TECHNOMETRICS, FEBRUARY 1995, VOL. 37, NO. 1

# Estimating the Correlation Between Destructively Measured Variables Using Proof-loading

**S. H. Steiner**

Department of Statistics
and Actuarial Sciences
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

**G. O. Wesolowsky**

Faculty of Business
McMaster University
Hamilton, Ontario L8S 4M4
Canada

Two simple proof-load testing procedures are suggested to estimate the correlation between two variables that can individually only be determined destructively. The first procedure assumes that the means and variances of the variables are known. The second testing procedure is more flexible and requires no prior information. By assuming that the variables have a bivariate normal distribution and considering only the number of units that fail at each proof-load, we determine the maximum likelihood estimate for the correlation coefficient. Theoretical and simulation results compare favorably with previously suggested more complicated procedures and guide the practitioner in appropriate choices for the proof-load levels.

KEY WORDS: Attribute testing; Censored data; Destructive testing; Jackknife methods; Maximum likelihood estimate (MLE).

Many materials used in construction and other applications can be characterized by two or more important physical strength properties. In assessing the acceptability of the materials, the correlation between the various strength properties can be very important. For physical structures subject to a variety of stresses, large correlations between strength modes have the effect of increasing the variability of a structure's load-carrying capacity, thus making it less reliable. Suddarth, Woeste, and Galligan (1978) and Galligan, Johnson, and Taylor (1979) studied the effect of the degree of correlation between bending and tensile strength in metal-plate wood trusses used in the roof structure of most homes. They concluded, based on theoretical and simulated results, that a large correlation may significantly affect the structure's reliability.

In many applications, however, the strength of an item can only be determined through destructive testing. Lumber, for example, has several physical properties—such as bending strength, tensile strength, shear strength, and compression strength—that can only be determined destructively. As a result, one is able to ascertain only the breaking strength in a single mode for each unit. In such situations, the correlations among the various strength properties cannot be measured directly and must be approximated. Several past studies (Amorim 1982; Amorim and Johnson 1986; Evans, Johnson, and Green 1984; Galligan et al. 1979; Green, Evans, and Johnson 1984; Johnson and Galligan 1983) have addressed the problem of estimating the correlation between two destructively determined variables by using proof-loading. Proof-loading involves stressing units only up to a prescribed (proof) load, thereby breaking only the weaker members of a population (Johnson 1980). This way, although some units break before the proof-load is reached, others survive and can be subjected to further testing in other strength modes.

The strategy employed in past studies (Amorim 1982; Evans et al. 1984) involves proof-loading units on the first mode followed by stressing the survivors until failure on a second mode and recording the exact load at failure. By assuming that the strength properties have a bivariate normal distribution with known means and standard deviations, Evans et al. (1984) and Amorim (1982) were able to solve numerically for the maximum likelihood estimate (MLE) of the correlation for various sample sizes and actual correlation values. A simulation study evaluated the mean and standard error of the MLE at different proof-load levels. Determining that the MLE was approximately unbiased, they compared the standard error of the MLE with the theoretical lower bound given by evaluating the reciprocal of the Fisher information.

Other researchers have considered extensions of the correlation estimation problem that use additional information from nondestructively measured properties. Bartlett and Lwin (1984) considered a variation in which a third property can be measured nondestructively. Johnson and Galligan (1983) and Galligan et al. (1979) estimated the correlation between two destructively measured properties in which each is a function of several properties that can be measured nondestructively. They presented results comparing the correlation estimate calculated ignoring the additional dependence on the nondestructively measured properties and estimates obtained using the additional information. The procedure was performed on real data, but the results were inconclusive due to a poor choice of proof-load levels. These methods, although theoretically useful, require much prior information that is often not available and so have not been successfully applied to real data.

Note that all procedures based on proof-loading implicitly assume that survivors of the proof-load are not damaged. Experimental studies by Madsen (1976) and Strickler, Pellerin, and Talbot (1970) suggested that this may be a reasonable assumption regarding the static strength of lumber, although a few pieces whose strength is only slightly greater than the proof-load stress will likely be weakened. In addition, according to cumulative damage theory (Gerhards 1979, p. 139), "the theoretical results suggest that some percentage of the population will fail during the proof-load, a very small additional percentage will be weakened, but the remainder will have residual strength virtually equal to original strength." These theoretical results are based on the reasonable assumption that the proof-loading is done at a rapid rate.

This article proposes two simple procedures that use only proof-loading to estimate the correlation between two variables that can individually only be measured destructively. In Section 1, we describe Procedure I, which is very simple but requires prior knowledge of the individual means and variances. In Section 2, we present Procedure II, which requires a slightly more complicated testing procedure but is more flexible because it does not require any prior information. In Section 3, we present an example of the use of the two procedures. In Section 4,we compare the efficiency of the proposed procedures with past approaches.

## 1. PROCEDURE I: ONE-WAY ESTIMATION

In developing Procedure I it is assumed that the two variables, denoted $A$ and $B$, have a bivariate normal distribution with known means and standard deviations $\mu_a$, $\mu_b$, $\sigma_a$, $\sigma_b$ and that survivors of proof-loads are not significantly damaged. To simplify the testing required, however, rather than recording the precise load at failure for each unit in the sample as in previous studies, we record only the number of units that fail each of the two proof-loads. As a result, because proof-loading generally does not require sophisticated measuring equipment and can be done quickly and easily, the proposed procedure would be cheaper and easier to apply in practice than previous procedures. Moreover, recording only either pass or fail is very natural for destructive strength tests because it is often difficult to measure breaking strength precisely. Procedure I is performed as follows:

1. Start with a sample of size $n$. Load each unit up to a proof-load of $\mathrm{PL}_a$ in variable $A$, letting $p_a$ equal the probability of failure on this load, that is, $p_a = \Phi((\mathrm{PL}_a - \mu_a)/\sigma_a)$, where $\Phi$ is the cumulative standard normal probability. Denote the number of units that break under this first proof-load as $n_a$.
2. Subject the remaining $n - n_a$ units to a proof-load of $\mathrm{PL}_a$ on mode $B$, where $p_a = \Phi((\mathrm{PL}_b - \mu_b)/\sigma_b)$ equals the probability of failure. Let $n_b$ equal the number of units that fail this second proof-load.

Note that $n - n_a - n_b$ units fail neither proof-load. Based on this testing procedure the likelihood of obtaining any given $n_a$ and $n_b$ is

$$\begin{aligned} L(&n, n_a, n_b) \\ &= p_a^{n_a}\big((1 - p_a)p_{b|\bar{a}}\big)^{n_b}\big((1 - p_a)(1 - p_{b|\bar{a}})\big)^{n-n_a-n_b} \\ &= p_a^{n_a}(p_b - p_{a\cap b})^{n_b}(1 - p_a - p_b + p_{a\cap b})^{n-n_a-n_b}, \end{aligned} \tag{1}$$

where $p_{b|\bar{a}}$ equals the probability of failure on mode $B$ given that the unit did not fail on mode $A$ and $p_{a\cap b}$ equals the theoretical probability that a unit would fail both proof-loads. Solving for $p^*_{a\cap b}$ the MLE of $p_{a\cap b}$, gives

$$p^*_{a\cap b} = p_b - \frac{n_b(1 - p_a)}{n - n_a}. \tag{2}$$

Using Equation (2) it is possible, although unlikely unless proof-load levels are very small, that the estimate $p^*_{a\cap b}$ is negative. This makes no physical sense, and we recommend that, if $p^*_{a\cap b} < 0$, $p^*_{a\cap b}$ is set equal to 0. Using this truncation, $p^*_{a\cap b}$ is technically no longer the MLE. As a result, subsequent analysis ignores this recommended truncation and is only valid when such truncation is unlikely. In any event, the practical effect of the truncation is to reduce the standard error of the estimate.

Because $p_{a\cap b}$ is a function of $p_a$, $p_b$, and $\rho_{ab}$, we can solve for the MLE of the correlation $\rho^*_{ab}$ that corresponds to the given values of $p_a$, $p_b$, and the MLE $p^*_{a\cap b}$. Define $g(x, y, \rho)$ as the standard bivariate normal probability function given by formula 26.3.1 of Abramowitz and Stegun (1972), and denote $h$ and $k$ as percentiles of a standard normal distribution such that $h = \Phi^{-1}(1 - p_a)$, $k = \Phi^{-1}(1 - p_b)$. Then

$$\begin{aligned} p_{a\cap b} &= \int_{y=-\infty}^{k}\int_{x=-\infty}^{h} g(x, y, \rho_{ab})\,dx\,dy \\ &= f(p_a, p_b, \rho_{ab}) \end{aligned} \tag{3a}$$

and

$$p_{a\cap b} = \frac{1}{2\pi}\int_{\cos^{-1}\rho_{ab}}^{\pi} \exp\Big(-\frac{h^2 + k^2 - 2hk\cos z}{2\sin^2 z}\Big)dz, \tag{3b}$$

where (3b) is due to Sheppard (1900). Based on (3b), the partial derivative of $f(p_a, p_b, \rho_{ab})$ with respect to $\rho_{ab}$ is given as (Drezner and Wesolowsky 1990)

$$\frac{\partial f(p_a, p_b, \rho_{ab})}{\partial \rho_{ab}} = g(h, k, \rho_{ab}). \tag{4}$$

This expression is strictly positive; therefore, $f(p_a, p_b, \rho_{ab})$ is a strictly increasing function as $\rho_{ab}$ ranges from $-1$ to $+1$. Thus we can search for the MLE $\rho^*_{ab}$ value that corresponds to the determined untruncated $p^*_{a\cap b}$ value using the method of bisection. Note that for any given values of $p_a$, $p_b$, and $\rho_{ab}$ a simple numerical procedure given by Drezner and Wesolowsky (1990) will determine the bivariate normal integral (3b). Alternatively we may use Expression (4) and Newton–Raphson (Press, Flannery, Teukolsky, and Vetterling 1988) to solve for $\rho_{ab}$.

In summary, for Procedure I, the MLE $\rho^*_{ab}$ is determined as follows:

1. Choose the sample size $n$ and failure probabilities $p_a$ and $p_b$. Determine $h = \Phi^{-1}(1-p_a), k = \Phi^{-1}(1-p_b)$.
2. Perform the experiment as outlined to obtain $n_a$ and $n_b$.
3. Use Expression (2) to obtain the MLE $p^*_{a\cap b}$. If $p^*_{a\cap b} < 0$, set $p^*_{a\cap b} = 0$.
4. Using the calculated $p^*_{a\cap b}$, $h$, and $k$ find the value of $\rho^*_{ab}$ that solves (3).

The preceding methodology is a two-step process; first we obtain $p^*_{ab}$, which is then translated to the corresponding $\rho^*_{ab}$ value. It is of interest to note that $\rho^*_{ab}$ obtained through this two-step process is the same as that obtained through direct maximization of (1) written in terms of $\rho_{ab}$. The two-step methodology is recommended because it gives a simple explicit formula for $p^*_{a\cap b}$, which may also be of interest, and the methodology is more easily extended.

### 1.1 Properties of the Estimates $p^*_{a\cap b}$ and $\rho^*_{ab}$ for Procedure I

The MLE $p^*_{a\cap b}$, as derived in Equation (2), is defined only if $n_a < n$—in other words, if not all units fail under the first proof-load. If $n_a = n$, none of the units are subjected to any proof-loads in the second strength mode, and as a result, no information is obtained about the correlation between strength modes. We shall proceed with the analysis assuming that $n_a < n$. Fortunately, because the proof-load levels can be set to any desired values, observing $n_a = n$ is very unlikely even for small sample sizes.

The mean and variance of $p^*_{a\cap b}$ can be derived through conditioning. Assuming $n_a < n$, we have

$$E\left(\frac{n_b}{n-n_a}\right) = E\left(\frac{E(n_b \mid n_a)}{n-n_a}\right) = E(p_{b\mid\bar{a}}) = p_{b\mid\bar{a}}.$$

Thus, by Equation (2), $E(p^*_{a\cap b}) = p_{a\cap b}$. Similarly the conditional variance formula gives

$$\begin{aligned}\operatorname{var}(p^*_{a\cap b}) &= (1-p_a)^2 \operatorname{var}\left(\frac{n_b}{n-n_a}\right)\\ &= (1-p_a)^2\left[\operatorname{var}\left\{E\left(\frac{n_b}{n-n_a}\Big| n_a\right)\right\}\right.\\ &\qquad \left.+ E\left\{\operatorname{var}\left(\frac{n_b}{n-n_a}\Big| n_a\right)\right\}\right]\\ &= (p_b - p_{a\cap b})(1-p_a-p_b+p_{a\cap b})\\ &\qquad \times E\left(\frac{1}{n-n_a}\right).\end{aligned}$$

Expressions of the form $E(x^{-1})$, where $x$ is a binomial variate bounded away from 0, have been studied (Johnson and Kotz 1969, sec. 3.10). Given that $n_a > n$, the variable $n - n_a$ is such a binomial variate with sample size $n$ and probability of success $1 - p_a$. An approximation suggested by Grab and Savage (1954)—namely, $E((n-n_a)^{-1}) \cong (n(1-p_a)-p_a)^{-1}$—gives two significant figures of accuracy for $n(1-p_a) > 10$ and is more than adequate for our application. Thus we have

$$\operatorname{var}(p^*_{a\cap b}) \cong \frac{(p_b - p_{a\cap b})(1-p_a-p_b+p_{a\cap b})}{n(1-p_a)-p_a}. \tag{5}$$

The mean and variance of $\rho^*_{ab}$ are more difficult to determine. An estimate $p^*_{a\cap b}$ is translated to $\rho^*_{ab}$ through the relation $p_{a\cap b} = f(\rho_{ab}; p_a, p_b)$. The function $f(\rho_{ab})$ is not a simple linear relation, however. Thus, unfortunately, the MLE $\rho^*_{ab}$ is not in general unbiased. In the next section, however, we show, using simulation studies, that the bias of $\rho^*_{ab}$ is approximately 0 and an insignificant part of the estimate's mean squared error for most proof-load levels. Fortunately, the standard error of $\rho^*_{ab}$ can be estimated using the $\delta$ method [method of statistical differentials (Johnson and Kotz 1989)]. For $\rho_{ab}$ values away from the extremes $\pm 1$, $f(\rho_{ab})$ can be closely approximated by a linear function, and $\operatorname{var}(\rho^*_{ab}) \cong (\partial f/\partial \rho_{ab})^{-2}\operatorname{var}(p^*_{a\cap b})$. Thus, through Equations (4) and (5), the standard error of $\rho^*_{ab}$ can be approximated as

$$\begin{aligned}\operatorname{se}(\rho^*_{ab}) &= \frac{\sqrt{\operatorname{var}(p^*_{a\cap b})}}{\partial f/\partial \rho_{ab}} \cong \frac{1}{g(h,k,\rho_{ab})}\\ &\qquad \times \sqrt{\frac{(p_b - p_{a\cap b})(1-p_a-p_b+p_{a\cap b})}{n(1-p_a)-p_a}}.\end{aligned} \tag{6}$$

Notice that Equation (6) cannot be computed unless the true correlation level $\rho_{ab}$ is known. Our experience has shown, however, that using the computed MLE's $\rho^*_{ab}$ and $p^*_{a\cap b}$ in Equation (6) provides a good estimate of $\operatorname{se}(\rho^*_{ab})$ based solely on the sample data.

### 1.2 Sensitivity Analysis for Procedure I

The following section explores how proof-load levels, actual correlation values, and sample size affect the bias and standard error of $\rho^*_{ab}$. The sensitivity of the estimate $\rho^*_{ab}$ to changes in proof-load failure rates $p_a$ and $p_b$ is shown in three-dimensional surface plots. Figure 1 and other simulations suggest that correlation estimates obtained with Procedure I are unbiased for a large range of $p_a$ and $p_b$ values; only for extreme combinations in which one proof-load is high and the other low does the bias deviate significantly from 0. Figure 2 suggests that

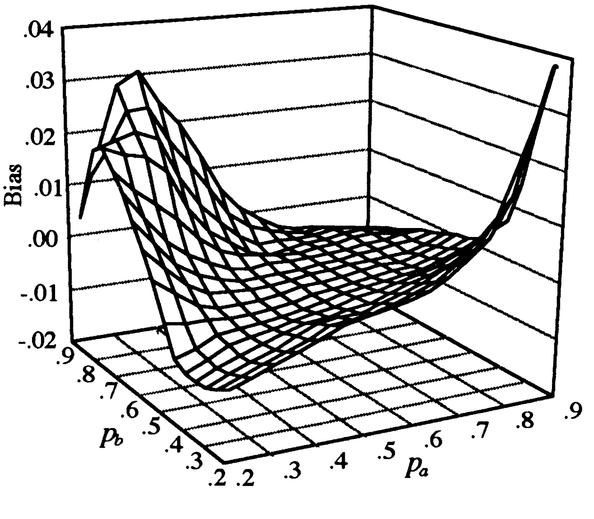


*Figure 1. Bias of Simulated Correlation Estimate: $n = 300$, $\rho_{ab} = .6$, 10,000 Samples.*

the standard error of $\rho^*_{ab}$ is relatively insensitive to changes in proof-load levels near the optimal values for $p_a$ and $p_b$. This is shown by the large flat section near the minimum. In general, simulation results of the estimate's standard error correspond very closely to Equation (6). Simulation results also suggest that $\rho^*_{ab}$ is approximately normally distributed so long as the sample size is fairly large and the estimate is not close to the extremes $-1$ and 1. In general, the normal approximation is good if the range $\rho^*_{ab} \pm 3se(\rho^*_{ab})$ does not include $-1$ or 1.

The effect of the sample size $n$ on the standard error of $\rho^*_{ab}$ is clearly demonstrated through Equation (6). The standard error of $\rho^*_{ab}$ decreases as a function of $1/\sqrt{n}$ as $n$ increases. The effect of $n$ on the bias of $\rho^*_{ab}$ is more difficult to quantify. Through a simulation study, however, we determined that the absolute value of the bias of $\rho^*_{ab}$ decreases approximately as a function of $1/n$ as $n$ increases. In any case, for near optimal proof-load levels the bias of $\rho^*_{ab}$ is insignificant compared with its standard error for any sample size.

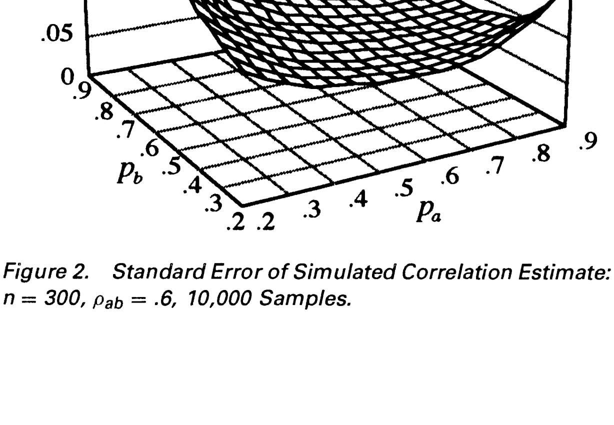


*Figure 2. Standard Error of Simulated Correlation Estimate: $n = 300$, $\rho_{ab} = .6$, 10,000 Samples.*

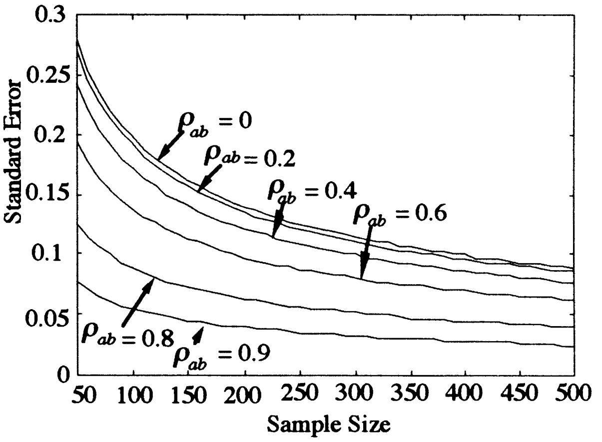


*Figure 3. Sample Size Versus Standard Error of $\rho^*_{ab}$: Assume Optimal Proof-load Levels From Section 1.3.*

Another factor that has a significant influence is the actual correlation value. Figure 3 explores, using (6), how changes in sample size and the actual $\rho_{ab}$ affect the standard error of $\rho^*_{ab}$. Note that the curves for negative correlations correspond almost exactly to the curves for positive correlations with the same absolute value. Figure 3 can guide the practitioner in choosing an appropriate sample size. Clearly, the estimation procedure works best when the real correlation is strongly positive or negative. Simulation studies suggest that the actual correlation value has little affect on the bias of $\rho^*_{ab}$ for near-optimal proof-load levels.

The effect of incorrect estimates for $p_a$ and $p_b$ on the correlation estimate is explored in Figure 4. For example, say we believe that our current proof-load level will result in $p_b = .4$ but in reality $p_b = .5$; that is, $\Delta p_b = -.1$. Figure 4 shows that this incorrect assumption leads to correlation estimates that are on average around .19 lower than the true value. Clearly, the effect of inaccurate $p_a$ or $p_b$ values can be very significant and is a major problem with any estimation procedure such as Procedure I and the Evans method (Evans et al. 1984) that relies on prior estimates. Notice also from Figure 4 that the bias of $\rho^*_{ab}$ is more sensitive to deviations in $p_b$. As a result, we recommend that the strength mode in which one has the most reliable prior information be used as mode $B$.

## 1.3 Optimal Proof-load Failure Rates $p_a$ and $p_b$ for Procedure I

Given an actual correlation $\rho_{ab}$ and the sample size $n$ we can find the values of $p_a$ and $p_b$ that minimize the predicted standard error given by (6), using the Nelder–Mead multidimensional simplex method (Press et al. 1988). Figure 5 plots the optimal $p_a$ and $p_b$ values for actual $\rho_{ab}$ values

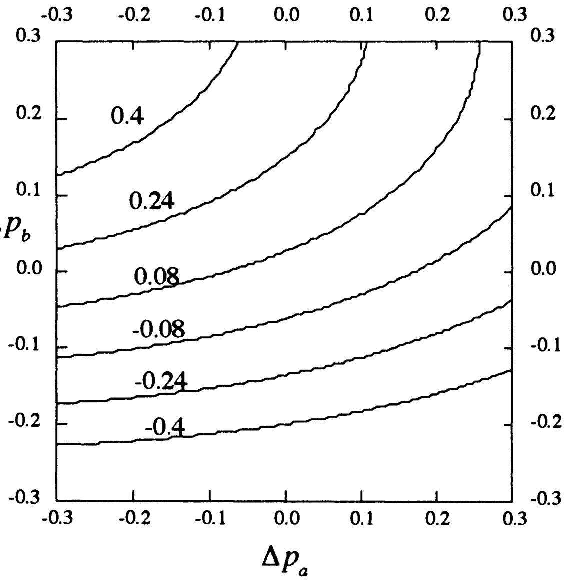


*Figure 4. Contour Plot of the Correlation Estimate Bias When Prior Estimates of $p_a$ and $p_b$ Are Incorrect. $\Delta p_a$ and $\Delta p_b$ equal the deviation of the prior estimates from the actual values: $p_a = .65$, $p_b = .5$, $\rho_{ab} = .6$.*

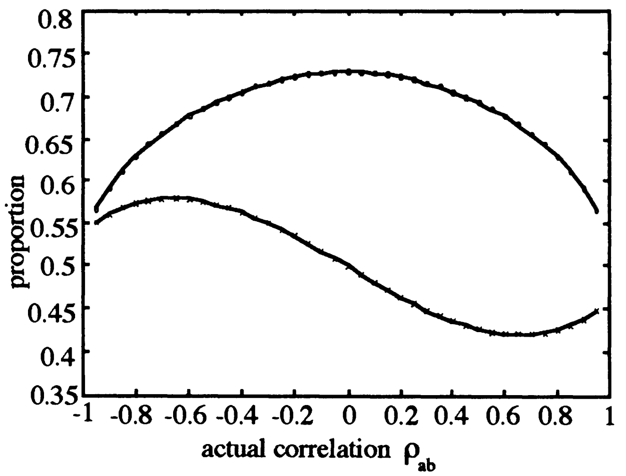


*Figure 5. Optimal $p_a$ and $p_b$ Values for Procedure I: o = Best $p_a$, x = Best $p_b$.*

between $-.95$ and $.95$. The curves showing the optimal $p_a$ and $p_b$ values are quadratic and cubic in nature, respectively, and can be very closely approximated from a regression analysis as

$$\begin{aligned} \text{optimal } p_a &= .733 - .165\rho_{ab}^2 \\ \text{optimal } p_b &= .499 - .184\rho_{ab} + .146\rho_{ab}^3. \end{aligned} \tag{7}$$

Unfortunately, determining the optimal values for $p_a$ and $p_b$ requires a prior estimate for the correlation $\rho_{ab}$. Due to the relative insensitivity of the standard error of $\rho_{ab}^*$ near the optimal $p_a$ and $p_b$ values, however, choosing good $p_a$ and $p_b$ values can be done even with little idea of the actual $\rho_{ab}$. With little prior information regarding the correlation level, we recommend proof-load levels close to $p_a = .65$ and $p_b = .5$.

## 2. PROCEDURE II: SYMMETRIC PROCEDURE

In developing the one-way estimation procedure, we assumed that good prior estimates of the means and variances of the two characteristics are available. In practice, these estimates may either be inaccurate due to quality changes or unavailable due to a lack of experience with the process. The symmetric procedure outlined here alleviates this difficulty by using the experimental results to estimate not only $p_{a\cap b}$ but also $p_a$ and $p_b$. The simplicity of the one-way procedure is retained by still considering only one proof-load level in each mode; however, to obtain more information we reverse the order of the proof-loads for some of the units. A similar type of procedure that extends the methodology of Amorim (1982) was suggested by Green and Evans (1983). They suggested subjecting half the sample to a proof-load in mode $A$ followed by stress until failure in mode $B$ and the other half to proof-load in mode $B$ followed by stress until failure in mode $A$. They reported good results in estimating all five parameters of a bivariate normal distribution but did not present their results or analysis.

Our proposed symmetric procedure is outlined as follows:

1. Start with a sample of size $n + m$.
2. Perform the one-way procedure of Section 1 on $n$ units.
3. Perform the one-way procedure in reverse order on the remaining $m$ units.

Denote the number of the first $n$ units that break under proof-loads $\mathrm{PL}_a$ and $\mathrm{PL}_b$ as $n_a$ and $n_b$, respectively, and denote the number of the $m$ units, in the reverse order test, that fail under the proof-loads $\mathrm{PL}_b$ and $\mathrm{PL}_a$ as $m_b$ and $m_a$, respectively. Using the notation of Section 1 the likelihood function for the symmetric procedure is

$$\begin{aligned} L(n, m, n_a, n_b, m_a, m_b) &= p_a^{n_a}(p_b - p_{a\cap b})^{n_b} \\ &\times (1 - p_a - p_b + p_{a\cap b})^{n-n_a-n_b} p_b^{m_b} \\ &\times (p_a - p_{a\cap b})^{m_a}(1 - p_a - p_b + p_{a\cap b})^{m-m_a-m_b}. \end{aligned} \tag{8}$$

This symmetric procedure provides more information than the one-way procedure. Solving for the MLE's gives

$$\begin{aligned} p_a^* &= \frac{n_a(n_a + n_b + m_a + m_b)}{(n+m)(n_a + n_b)} \\ p_b^* &= \frac{m_b(n_a + n_b + m_a + m_b)}{(n+m)(m_a + m_b)} \\ p_{a\cap b}^* &= \frac{(m_b n_a - m_a n_b)(n_a + n_b + m_a + m_b)}{(n+m)(n_a + n_b)(m_a + m_b)}. \end{aligned} \tag{9}$$

Note that $p_{a\cap b}^*$ may be less than 0, although this makes no physical sense. Unless $p_a$ and $p_b$ are quite small, however, that is very unlikely because normally we will observe $n_a > m_a$ and $m_b > n_b$. If the experiment results in $p_{a\cap b}^* < 0$, we recommend setting $p_{a\cap b}^* = 0$. As in Section 1, the truncated $p^*_{a\cap b}$ is technically no longer the MLE. Subsequent analysis ignores this recommended truncation and is only valid when such truncation is unlikely. In any event, the practical effect of the truncation is to reduce the variance of the estimate.

Given $p^*_{a\cap b}$ the corresponding $\rho^*_{ab}$ can be obtained using the methodology of Section 1. Rather than using known values for $p_a$ and $p_b$, however, we use the MLE's $p^*_a$ and $p^*_b$ that is, solve $p^*_{a\cap b} = f(p^*_a, p^*_b, \rho^*_{ab})$ for $\rho^*_{ab}$. The method is summarized as follows:

1. Choose sample sizes $n$ and $m$ and proof-load levels in modes $A$ and $B$.
2. Perform the experiment as outlined to obtain $n_a$, $n_b$, $m_a$, and $m_b$.
3. Use Expressions (9) to obtain the MLE's $p^*_a$, $p^*_b$, and $p^*_{a\cap b}$, and determine $h^* = \Phi^{-1}(1 - p^*_a)$ and $k^* = \Phi^{-1}(1 - p^*_b)$. If $p^*_{a\cap b} < 0$, set $p^*_{a\cap b} = 0$.
4. Using the calculated $p^*_{a\cap b}$, $h^*$, $k^*$, find the corresponding $\rho^*_{ab}$ through (3).

Substituting (3a) into (8) and numerically searching for the values of $p_a$, $p_b$, and $\rho_{ab}$ that maximize (8) is an alternative approach to finding the MLE's. This direct method yields the same results as our two-step approach, however, and has the disadvantage of requiring a simultaneous search in three dimensions.

## 2.1 Properties of the MLE's for Procedure II

We restrict analysis to the case in which $n_a + n_b \neq 0$ and $m_a + m_b \neq 0$. This restriction is necessary because if either $n_a + n_b = 0$ or $m_a + m_b = 0$—that is, no units fail in either proof-load—not enough information is obtained and two or more of the MLE's are undefined. The likelihood of observing either $n_a + n_b = 0$ or $m_a + m_b = 0$ is in most cases very small.

Given the restriction, the MLE's given by Equations (9) can be shown to be unbiased, and the standard error and bias of $\rho^*_{ab}$ can be estimated based solely on the observed sample using the jackknife method (Efron 1981). Consider the correlation estimate obtained from each subsample of the original sample that has one observation removed. We can think of $\rho^*_{ab}$ as being a function of the experimental outcome; that is, $\rho^*_{ab} = h(n_a, n_b, m_a, m_b, n + m)$. Fortunately, in our case, due to the discreteness of our data we need consider only five distinct cases $\rho^i_{ab}$ with corresponding weights $w_i$:

$$
\begin{aligned}
\rho^1_{ab} &= h(n_a - 1, n_b, m_a, m_b, n + m - 1), & w_1 &= n_a\\
\rho^2_{ab} &= h(n_a, n_b - 1, m_a, m_b, n + m - 1), & w_2 &= n_b\\
\rho^3_{ab} &= h(n_a, n_b, m_a - 1, m_b, n + m - 1), & w_3 &= m_a\\
\rho^4_{ab} &= h(n_a, n_b, m_a, m_b - 1, n + m - 1), & w_4 &= m_b\\
\rho^5_{ab} &= h(n_a, n_b, m_a, m_b, n + m - 1), & &\\
& & w_5 &= n + m - n_a - n_b - m_a - m_b.
\end{aligned}
\tag{10}
$$

Based on these additional correlation estimates, the jackknife estimate of the bias BIAS, and standard error S of $\rho^*_{ab}$ are

$$
\mathrm{BIAS} = \frac{n+m-1}{n+m}\sum_{k=1}^{5} w_k(\rho^*_{ab} - \rho^k_{ab})
$$

$$
\mathrm{S} = \sqrt{\frac{n+m-1}{n+m}\sum_{k=1}^{5} w_k\left[\sum_{i=1}^{5}\left(\frac{w_i}{n+m}\rho^i_{ab}\right) - \rho^k_{ab}\right]^2}. \tag{11}
$$

Because in most cases the bias of $\rho^*_{ab}$ is quite small, the jackknife bias adjustment is of little value. On the other hand, simulation results suggest that the jackknife estimates for the standard error of the correlation estimate are usually very good.

## 2.2 Sensitivity Analysis for the Symmetric Procedure

It is of interest to determine the effect of sample size, proof-load levels, and actual correlation values on the bias and standard error of the correlation estimate $\rho^*_{ab}$ derived by Procedure II. Simulation results suggest that the bias of $\rho^*_{ab}$ using the symmetric procedure follows a very similar pattern to that exhibited by Procedure I (see Fig. 1). In other words, the bias is very small unless one or both the proof-loads are small ($p_a$ or $p_b < .2$). Figure 6 explores the simulated standard error of $\rho^*_{ab}$ for various proof-load levels.

Based on Figure 6 and additional simulation studies, we recommend trying to choose proof-load levels that result in about 60% failures. Near these proof-load levels, the standard error of $\rho^*_{ab}$ is small and relatively insensitive to changes in $p_a$ and $p_b$. This insensitivity is very important because when the individual means and standard deviations are not known with certainty it is impossible to set proof-load levels exactly at a desired level.

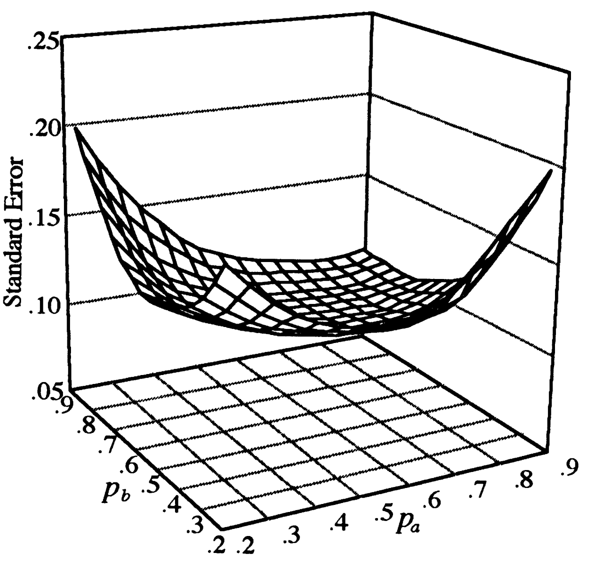


*Figure 6. Standard Error of the Simulated Correlation Estimate: $n = m = 150$, $\rho_{ab} = .6$, 10,000 Samples.*

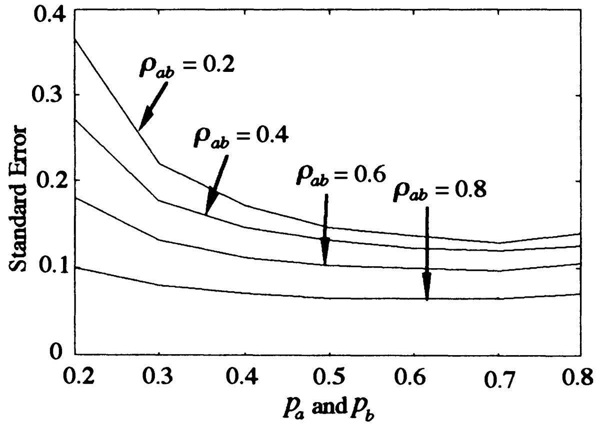


*Figure 7. Standard Error as a Function of Proof-load Levels, $n = m = 150$.*

The symmetric procedure also exhibits performance similar to Procedure I with regards to the effect of sample size (see Fig. 3). Because for the symmetric procedure we are unable to set the proof-load probabilities $p_a$ and $p_b$, however, we are very interested in the effect they have on the procedure's estimates. Figure 7 shows simulated results of standard error of $\rho^*_{ab}$ as a function of proof-load probabilities and true correlation levels.

Clearly, as shown in Figure 7, the estimation procedure works best when $\rho_{ab}$ is large and $p_a$ and $p_b$ are not small. In our experience, the estimation procedure works very well unless $p^*_{a\cap b}$ (or $m_b n_a - n_b m_a$) is likely to be negative or close to 0. When $p^*_{a\cap b}$ is close to 0, the procedure is not very stable because small changes in the experimental outcome $(n_a, n_b, m_a, m_b)$ result in large changes in the corresponding $\rho^*_{ab}$ value. Fortunately, $m_b n_a - n_b m_a$ is only likely to be close to 0 when $p_a$ or $p_b$ or both are small and $\rho_{ab}$ is not strongly positive.

## 3. EXAMPLE

To illustrate the two procedures, suppose that we are interested in estimating the correlation between the bending and tension strength of lumber.

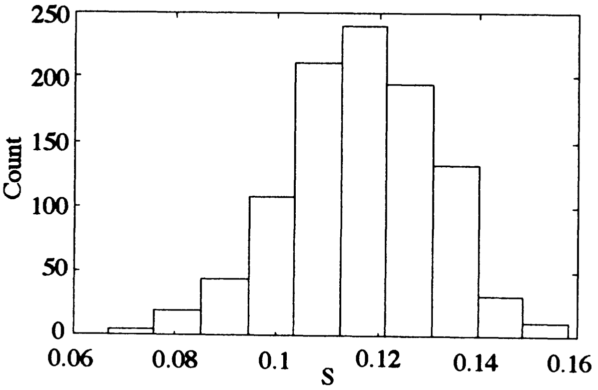


*Figure 8. Histogram of the Jackknife Standard Error Estimate.*

Following the summary of Procedure I resulted in the following steps:

1. A sample of 300 lumber specimens were taken (i.e., $n = 300$), and the proof-load levels in bending and tension were chosen so that $p_a = .65$ and $p_b = .45$; that is, $h = -.385$ and $k = .126$
2. The experiment resulted in 190 units failing under the proof-load in bending and 23 units of the remaining 110 units failing under the proof-load in tension. Thus $n_a = 190$ and $n_b = 23$.
3. Using these variables and Equations (2) and (3) gave $p^*_{a\cap b} = .377$ and $\rho^*_{ab} = .557$. From Equation (6), we estimated $se(\rho^*_{ab}) = .088$.

Following the summary of Procedure II, we performed the following steps:

1. A sample of 300 lumber specimens was taken choosing $n = m = 150$.
2. Running the experiment gave $n_a = 96$ and $n_b = 11$, $m_a = 42$ and $m_b = 65$.
3. Substituting into Equations (9) gave $p^*_a = .64$, $p^*_b = .43$, and $p^*_{a\cap b} = .36$. Thus $h = -.359$ and $k = .168$.
4. Inverting (3) gave $\rho^*_{ab} = .547$. The jackknife procedure, Equations (10) and (11), gave $\rho^1_{ab} = .546$, $\rho^2_{ab} = .573$, $\rho^3_{ab} = .556$, $\rho^4_{ab} = .549$, and $\rho^5_{ab} = .544$, which results in a standard error estimate of S = .113.

In both cases the example was generated randomly using a bivariate normal distribution with $p_a = .65$, $p_b = .45$, and $\rho_{ab} = .5$ (thus $p_{a\cap b} = .368$). Simulation results suggest that for Procedure I $se(\rho^*_{ab}) = .091$ and for Procedure II $se(\rho^*_{ab}) = .1168$.

For comparison, we also include Figure 8, which shows simulation results of the jackknife estimate for the standard error of Procedure II. Figure 8 shows a histogram of 1,000 trials with a mean jackknife standard error estimate of .116.

## 4. COMPARISON OF RESULTS WITH EVANS'S METHOD

Table 1 compares the best standard errors of $\rho^*_{ab}$ obtained by simulating the Evans et al. (1984) method and our proposed one-way and symmetric procedures. As expected, the standard errors of $\rho^*_{ab}$ for our Procedures

*Table 1. Best Standard Error of $\rho^*_{ab}$*

| Actual $\rho_{ab}$ | Evans's procedure $n = 300$ | Proposed Procedure I $n = 300$ | Proposed Procedure I $n = 365$ | Proposed Procedure II $n + m = 300$ |
|---|---|---|---|---|
| .2 | .0895 | .1101 | .0972 | .1328 |
| .6 | .0649 | .0804 | .0709 | .0982 |
| .8 | .0400 | .0520 | .0456 | .0701 |
| .9 | .0257 | .0310 | .0284 | .0444 |

I and II are higher (about 20–25% and 40–60%, respectively). This decrease in efficiency, however, will be compensated by the lower cost and greater ease of implementation. Both Procedures I and II have the advantage of not requiring the determination of precise breaking strengths because only proof-loading is used.

Moreover, because the proposed procedures use only proof-loading, some of the units tested are not destroyed and can be returned to the population of unused samples. The average number of units not destroyed is easily approximated. For example, using Procedure I with $p_a = .65$ and $p_b = .5$ results in the survival of 17.5% of the units on average. These units not only survive the testing but are in fact the strongest units in the sample because they have withstood two proof-loads. For comparison, Table 1 also shows results for Procedure I with $n = 365$ (if 17.8% of the units do not fail, the procedure will destroy 300 units). The standard errors for Procedure I are now only about 10% higher than those obtained with the Evans method. Procedure II, the symmetric procedure, has the advantage of not requiring accurate prior mean and standard-deviation estimates. When such prior estimates are unavailable or incorrect, the symmetric procedure may well outperform the one-way procedure and previously developed procedures.

## 5. RECOMMENDATIONS AND CONCLUSIONS

For the practitioner, these two new procedures offer simple and practical ways to obtain good estimates of the correlation between two variables that can only be determined destructively. The proposed procedures require only one pass/fail proof-load in each strength mode. In both cases, optimal proof-load levels depend on the unknown correlation coefficient; however, both suggested procedures are not overly sensitive to changes in the proof-load levels near the optimal levels. Therefore, for the one-way procedure, a good rule of thumb is to choose the proof-loads such that on average 65% of the units break on the first test, and an average of 50% break on the second test (i.e., 50% would break if the second test were done first). When the means and standard deviations of the individual variables are unknown, the symmetric procedure is applicable, but choosing good proof-load levels is more difficult. In this case, the practitioner should aim for proof-load levels at which 60% fail proof-load levels in each mode.

## ACKNOWLEDGMENTS

This research was supported, in part, by the Natural Sciences and Engineering Research Council of Canada. We thank two anonymous referees and an associate editor for their many helpful comments and suggestions that considerably improved this article.

*[Received September 1992. Revised June 1994.]*

## REFERENCES

Abramowitz, M., and Stegun, I. A. (eds.) (1972), *Handbook of Mathematical Functions* (Applied Mathematics Series 55), Washington, DC: National Bureau of Standards.

Amorim, S. (1982), "Experimental Designs for Estimating the Correlation Between Two Destructively Tested Variables," unpublished Ph.D. thesis, University of Wisconsin-Madison, Dept. of Statistics.

Amorim, S., and Johnson, R. A.(1986), "Experimental Designs for Estimating the Correlation Between Two Destructively Tested Variables," *Journal of the American Statistical Association*, 81, 807–811.

Bartlett, N. R., and Lwin, T. (1984), "Estimating a Relationship Between Different Destructive Tests on Timber," *Applied Statistics*, 33, 65–72.

Drezner, Z., and Wesolowsky, G. O.(1990), "On the Computation of the Bivariate Normal Integral," *Journal of Statistical Computation and Simulation*, 35, 101–107.

Efron, B. (1981), *The Jackknife, the Bootstrap, and Other Resampling Plans* (CBMS Monogram No. 38), Philadelphia: Society of Industrial and Applied Mathematics.

Evans, J. W., Johnson, R. A., and Green D. W. (1984), "Estimating the Correlations Between Variables Under Destructive Testing, or How to Break the Same Board Twice," *Technometrics*, 26, 285–290.

Galligan, W. L., Johnson, R. A., and Taylor, J. R. (1979), "Examination of the Concomitant Properties of Lumber," in *Proceedings of the Metal Plate Wood Truss Conference*, St. Louis, MO, Madison, WI: Forest Products Research Society. pp. 65–70.

Gerhards, C. C. (1979), "Time-Related Effects of Loading on Wood Strength: A Linear Cumulative Damage Theory," *Wood Science*, 11, 139–144.

Grab, E. L., and Savage, I.R. (1954), "Tables of the Expected Value of 1/X for Positive Bernoulli and Poisson Variables," *Journal of the American Statistical Association*, 169–177.

Green, D. W., and Evans, J. W. (1983), "Estimating Correlation Between Strength Properties," in *Proceedings of the Fourth Engineering Mechanics Division Specialty Conference: Recent Advances in Engineering Mechanics and Their Impact on Civil Engineering Practice* (Vol. 2), West Lafayette, IN: American Society of Civil Engineers, Purdue University, pp. 936–939.

Green, D. W., Evans, J. W., and Johnson, R.A. (1984), "Investigation of the Procedure for Estimating the Concomitance of Lumber Strength Properties," *Wood and Fiber Science*, 16, 427–440.

Johnson, N. L., and Kotz, S. (1969), *Distributions in Statistics: Discrete Distributions*, Boston: Houghton Mifflin.

——— (1989), *Encyclopedia of Statistical Sciences*, New York: John Wiley.

Johnson, R. A. (1980), "Current Statistical Methods for Estimating Lumber Properties by Proof-loading," *Forest Products Journal*, 30, 14–22.

Johnson, R. A., and Galligan, W. L. (1983), "Estimating the Concomitance of Lumber Strength Properties," *Wood and Fiber Science*, 15, 235–244.

Madsen, B. (1976), "In-Grade Testing: Degree of Damage Due to Proofloading of Lumber in Bending," Report 17, Structural Research Series,University of British Columbia, Dept. of Civil Engineering.

Press, W. H., Flannery. B. P., Teukolsky, S.A., and Vetterling, W.T. (1988), *Numerical Recipes*, Cambridge, U.K.: Cambridge University Press.

Sheppard, W. F. (1900), "On the Calculation of the Double Integral Expressing Normal Correlation," *Transactions of the Cambridge Philosophical Society*, 19, 23–69.

Suddarth, S. K., Woeste, F. E., and Galligan, W. L. (1978), "Probabilistic Engineering Applied to Wood Members in Bending/Tension," Research Paper FPL 302, U.S. Dept. of Agriculture, Forest Products Laboratory, Madison, WI.

Strickler, M. D., Pellerin, R. F., and Talbot, J. W. (1970), "Experiments in Proofloading Structural End Joints Lumber," *Forest Products Journal*, 20, 29–35.