching experience. These consist of situations in which students are asked to generate bivariate data for different types of variables. Students sometimes recognize that there are correlation bounds, and here we indicate and illustrate these for the simpler cases. However, for more general situations, as we describe, closed-form bounds are not always obtainable and students typically have no idea about how to obtain the correlation bounds. Also, while here we focus on the bivariate case, for multivariate settings correct specification of the correlation matrices requires that the specified correlation matrix is positive semidefinite, in addition to the requirement that each pairwise correlation be within the corresponding bivariate correlation range. We further discuss the multivariate case in Section 4.2.

3.1 Binary-Normal Example

The first example considers dichotomization of one of the variables. Suppose that students are to generate a binary-normal combination with specified values of correlation and proportion for the dichotomized variable. Let X and Y follow a bivariate normal distribution with a correlation δ_{XY} . If X is dichotomized to produce X_D , then the resulting correlation between X_D and Y can be designated as $\delta_{X_D Y}$ (point-biserial correlation). The effect of dichotomization on δ_{XY} (biserial or tetrachoric correlation) is given by $\delta_{X_D Y} = \delta_{XY}(h/\sqrt{pq})$ where p and q are the proportions of X values above and below the point of dichotomization, respectively, and h is the ordinate of the normal curve at the same point (MacCallum et al. 2002). Here, it is well known that the correlation range for $\delta_{X_D Y}$ is $\mp h/\sqrt{pq}$. Lower and upper bounds with respect to different values of p are shown in Figure 1. Even in the least extreme

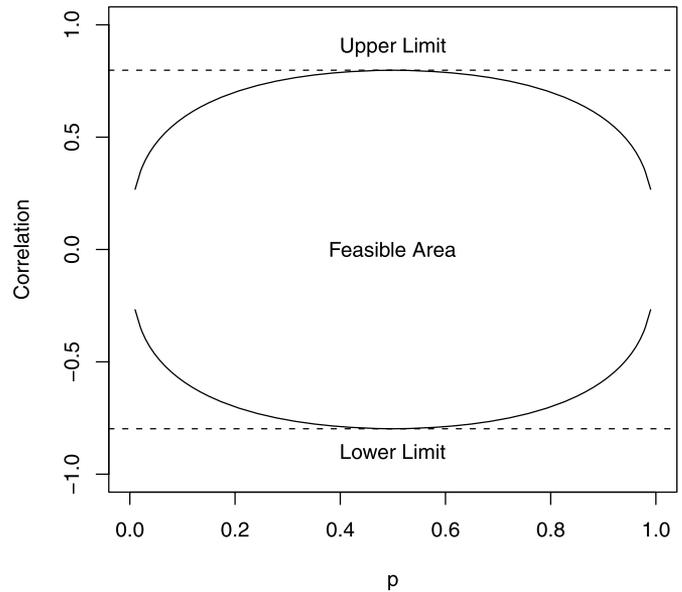


Figure 1. Upper and lower correlation bounds for the binary-normal example with respect to different values of p . The dotted lines represent the correlation bounds, $\mp\sqrt{\frac{2}{\pi}} = \mp 0.7979$, when $p = 0.5$.

case ($p = 0.5$), the correlation bounds are different from -1 and $+1$; and when p gets closer to 0 or 1, the feasible correlation range becomes narrower, as can be seen in Figure 1.

3.2 Bivariate Binary Example

The second example considers generation of correlated binary data. If students are given a task to simulate bivariate binary outcomes, Y_1 and Y_2 , such that $E[Y_j] = p_j$ where $j = 1, 2$ and $\text{corr}(Y_1, Y_2) = \delta_{12}$, then δ_{12} is bounded below by $\max[-(\frac{p_1 p_2}{(1-p_1)(1-p_2)})^{1/2}, -(\frac{(1-p_1)(1-p_2)}{p_1 p_2})^{1/2}]$ and above by $\min[(\frac{p_1(1-p_2)}{(1-p_1)p_2})^{1/2}, (\frac{(1-p_1)p_2}{p_1(1-p_2)})^{1/2}]$ (Emrich and Piedmonte 1991). Lower and upper bounds with respect to p_2 while fixing p_1 at 10 different values (from 0.05 to 0.95 with increments of 0.10) are shown in Figure 2. As $|p_1 - p_2|$ becomes larger, the upper bound moves away further from $+1$. When the absolute difference gets smaller, then the lower bound moves away further from -1 , as demonstrated in Figure 2. A comprehensive table of bounds in this setting is given by Demirtas (2006).

3.3 Bivariate Lognormal Example

Suppose Y_1 and Y_2 are jointly normally distributed with zero means, variances, σ_i^2 , $i = 1, 2$, and correlation δ_y . Let $X_i = \exp(Y_i)$ for $i = 1, 2$. Then, X_1 and X_2 are jointly lognormally distributed with correlation $\delta_x = \frac{e^{\delta_y \sigma_1 \sigma_2}}{\sqrt{e^{\sigma_1^2} - 1} \sqrt{e^{\sigma_2^2} - 1}}$. Bounds for δ_x can be obtained by substituting -1 and $+1$ for δ_y . Lower and upper bounds with respect to σ_2 while fixing σ_1 at nine different values (from 0.1 to 0.9 with increments of 0.1) are shown in Figure 3. The most striking trend is seen when σ_2 becomes large (e.g., 2.5): the feasible correlation range is merely a small subset of the nominal $(-1, +1)$ range.

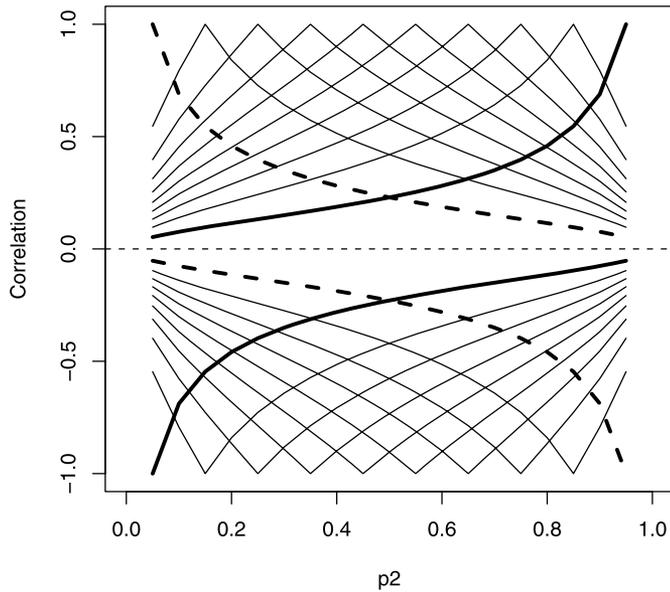


Figure 2. Upper and lower correlation bounds for the binary–binary example with respect to p_2 when p_1 is fixed at 10 different values ranging from 0.05 to 0.95 with increments of 0.1. Thick dashed and solid curves correspond to $p_1 = 0.05$ and 0.95 , respectively.

As the above three examples have illustrated, the correlation bounds can often be dramatically different from the usual $(-1, +1)$ range; and while, as indicated, there are established formulas for the bounds, students are not always aware of these. Alternatively, the approach we describe in Section 4, which students can readily apply, yields the admissible bounds in a quick and simple manner.

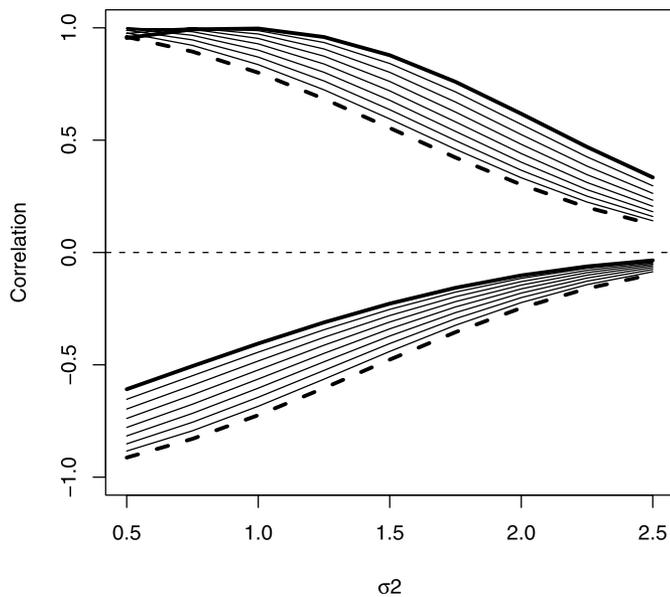


Figure 3. Upper and lower correlation bounds for the bivariate lognormal example with respect to σ_2 when σ_1 is fixed at nine different values ranging from 0.1 to 0.9 with increments of 0.1. Thick dashed and solid curves correspond to $\sigma_1 = 0.1$ and 0.9 , respectively.

3.4 Possible Consequences of Violating Correlation Bounds: Bivariate Lognormal Case

Here, we illustrate some consequences of violating the correlation bounds for the lognormal case. To do so, we employ a rather general data generation technique that is often utilized in simulating multivariate continuous data, namely, the use of Fleishman polynomials (Fleishman 1978; Vale and Maurelli 1983; Demirtas and Hedeker 2008; Headrick 2010). This approach is a moment-matching procedure where any given continuous variable in the system (i.e., set of variables that are to be generated) is expressed by the sum of linear combinations of powers of a standard normal variate. It involves solving a set of nonlinear equations via an optimization routine such as Newton–Raphson to compute the coefficients of polynomials marginally, and calculating intermediate correlations among normal variables that form a basis for desired nonnormal continuous distributions to attain the specified correlation structure (Headrick 2010).

Considering the lognormal case, one reason that we have for resorting to the Fleishman polynomials to generate such data is to elucidate a major point: in one specification ($\sigma_1 = 0.6$, $\sigma_2 = 0.3$), students would get a warning message and “unavailable” results for a subset of the specified correlation values due to range violation (see Figure 4), whereas another specification ($\sigma_1 = 0.65$, $\sigma_2 = 0.65$) can lead to false correlations (see Figure 5). [These comments pertain to programs written in R version 2.12.0 for Windows (R Development Core Team 2010), and codes for implementing multivariate Fleishman polynomials. Other computing environments (programs and/or operating systems) or applications may give rise to different consequences.] Here, we specify a correlation value, generate data,

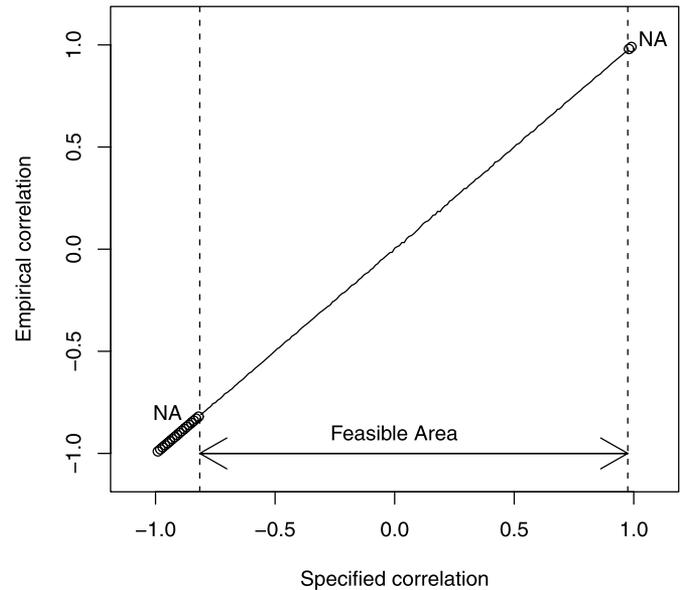


Figure 4. Plot of the specified (true) versus empirical (computed) correlations for bivariate lognormal data with $\sigma_1 = 0.6$ and $\sigma_2 = 0.3$ generated by Fleishman polynomials. Empirical correlations cannot be computed when the specified correlation is below -0.81 and above 0.97 (i.e., the computer program gives a warning message).

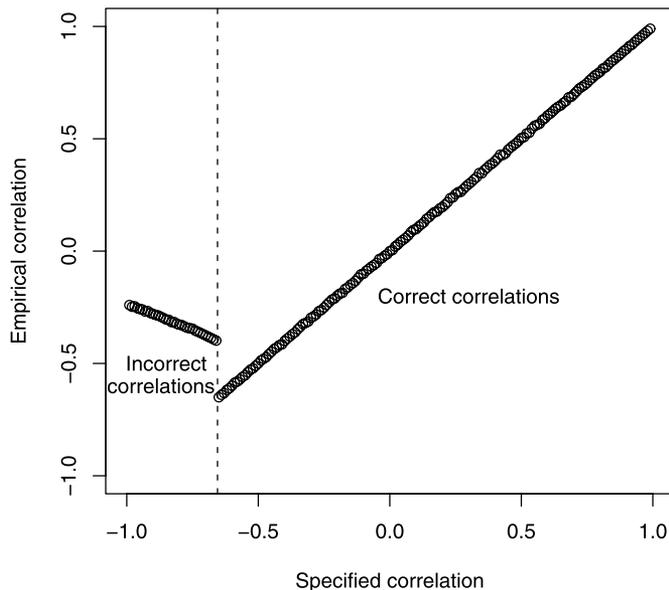


Figure 5. Plot of the specified (true) versus empirical (computed) correlations for bivariate lognormal data with $\sigma_1 = \sigma_2 = 0.65$ generated by Fleishman polynomials. Empirical correlations are computed incorrectly when the specified correlation is below -0.65 , (i.e., the computer program does not give a warning message).

and compute sample (empirical) correlations using the simulated data. If a procedure is working well, the specified (true) and empirical (computed) correlations should be close to each other. Specifically, in Figure 4, we plot specified versus empirical correlations after generating data using the first set of parameter values above, and see that correlations are not computable below and above the two thresholds with a “not available” message. That is, in these ranges, students would get an error message and data would not be generated. This is a lucky case, because students become aware of the misspecification problem. To demonstrate how bad the situation can become, we generate a similar plot (Figure 5) with the second set of parameter values, and see that the empirical correlations turn out to be badly misleading below a certain desired correlation value. For instance, when the true value is -0.8 , the empirical value is way off (-0.43). As can be seen in Figure 5, when the true correlation is below -0.65 , the sample correlation of the generated data is far from the truth by any acceptable standards. Here, the computer program gives no indication that there is anything wrong, hence one can easily proceed with implausibly generated data with false correlations under this second set of specifications. Considering the fact that random number generation is only the first step of a much bigger simulation scheme in most cases, there is a serious potential for obtaining incorrect results from the subsequent stages. In Section 4.1, we provide further explanation of the problematic behavior in Figures 4 and 5.

3.5 More Complicated Cases

For the above scenarios, the correlation bounds are derivable. However, this is not always the case. What if the correlation bounds are not derivable in closed form (which is certainly plausible for many real-life applications)? In these situations, how can one obtain the range of correlation values

that are within the distributional limits for simulation purposes? As an example, suppose that there are two variables where one follows a normal mixture distribution, $\pi N(\mu_1, \sigma_1^2) + (1 - \pi)N(\mu_2, \sigma_2^2)$ with the mixing proportion π , and the other follows a generalized lambda distribution (Ramberg et al. 1979) whose quantile function is $\lambda_1 + [p^{\lambda_3} - (1 - p)^{\lambda_4}]/\lambda_2$ where $0 \leq p \leq 1$. In this particular case, the boundary values for correlation cannot be expressed analytically.

In the next section, we outline a flexible, easy-to-implement sorting algorithm to compute the empirical correlation bounds among virtually any type of variables that can be encountered in practice. We show that the empirical bounds found by this technique agree closely to the theoretical correlation bounds when these are known. We further elucidate some of the advantages and extensions of our approach in more general situations.

4. GENERATE, SORT, AND CORRELATE (GSC)

The GSC algorithm for the bivariate case is as follows:

1. Generate random samples from the intended distributions independently using a large number of observations (e.g., $N = 100,000$).
2. To estimate the lower bound, sort the two variables in opposite directions (i.e., from smallest to largest for one of the variables, and from largest to smallest for the other). Then, compute the sample correlation.
3. To estimate the upper bound, sort the two variables in the same direction, and compute the sample correlation.

Steps 2 and 3 correspond to maximal anticorrelation and correlation, respectively. In effect, this is equivalent to the inverse cdf method. If a random number generation routine is available to simulate samples from the marginal distributions of interest, then one can generate random realizations, and sort them in these two different ways, as described above, to compute the approximate correlation limits.

4.1 Agreement With Theoretical Correlation Bounds

Following the examples given in Section 3, we now illustrate the numerical efficiency of this approach with $N = 100,000$ observations. For the binary-normal case, under a standard bivariate normal distribution assumption, let one of the variables be dichotomized at the mean. In this situation, the correlation bounds are known to be $\mp\sqrt{\frac{2}{\pi}} = \mp 0.7979$, which is exactly what the GSC approach yields on average across 1000 replications, where the empirical, central 95% confidence interval is $(\mp 0.7967, \mp 0.7991)$. In what follows, we report these intervals in parentheses. For the binary-binary case, if $p_1 = 0.5$ and $p_2 = 0.6$, the theoretical upper and lower limits are 0.8165 and -0.8165 , respectively. Using the GSC approach, we obtain limits of 0.8164 (0.8107, 0.8220) and -0.8165 ($-0.8222, -0.8108$). For the lognormal case, Greenland (1996) described two scenarios where $\sigma_1^2 = \sigma_2^2 = 0.5$ or 3 with corresponding lower limits of -0.6065 and -0.0498 , respectively (upper limits are 1). Employing the GSC approach by generating two lognormally distributed variables independently, then reverse-sorting and correlating these

variables yields values of -0.6066 ($-0.6128, -0.6001$) and -0.0499 ($-0.0656, -0.0376$), respectively, which are near-identical to the theoretical lower bounds on average. Depending on the underlying distributional shapes, some variation is naturally expected. However, in simulation practice, people vary correlation specifications with nice incremental numbers, and slight deviations from the true bounds are unlikely to cause misspecifications that arise from bound violations; one can get closer to the true bounds with a little extra effort (i.e., by running the proposed procedure sufficiently many times and then using the average), which would not be computationally expensive as it will be done only once before specifying a plausible correlation range in simulations.

Now, returning to the two interesting specifications described in Section 3.4, that could lead to bad consequences. When $\sigma_1 = 0.6$ and $\sigma_2 = 0.3$, the empirical (based on the GSC algorithm) and theoretical (based on the formula in Section 3.3) approaches both yield the lower and upper bounds of -0.8154 and 0.9763 , respectively. This is consistent with the vertical lines that indicate the feasible range in Figure 4. Similarly, when $\sigma_1 = \sigma_2 = 0.65$, the lower bound is -0.6554 , below which, as noted, incorrect empirical correlations would be obtained (Figure 5). Thus, using the GSC approach, students would realize that specification of a correlation below -0.6554 is not possible (and so would avoid generating data with “incorrect” correlations).

4.2 Advantages and Extensions

This method has a few key advantages. First of all, it works for many kinds of variables (e.g., continuous, binary, ordinal, count). While there may be a better measure of association depending on the types of variables, it is still good to know the limits of the Pearson correlation. Second, the variables need not come from the same distribution. As illustrated above, we obtained the correlation between a continuous and a binary variable, assuming underlying normality for both, but the two distributions could be completely different. Third, in some cases, a mathematical solution for the correlation bounds may be very difficult to obtain, or may not even exist in closed form. In such cases, the proposed approach is very useful by providing results for these intractable situations. Finally, these boundary arguments extend to the multivariate case. The overall correlation bound matrices can be formed using the bivariate correlation bounds. In this situation, one must realize that ensuring all bivariate correlations are within the feasible limits does not guarantee positive semidefiniteness, so one must still pay attention to this.

To further elaborate and illustrate the above-mentioned advantages, we consider a multivariate setting consisting of five variables that have different distributional properties. Specifically, we assume $X_1 \sim \text{Bernoulli}(0.8)$; $X_2 \sim \text{Poisson}(1)$; X_3 is ordinal, $P(X_3 = i) = p_i$, where $p_i = (0.1, 0.4, 0.1, 0.4)$ for $i = 1, 2, 3, 4$; $X_4 \sim \pi N(\mu_1, \sigma_1^2) + (1 - \pi)N(\mu_2, \sigma_2^2)$ (normal mixture) where $\pi = 0.2$, $\mu_1 = 0$, $\mu_2 = 4$, $\sigma_1^2 = 3$, $\sigma_2^2 = 8$; and X_5 follows a generalized lambda distribution where $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0, 1, 1, 10)$ (see Section 3.5). Admittedly, this is a rather unusual situation, but helps to illustrate the generality of our approach. Approximate correlation limits are shown

Table 1. Lower (below the diagonal) and upper (above the diagonal) correlation bounds.

	X_1	X_2	X_3	X_4	X_5
X_1	–	0.5000	0.6035	0.3765	0.8134
X_2	–0.7592	–	0.7759	0.8855	0.7906
X_3	–0.5571	–0.8721	–	0.6745	0.8860
X_4	–0.9441	–0.6311	–0.7414	–	0.6909
X_5	–0.5511	–0.9562	–0.8311	–0.9421	–

in Table 1, with upper- and lower-diagonal elements representing upper and lower correlation bounds, respectively. It is interesting to observe that these bounds are often quite a bit less, in absolute value, than the bounds of -1 and 1 . Note that both upper- and lower-diagonal parts of Table 1 should not be regarded as legitimate upper and lower parts of a correlation matrix; Table 1 is constructed merely for demonstration of correlation bounds for each pair of variables. Of course, any specified correlation matrix must also be checked for positive semidefiniteness.

To illustrate how these bounds might be used in a multivariate setting, consider the following. What often happens in simulation practice is that one specifies equally spaced correlation values between -1 and $+1$ for each pair. However, when the limits are different from -1 and $+1$ (e.g., those given in Table 1), this naive approach will inevitably yield problems. Instead, if one uses the derived bounds from the GSC method, then the specified range could be defined accordingly to avoid such range violations. We hasten to add that closed-form solutions do not exist for many pairwise combinations (e.g., those in Table 1). Furthermore, this method is only designed to produce the approximate correlation bounds that should not be exceeded; availability of bivariate or multivariate data generation routines for a particular setting is a separate issue. Finally, as correctly pointed out by a reviewer, even when positive semidefiniteness holds and the correlation range is not violated, some correlation structures may not be achievable due to operational and/or procedural problems that are specific to the random number generation method in use [e.g., the tetrachoric correlation matrix may not be positive semidefinite even when there is no misspecification in the correlated binary data generation technique of Emrich and Piedmonte (1991)].

Another area of application of this approach is in situations where there are insufficient data available to accurately compute a correlation matrix. In such cases, the correlations are frequently derived from expert opinion, usually the combined opinions of many experts (Lurie and Goldberg 1998). With fixed marginals, it is important to check whether the expert-derived correlations violate the range constraints.

Finally, as mentioned, in designing simulation studies, researchers typically identify a range for model parameters including correlations. In doing so, researchers often simulate data under relatively extreme cases. A correlation range that includes impossible values may lead to inferential biases, especially if the computational procedure that is being assessed does not give error or warning messages to alert the simulation designer. Clearly, in this situation, checking the correlation bounds before simulating the data using our simple approach could prevent problems and/or subsequent misinterpretations regarding the computational procedure.

5. CONCLUSION

In this article, we have described a simple approach that can be used to obtain approximate correlation bounds under a wide variety of situations. In our teaching experience in computational statistics, this approach can help students from making unintended errors when devising simulations involving specification of a correlation matrix. As long as one is capable of generating the univariate marginals, it is straightforward to perform the sorting operation to identify the correlation bounds regardless of data type or distribution. We believe that this simple technique is a minor but valuable addition to the teachers' toolkit. Additionally, this approach is also potentially useful to statisticians who are evaluating models or methods via simulation.

[Received May 2010. Revised April 2011.]

REFERENCES

- Demirtas H. (2006), "A Method for Multivariate Ordinal Data Generation Given Marginal Distributions and Correlations," *Journal of Statistical Computation and Simulation*, 76, 1017–1025. [105]
- Demirtas, H., and Hedeker, D. (2008), "Multiple Imputation Under Power Polynomials," *Communications in Statistics—Simulation and Computation*, 37, 1682–1695. [106]
- Dobrushin, R. L. (1970), "Prescribing a System of Random Variables by Conditional Distributions," *Theory of Probability and Its Applications*, 15, 458–486. [104]
- Emrich, L. J., and Piedmonte, M. R. (1991), "A Method for Generating High-Dimensional Multivariate Binary Variates," *The American Statistician*, 45, 302–304. [105,108]
- Fleishman, A. I. (1978), "A Method for Simulating Non-Normal Distributions," *Psychometrika*, 43, 521–532. [106]
- Frechet, M. (1951), "Sur les tableaux de corrélation dont les marges sont données," *Annales de l'Université de Lyon Section A*, 14, 53–77. [104]
- Greenland, S. (1996), "A Lower Bound for the Correlation of Exponentiated Bivariate Normal Pairs," *The American Statistician*, 50, 163–164. [107]
- Headrick, T. C. (2010), *Power Method Polynomials and Other Transformations*, Boca Raton, FL: Chapman & Hall/CRC. [106]
- Hoeffding, W. (1940), "Scale-Invariant Correlation Theory," in *The Collected Works of Wassily Hoeffding*, eds. N. I. Fisher and P. K. Sen, New York: Springer-Verlag, pp. 57–107. [104]
- Lurie, P. M., and Goldberg, M. S. (1998), "An Approximate Method for Sampling Correlated Random Variables From Partially-Specified Distributions," *Management Science*, 44, 203–218. [108]
- MacCallum, R. C., Zhang, S., Preacher, K. J., and Rucker, D. D. (2002), "On the Practice of Dichotomization of Quantitative Variables," *Psychological Methods*, 7, 19–40. [105]
- R Development Core Team (2010), *R: A Language and Environment for Statistical Computing*, Vienna, Austria: Foundation for Statistical Computing. [106]
- Rachev, S. T., and Ruschendorf, L. (1998), *Mass Transportation Problems, Volume I: Theory*, New York: Springer-Verlag. [105]
- Ramberg, J. S., Dudewicz, E. J., Tadikamalla, P. R., and Mykytka, E. F. (1979), "A Probability Distribution and Its Uses in Fitting Data," *Technometrics*, 21, 201–214. [107]
- Vale, C. D., and Maurelli, V. A. (1983), "Simulating Multivariate Nonnormal Distributions," *Psychometrika*, 48, 465–471. [106]