Multivariate Behavioral Research, 24 (3), 365-395

A Comparison of Alternatives to Conducting Monte Carlo Analyses for Determining Parallel Analysis Criteria

Gary J. Lautenschlager
University of Georgia

The parallel analysis method for determining the number of components to retain in a principal components analysis has received a recent resurgence of support and interest. However, researchers and practitioners desiring to use this criterion have been hampered by the required Monte Carlo analyses needed to develop the criteria. Two recent attempts at presenting regression estimation methods to determine eigenvalues were found to be deficient in several respects, and less accurate in general, than a simple linear interpolation of tabled random data eigenvalues generated through Monte Carlo simulation. Other methods for determining the parallel analysis criteria are discussed.

Researchers using principal components analysis, either as a data reduction method or as a preliminary step in common factor analysis, have available a number of criteria for determining the number of factors to retain for rotation and/or interpretation. Perhaps the most prevalent method in use is Kaiser's (1960) criterion of retaining eigenvalues greater than one for determining the number of components, or common factors, to retain. However, applying Kaiser's rule, or any of several other rules that are available, may not be as accurate as using the parallel analysis (PA) criterion. It has been suggested that both the minimum average partial correlation method (Velicer, 1976) and the PA method might be employed to reach a decision about the number of components to retain (Zwick and Velicer, 1986). Cliff (1988) has also called into question the use of eigenvalues-greater-than-one to determine the number of components or common factors. Horn (1965) initially proposed the PA method as a sample-based adaption of Kaiser's (1960) rule of retaining all

---

The author wants to thank Rolf E. Bargmann for invaluable discussions and help in developing a computer program and conducting analyses. In addition, the comments of three anonymous reviewers and discussions with Richard J. Harris were also helpful.

This research was facilitated by additional equipment support provided in August, 1985 by the Office of the Vice President For Research, The University of Georgia.

Portions of this work were presented at the annual meeting of the American Psychological Association, Atlanta, August, 1988.

Correspondence concerning this article should be addressed to Gary J. Lautenschlager, Department of Psychology, University of Georgia, Athens, GA, 30602, USA.

eigenvalues greater than or equal to 1.00. The basic rationale underlying the PA criterion is that "meaningful" components extracted from actual sample data should tend to have eigenvalues larger in size than eigenvalues of the same order obtained from random normal variates generated to simulate the same sample size and number of variables. Since it is the expected value of eigenvalues of random data that are used for the comparison, it is necessary to simulate many such datasets and average the respective eigenvalues obtained to estimate the expected values. A stumbling block for those desiring to implement the PA criterion in practice is the required number of Monte Carlo analyses.

Analytical Determination of Eigenvalues

There is good reason to initially question whether using Monte Carlo simulations to estimate expectations for eigenvalues is even warranted. Much analytical work on the eigenvalues of random Wishart matrices is available in the statistical literature (Hsu, 1939; Roy, 1953; Krishnaiah, 1980; Clemm, Chattopadhyay & Krishnaiah 1973; Clemm, Krishnaiah & Waikar, 1973; Krishnaiah & Chang, 1971).[1] Approximations and very precise numerical methods have been used to generate tables of percentage points for extreme and intermediate roots of such matrices, and much of this work has been summarized by Krishnaiah (1980). Tables are available which present percentiles of the individual roots of certain classes of Wishart matrices. However, an analytical solution for the distribution of the characteristic roots of correlation matrices, even for the case of independence, is not available. Of course, Krishnaiah's tables of percentiles applied to a sample covariance matrix with unit diagonals, could serve as a useful approximation.

Even if the available tables for the covariance matrix case were applicable there would still remain several problems with implementing them for use as PA criteria. First, the analytically derived tables (Krishnaiah, 1980) cover only a very limited range of values of P and N, much more limited than was covered in the Monte Carlo data simulated in the present study. Second, these tables were derived for the simple Wishart case where the matrix of covariances is used under the assumption that the variance of each random variate was two. Third, the values in the tables represent the extreme upper percentage points of individual roots. Horn (1965) had based his development on using the expectation for the individual random data eigenvalues as the PA criterion. Medians could be easily obtained from the probability distributions, but this does not guarantee that the expectations can be easily derived. Regardless of the significance of the preceding matters, the critical problem with using the

[1] I would like to thank an anonymous reviewer for pointing out the relevance of P. R. Krishnaiah's (1980) work.

analytically-derived tables is that they were derived from variance-covariance matrices and are not applicable to the correlation matrix case. Numerical integration based on an analytical solution for the roots of random correlation matrices appears to be an intractable problem (R. E. Bargmann, personal communication, January 9, 1989).

Regression Estimates of Eigenvalues

Allen and Hubbard (1986) have presented a means by which parallel analysis could ostensibly be made more accessible to practitioners. Their work was based on earlier work by Montanelli and Humphreys (1976; Humphreys & Montanelli, 1975) which had presented a general form of a recursive regression equation to be used for estimating the eigenvalues of random data correlation matrices with squared multiple correlations on the diagonal. Allen and Hubbard derived a general form of a regression equation designed to predict values of all but the two smallest (last) eigenvalues of a random data correlation matrix with unities on the diagonal. The Allen and Hubbard equation is:

$$[1] \qquad \log(\lambda_k) = a_k + b_k \log(N\text{-}1) + c_k \log\{(P\text{-}k\text{-}1)(P\text{-}k\text{+}2)/2\} + d_k \log(\lambda_{k\text{-}1})$$

where log is the natural logarithm, N is sample size, P is the number of variables, k [$k = 1,2,...,(P\text{-}2)$] indexes the k_{th} eigenvalue λ_k, ($\lambda_0 = 1.0$) and a_k, b_k, c_k and d_k are regression model parameters. The term weighted by c_k was derived from a degrees-of-freedom rationale offered by Bartlett (1951) and Lawley (1956). The inclusion of this term necessarily restricts eigenvalue estimates to the first P-2 eigenvalues. An excellent fit was obtained between predicted and observed values for all but the first eigenvalue by including the d_k term. Allen and Hubbard's (1986) empirical equations (hereafter referred to as the *AH* equations) have been implemented in the computer program PARALLEL (Hays, 1987).

Recently, Lautenschlager, Lance and Flaherty (1989; Flaherty, Lautenschlager & Lance, 1988) have improved upon Allen and Hubbard's (1986) general equation to provide a more nearly exact estimate of the first eigenvalue. They employed the following revised equation:

$$[2] \qquad \log(\lambda_k) = a_k + b_k \log(N\text{-}1) + c_k \log\{(P\text{-}k\text{-}1)(P\text{-}k\text{+}2)/2\} + d_k \log(\lambda_{k\text{-}1}) + e_k P/N$$

which involves one additional predictor term in addition to the predictors used in equation 1, namely the ratio of the number of variables to the sample size. Allen and Hubbard reported an $R^2 = .931$ between the actual first eigenvalues and predicted first eigenvalues for their data using equation 1. Lautenschlager

et al. (1989) obtained an $R^2 = .933$ for their data using equation 1, and obtained an $R^2 = .993$ using equation 2. The increment in R^2 was statistically significant ($F_{(1,92)} = 782.01, p<.0001$) using equation 2 over equation 1. Revised empirical regression equations (hereafter referred to as the *LLF* equations) would appear to have the advantage of making better predictions of the first eigenvalue, and hence better predictions all around, owing to the fact that the regression estimates are recursive, meaning estimates of initial eigenvalues figure into estimates of later eigenvalues in applications of equations 1 and 2. Both equations produce regression estimates of eigenvalues that ostensibly could be used as criteria for a parallel analysis.

Available procedures for implementing PA criteria in practice were compared in the present study. This involved the examination of regression equation methods that can be used to estimate random data eigenvalues from known values of the sample size and number of variables. Problems inherent in the application of these regression estimation procedures are described, and more internally accurate methods for determining PA criteria in practice are presented.

Method

Procedure

Data were generated in a series of Monte Carlo simulations in which the number of variables (P) ranged from 5 to 50 in steps of 5, and sample sizes (N) were chosen as 50, 75, 100, 150, 200, 300, 400, 500, 750, 1000, 1500 and 2000 to reflect a range typically found in published research. Values of P of 60, 70 and 80 were also examined for sample sizes of 100, 150, 200, 300, 400, 500, 750, 1000, 1500 and 2000. For each N, P combination where $P \leq 50$ and $N \leq 1000$ the SYSTAT (Wilkinson, 1986) analysis package was used to (a) generate random data for N cases on each of P variables from a $N(0,1)$ population, (b) create a correlation matrix based on these data, and (c) conduct a principal components analysis of the correlation matrix (Lautenschlager, in press). Characteristic roots for values of P 60 and for all values of $N > 1000$ were obtained using a program which generates sample Wishart matrices, W (from $\Sigma = I$), by generating a triangular matrix T, such that $TT' = W$, where the diagonal elements are chi-variate and the lower triangular values are standard normal for a given N, P combination (Bartlett, 1933; Roy, 1957; Kendall & Stuart, 1969). The matrix was standardized and all characteristic roots of the the resulting correlation matrix were obtained.[2] A check on the accuracy of the chi-variate simulations revealed fairly good approximation of the expectation for the

[2] R. E. Bargmann wrote a Pascal program for these analyses. Analyses were conducted on an 80386 PC with a numeric coprocessor, requiring about 8 hrs for 100 matrices for $P = 70$.

determinant of the sample correlation matrix for the random variate case even when $P = 80$ and N as low as 100. A check on the accuracy of all simulations revealed that deviations from the expectation of the determinant of the sample correlation matrix were greatest when the value of N was close to the value of P, as expected. The overall fit regarding the expectation for the sample determinant was generally excellent. For values of $P > 10$, one hundred replications were generated; for values of $P \leq 10$, two hundred replications were generated. A total of 17,000 unique datasets were created.

Within each fixed N, P combination the results were combined over replications and averaged to produce an average first eigenvalue based on 100 (or 200) values, an average second eigenvalue, etc. These averaged values were then used to form tables of averaged eigenvalues. These datasets had been used by Lautenschlager et al. (1989) under the restriction that only those N, P combinations where $P \leq 50$ and $N \leq 1000$ which satisfied $N \quad 3P/2$ were used in the development of the *LLF* equations.

Proposed Comparisons

The generated data provided the empirical criteria for the comparison of the regression equation estimates using the *AH* and *LLF* estimation equations. In addition, tabled values of eigenvalues based on the generated data were used to interpolate values for eigenvalues not specifically covered by the tables. The accuracy of some of these interpolations was investigated vis-à-vis the regression estimates by conducting additional Monte Carlo simulations to serve as criteria. The root mean squared error (RMSE), defined as the square root of the average squared deviation of the interpolations and the regression estimates from the averaged eigenvalues obtained from these additional simulations, was calculated in each case to permit comparisons. Because the regression estimation methods only generate P-2 estimates for any given case, all RMSE calculations involved at most only the first P-2 eigenvalues. In addition, comparisons involving the regression estimation methods were limited to values of P and N used in developing those equations. Given the problems associated with extending nonlinear models beyond available data points, this restriction was assumed desirable and necessary to enable a fair comparison.

Results

Problems with Regression Estimates of Eigenvalues

Those who wish to use *AH* or *LLF* regression equation methods to estimate the eigenvalues of random data correlation matrices can use the empirical

equations as presented in a table found in each of the respective papers. In order to produce *AH* regression estimates, one can use the PARALLEL program (Hays, 1987). The *LLF* equations have been programmed and can be implemented through use of the PARANAL program (Lautenschlager, 1987). At first blush it would appear that the *AH* equations were at a distinct disadvantage in comparison with the *LLF* equations, as the latter were derived from the generated data described above. Neither regression estimation method was found to be generally useful across the ranges of N and P from which the equations were developed.

Although for certain N, P combinations the results produced by regression estimation methods tended to agree with the Monte Carlo simulations, numerous other combinations produced clearly divergent, and often unreasonable results. As a case in point, fixing $P = 50$ (i.e., 50 variables were involved in the analysis), whenever $N > 170$ the *AH* equations began to produce predictions of subsequent eigenvalues that were larger than preceding eigenvalues (e.g. when $N = 171$ estimated eigenvalue #45 > estimated eigenvalue #44). As sample size grew larger while holding $P = 50$, even more peculiar things happened. Estimated eigenvalues decreased beyond the first value up to a point but then consistently increased. In addition, predicted eigenvalues were obtained for eigenvalues past the $(P/2)$th value that exceed the number of variables in the analysis. At the extreme high end where $N = 1000$, the estimated first eigenvalue was reasonably close to the empirical value. However, the estimate for the 48th eigenvalue was slightly over 3,849,433,795. Admittedly, this was at the extreme upper bound of both N and P, but similar problems occurred for other combinations of sample sizes and number of variables. Obviously the usefulness of these empirically obtained regression equations was not uniform across the N, P combinations studied by Allen and Hubbard (1986).

Were the *LLF* equations any better at providing estimates of eigenvalues? The answer was a qualified yes. In general, the estimates based on the empirical equations provided by Lautenschlager et al. (1989) were better behaved, but not well-behaved in terms of being generally useful to researchers. The *LLF* empirical equations tended to produce better estimates than the *AH* equations, in the sense that estimates tended to be closer to the simulated empirical data (from which the estimation equations were developed), but these equations also sufferred from the problems described in the preceding paragraph. However, for some combinations of N and P the *AH* estimates were better. The differences in precision may have been due in part to the differences in the N, P combinations employed in the two studies and to the number of replications. Although the *LLF* equations were technically "better" than the *AH* equations in terms of overall fit, they also posed problems for researchers desiring to implement the PA criterion in the numerous research situations ostensibly covered by the range

of *N* and *P* values used to derive the *AH* and *LLF* equations. It should be noted that these pronounced problems described above, which were easily identified at the extremes, were also present elsewhere in a less obvious form.

Factors Affecting Regression Estimates of Eigenvalues

One might wonder why the regression estimates of eigenvalues can be so far off, even well within the bounds of *N* and *P* used to generate the equations. After all, the smallest R^2 reported by Allen and Hubbard (1986) accounted for over 93% of the variance of the first eigenvalue, and that reported by Lautenschlager et al. (1989) accounted for over 99% of the variance. First, and obviously, the estimates provided by either set of empirically derived regression equations must produce estimates that are somewhat closer to the mean of the set of specific eigenvalues of a given order (e.g., all first eigenvalues) used in generating the equation. Given the large R^2s, this may seem a trivial factor. However, what is estimated is the natural logarithm of the eigenvalue. Small differences in the estimation accuracy of a logarithm will have a more pronounced impact on the estimation accuracy of the eigenvalue itself.

Second, as noted by Lautenschlager et al. (1989), errors in prediction are cumulative (and likely multiplicative) in effect. This is because a series of different, but dependent equations must be applied in sequence to estimate a range of eigenvalues. The prediction of the third eigenvalue must suffer since the prediction of first and second eigenvalues were both somewhat in error themselves. Because estimates of preceding eigenvalues must be used to estimate subsequent eigenvalues, the proportions of variance accounted for, as reported in the development of the regression equations in these articles, are potentially deceiving regarding this deterioration.

A third source for lack of fit is likely due to the fact that eigenvalues must be estimated within a specific *N*, *P* combination. On the other hand, the empirical equations were developed by extracting an eigenvalue of a specific order from across all *N*, *P* combinations. Variance between *N*, *P* conditions for a particular order eigenvalue is clearly accounted for, but some variance across eigenvalues within conditions is not captured in the process.

A fourth source of lack of fit is associated with the assumption about the error component implicit in the regression estimation procedures. The natural logarithm transformation employed in both regression estimation methods is based on the assumption that the error term structure is multiplicative. The potential consequences of the linearization of a nonlinear model include both unreasonable error structures and poorer estimates of model parameters (Myers, 1986).

Another Option for Determining Parallel Analysis Criteria

Although the preceding discussion may seem to imply that the PA criterion can only be employed by those willing to conduct their own Monte Carlo simulations, there was another alternative. Given that current regression estimation methods could not generally be recommended even across the range of the N, P values from which the equations themselves were developed, a better approach was desired. An alternative means for applying the PA method was derived through the direct use of the available simulated data itself. Tables 1 through 13 were constructed and represent the averaged eigenvalues from principal components analyses of the 17,000 random data correlation matrices. The tables reflect all N, P combinations described earlier, and were arranged so that each table covered a specific value of P and presented the average eigenvalues arranged in decreasing order of derivation, obtained over replications for each specific sample size. These tables represent an alternative implementation of PA criteria, one similar in spirit to that intended by either regression estimation method. The tables were subsequently used to examine the merits of the regression estimation methods described above versus simple linear interpolation of values contained within the tables.[3]

Table 1
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 5

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	1.392	1.330	1.297	1.234	1.201	1.166	1.143	1.127	1.104	1.089	1.068	1.064
2	1.161	1.134	1.120	1.096	1.086	1.070	1.060	1.052	1.046	1.039	1.031	1.027
3	0.990	0.984	0.989	0.996	0.993	0.995	1.001	0.999	0.999	0.999	1.001	0.998
4	0.818	0.853	0.867	0.893	0.912	0.925	0.937	0.944	0.953	0.960	0.969	0.972
5	0.640	0.699	0.729	0.783	0.809	0.844	0.860	0.880	0.900	0.914	0.932	0.938

[3] Copies of the tables are available on an MS-DOS compatible diskette from the author.

Table 2
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 10

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	1.787	1.611	1.528	1.435	1.366	1.294	1.260	1.230	1.186	1.158	1.128	1.112
2	1.529	1.421	1.358	1.297	1.249	1.205	1.179	1.158	1.129	1.110	1.090	1.080
3	1.330	1.268	1.229	1.188	1.161	1.136	1.118	1.104	1.084	1.074	1.059	1.054
4	1.159	1.145	1.122	1.102	1.087	1.074	1.067	1.059	1.048	1.043	1.035	1.031
5	1.012	1.025	1.029	1.021	1.021	1.020	1.016	1.015	1.013	1.015	1.010	1.009
6	0.876	0.915	0.926	0.945	0.957	0.967	0.969	0.975	0.980	0.985	0.988	0.989
7	0.760	0.810	0.841	0.870	0.896	0.914	0.925	0.933	0.947	0.954	0.963	0.966
8	0.633	0.710	0.753	0.800	0.828	0.858	0.877	0.892	0.913	0.923	0.938	0.945
9	0.519	0.605	0.658	0.719	0.757	0.802	0.826	0.847	0.873	0.892	0.911	0.922
10	0.399	0.493	0.559	0.627	0.680	0.732	0.765	0.790	0.830	0.850	0.876	0.892

Table 3
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 15

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	2.065	1.861	1.724	1.573	1.491	1.399	1.349	1.306	1.253	1.212	1.170	1.150
2	1.799	1.636	1.551	1.442	1.378	1.310	1.268	1.239	1.194	1.167	1.133	1.118
3	1.605	1.486	1.417	1.347	1.296	1.235	1.212	1.188	1.154	1.131	1.106	1.091
4	1.438	1.350	1.308	1.252	1.219	1.172	1.158	1.143	1.115	1.099	1.081	1.070
5	1.281	1.231	1.215	1.181	1.150	1.122	1.110	1.103	1.084	1.071	1.060	1.051
6	1.140	1.125	1.119	1.106	1.093	1.077	1.071	1.063	1.054	1.045	1.039	1.033
7	1.024	1.033	1.039	1.036	1.035	1.032	1.031	1.025	1.022	1.022	1.017	1.014
8	0.915	0.946	0.961	0.973	0.984	0.987	0.991	0.991	0.993	0.998	0.997	0.997
9	0.815	0.867	0.882	0.913	0.929	0.944	0.952	0.958	0.966	0.972	0.977	0.981
10	0.710	0.777	0.811	0.853	0.877	0.901	0.913	0.923	0.938	0.948	0.959	0.964
11	0.609	0.701	0.743	0.790	0.821	0.855	0.874	0.889	0.907	0.923	0.939	0.946
12	0.529	0.616	0.669	0.726	0.772	0.811	0.836	0.854	0.880	0.897	0.917	0.926
13	0.444	0.537	0.601	0.673	0.714	0.770	0.793	0.816	0.849	0.870	0.894	0.908
14	0.361	0.461	0.527	0.603	0.657	0.720	0.749	0.773	0.816	0.843	0.872	0.889
15	0.264	0.373	0.433	0.534	0.585	0.665	0.692	0.729	0.775	0.803	0.840	0.861

Table 4
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 20

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	2.334	2.050	1.909	1.716	1.604	1.492	1.422	1.371	1.301	1.259	1.208	1.177
2	2.041	1.837	1.710	1.577	1.483	1.401	1.337	1.300	1.247	1.212	1.172	1.148
3	1.835	1.673	1.584	1.469	1.402	1.330	1.283	1.251	1.206	1.178	1.145	1.125
4	1.669	1.543	1.470	1.378	1.332	1.274	1.238	1.208	1.172	1.148	1.121	1.105
5	1.513	1.432	1.373	1.301	1.266	1.217	1.192	1.171	1.140	1.122	1.099	1.085
6	1.379	1.329	1.287	1.233	1.209	1.164	1.146	1.131	1.108	1.095	1.081	1.068
7	1.255	1.221	1.196	1.170	1.151	1.123	1.111	1.097	1.080	1.071	1.061	1.051
8	1.139	1.130	1.117	1.106	1.097	1.080	1.075	1.069	1.053	1.049	1.043	1.036
9	1.029	1.041	1.038	1.046	1.044	1.039	1.039	1.036	1.029	1.026	1.024	1.020
10	0.940	0.956	0.970	0.991	0.994	1.005	1.003	1.006	1.005	1.004	1.005	1.005
11	0.840	0.881	0.908	0.936	0.948	0.964	0.966	0.975	0.980	0.983	0.987	0.990
12	0.747	0.812	0.843	0.878	0.901	0.925	0.935	0.947	0.955	0.961	0.970	0.975
13	0.667	0.740	0.781	0.827	0.856	0.885	0.900	0.914	0.931	0.941	0.953	0.960
14	0.585	0.671	0.720	0.779	0.812	0.845	0.870	0.888	0.908	0.922	0.936	0.945
15	0.504	0.602	0.661	0.728	0.768	0.808	0.835	0.858	0.882	0.901	0.916	0.929
16	0.433	0.544	0.606	0.679	0.718	0.772	0.804	0.825	0.857	0.877	0.899	0.912
17	0.369	0.480	0.550	0.629	0.675	0.734	0.769	0.795	0.831	0.853	0.879	0.896
18	0.304	0.417	0.490	0.577	0.630	0.695	0.733	0.759	0.804	0.830	0.858	0.880
19	0.244	0.359	0.428	0.521	0.583	0.650	0.695	0.722	0.774	0.803	0.834	0.859
20	0.175	0.284	0.358	0.459	0.526	0.597	0.646	0.679	0.736	0.767	0.808	0.833

Table 5
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 25

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	2.588	2.257	2.053	1.840	1.713	1.574	1.494	1.432	1.345	1.298	1.239	1.206
2	2.289	2.037	1.878	1.689	1.608	1.486	1.412	1.365	1.293	1.255	1.205	1.174
3	2.064	1.853	1.745	1.586	1.517	1.417	1.351	1.311	1.254	1.220	1.178	1.152
4	1.870	1.717	1.627	1.506	1.440	1.357	1.309	1.270	1.219	1.190	1.155	1.134
5	1.722	1.592	1.519	1.424	1.373	1.297	1.259	1.234	1.189	1.163	1.134	1.115
6	1.584	1.487	1.429	1.357	1.312	1.248	1.217	1.197	1.159	1.141	1.114	1.098
7	1.461	1.383	1.344	1.289	1.256	1.208	1.178	1.163	1.134	1.115	1.094	1.083
8	1.342	1.292	1.268	1.221	1.199	1.162	1.142	1.130	1.108	1.094	1.078	1.067
9	1.225	1.214	1.182	1.166	1.148	1.119	1.107	1.101	1.083	1.074	1.060	1.053
10	1.121	1.121	1.110	1.109	1.096	1.082	1.073	1.069	1.061	1.053	1.043	1.039
11	1.029	1.039	1.047	1.054	1.047	1.046	1.042	1.042	1.037	1.033	1.028	1.025
12	0.934	0.962	0.980	1.002	1.003	1.008	1.008	1.013	1.013	1.012	1.011	1.010
13	0.843	0.893	0.918	0.952	0.959	0.973	0.979	0.985	0.993	0.993	0.995	0.997
14	0.761	0.828	0.868	0.901	0.916	0.940	0.947	0.958	0.968	0.974	0.979	0.984
15	0.688	0.761	0.806	0.854	0.877	0.904	0.921	0.930	0.944	0.954	0.964	0.969
16	0.615	0.700	0.749	0.805	0.835	0.872	0.893	0.904	0.923	0.934	0.948	0.956
17	0.543	0.641	0.697	0.763	0.795	0.838	0.862	0.878	0.903	0.916	0.932	0.942
18	0.480	0.583	0.646	0.718	0.757	0.805	0.835	0.851	0.881	0.897	0.919	0.928
19	0.424	0.535	0.599	0.673	0.718	0.774	0.806	0.825	0.860	0.877	0.902	0.914
20	0.361	0.485	0.551	0.629	0.677	0.738	0.775	0.796	0.836	0.858	0.885	0.900
21	0.309	0.432	0.500	0.585	0.636	0.706	0.746	0.770	0.813	0.837	0.867	0.885
22	0.255	0.373	0.448	0.542	0.600	0.672	0.715	0.743	0.788	0.815	0.850	0.869
23	0.210	0.326	0.396	0.495	0.558	0.634	0.681	0.714	0.763	0.794	0.831	0.853
24	0.165	0.274	0.346	0.451	0.509	0.593	0.646	0.681	0.734	0.766	0.809	0.835
25	0.116	0.216	0.294	0.389	0.453	0.546	0.602	0.640	0.698	0.738	0.782	0.812

Table 6
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 30

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	2.813	2.434	2.181	1.945	1.812	1.643	1.546	1.481	1.387	1.333	1.269	1.229
2	2.517	2.192	1.994	1.812	1.686	1.560	1.470	1.414	1.331	1.291	1.232	1.200
3	2.301	2.035	1.869	1.700	1.594	1.483	1.411	1.366	1.295	1.259	1.208	1.178
4	2.109	1.879	1.747	1.612	1.521	1.422	1.365	1.323	1.262	1.227	1.186	1.159
5	1.932	1.755	1.646	1.531	1.456	1.370	1.321	1.288	1.232	1.202	1.164	1.140
6	1.774	1.648	1.560	1.455	1.394	1.320	1.279	1.250	1.206	1.178	1.145	1.125
7	1.642	1.546	1.479	1.394	1.339	1.276	1.243	1.216	1.179	1.154	1.126	1.109
8	1.527	1.447	1.403	1.326	1.284	1.239	1.207	1.186	1.153	1.132	1.110	1.095
9	1.404	1.362	1.323	1.269	1.235	1.199	1.174	1.155	1.128	1.112	1.094	1.081
10	1.295	1.273	1.247	1.216	1.187	1.158	1.141	1.128	1.104	1.092	1.077	1.066
11	1.194	1.199	1.175	1.156	1.142	1.122	1.112	1.098	1.083	1.073	1.062	1.054
12	1.096	1.114	1.120	1.104	1.099	1.089	1.080	1.070	1.061	1.055	1.047	1.040
13	1.014	1.037	1.049	1.059	1.058	1.052	1.050	1.044	1.041	1.036	1.031	1.027
14	0.933	0.967	0.993	1.008	1.018	1.019	1.021	1.020	1.021	1.018	1.015	1.014
15	0.852	0.905	0.937	0.960	0.975	0.985	0.994	0.994	0.998	1.001	1.001	1.003
16	0.775	0.840	0.884	0.913	0.935	0.955	0.964	0.970	0.979	0.982	0.985	0.991
17	0.697	0.778	0.829	0.869	0.895	0.922	0.937	0.945	0.957	0.966	0.971	0.977
18	0.631	0.724	0.779	0.825	0.860	0.893	0.910	0.921	0.938	0.949	0.957	0.965
19	0.567	0.669	0.733	0.782	0.823	0.861	0.882	0.897	0.917	0.931	0.943	0.953
20	0.506	0.613	0.681	0.745	0.790	0.831	0.855	0.874	0.898	0.913	0.930	0.941
21	0.449	0.563	0.633	0.709	0.750	0.800	0.827	0.851	0.879	0.894	0.915	0.928
22	0.396	0.514	0.586	0.668	0.715	0.768	0.801	0.828	0.860	0.878	0.902	0.916
23	0.345	0.465	0.543	0.629	0.680	0.740	0.776	0.803	0.840	0.859	0.887	0.902
24	0.296	0.422	0.499	0.589	0.647	0.709	0.749	0.778	0.819	0.841	0.873	0.889
25	0.251	0.378	0.458	0.554	0.610	0.679	0.721	0.752	0.799	0.822	0.856	0.876
26	0.210	0.333	0.415	0.516	0.577	0.649	0.693	0.728	0.777	0.805	0.841	0.863
27	0.171	0.292	0.372	0.476	0.540	0.617	0.666	0.703	0.753	0.785	0.824	0.847
28	0.134	0.248	0.332	0.437	0.502	0.583	0.635	0.672	0.730	0.765	0.807	0.830
29	0.100	0.209	0.291	0.395	0.460	0.551	0.604	0.642	0.703	0.739	0.784	0.813
30	0.068	0.161	0.242	0.345	0.416	0.505	0.566	0.604	0.670	0.709	0.759	0.791

Table 7
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 35

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	3.050	2.588	2.326	2.058	1.878	1.708	1.602	1.536	1.428	1.365	1.297	1.256
2	2.721	2.353	2.152	1.914	1.769	1.624	1.533	1.463	1.376	1.326	1.259	1.227
3	2.499	2.175	2.001	1.799	1.685	1.553	1.471	1.415	1.337	1.294	1.233	1.205
4	2.306	2.039	1.884	1.705	1.610	1.492	1.419	1.374	1.303	1.267	1.211	1.186
5	2.128	1.917	1.781	1.624	1.540	1.443	1.374	1.337	1.273	1.238	1.192	1.166
6	1.977	1.801	1.690	1.561	1.473	1.393	1.337	1.300	1.243	1.215	1.174	1.151
7	1.832	1.689	1.595	1.489	1.418	1.349	1.299	1.272	1.218	1.192	1.157	1.136
8	1.703	1.595	1.519	1.423	1.371	1.306	1.263	1.242	1.194	1.171	1.141	1.119
9	1.582	1.499	1.431	1.363	1.318	1.262	1.231	1.211	1.171	1.149	1.123	1.106
10	1.484	1.406	1.362	1.306	1.270	1.225	1.200	1.180	1.149	1.130	1.107	1.094
11	1.374	1.327	1.303	1.254	1.227	1.192	1.168	1.152	1.127	1.111	1.092	1.080
12	1.267	1.251	1.231	1.207	1.182	1.155	1.137	1.127	1.104	1.092	1.078	1.066
13	1.176	1.179	1.171	1.153	1.143	1.122	1.109	1.099	1.084	1.076	1.063	1.054
14	1.086	1.108	1.107	1.101	1.099	1.089	1.083	1.074	1.066	1.057	1.048	1.042
15	0.997	1.039	1.051	1.056	1.064	1.057	1.057	1.048	1.046	1.040	1.034	1.029
16	0.916	0.972	0.993	1.010	1.025	1.024	1.030	1.026	1.026	1.022	1.020	1.017
17	0.840	0.910	0.945	0.967	0.988	0.993	1.000	1.003	1.005	1.005	1.007	1.006
18	0.767	0.849	0.893	0.928	0.947	0.964	0.975	0.977	0.987	0.989	0.992	0.994
19	0.704	0.796	0.838	0.889	0.913	0.935	0.949	0.956	0.968	0.974	0.979	0.983
20	0.638	0.738	0.789	0.846	0.878	0.908	0.923	0.933	0.949	0.956	0.967	0.972
21	0.573	0.684	0.744	0.807	0.844	0.879	0.899	0.911	0.930	0.940	0.953	0.960
22	0.518	0.636	0.699	0.767	0.805	0.847	0.874	0.889	0.912	0.923	0.940	0.949
23	0.464	0.582	0.655	0.730	0.773	0.820	0.851	0.868	0.893	0.907	0.926	0.937
24	0.410	0.532	0.615	0.697	0.742	0.792	0.825	0.846	0.875	0.891	0.913	0.926
25	0.362	0.492	0.571	0.662	0.709	0.766	0.800	0.824	0.857	0.875	0.900	0.914
26	0.319	0.453	0.533	0.625	0.676	0.739	0.775	0.799	0.838	0.858	0.887	0.903
27	0.273	0.411	0.496	0.589	0.645	0.713	0.748	0.779	0.819	0.841	0.872	0.891
28	0.240	0.373	0.459	0.556	0.618	0.684	0.725	0.757	0.801	0.826	0.859	0.878
29	0.202	0.337	0.419	0.520	0.587	0.655	0.702	0.732	0.783	0.809	0.844	0.865
30	0.167	0.300	0.384	0.487	0.554	0.627	0.675	0.709	0.763	0.792	0.830	0.852
31	0.139	0.262	0.348	0.454	0.521	0.596	0.651	0.686	0.742	0.775	0.815	0.839
32	0.108	0.229	0.309	0.418	0.487	0.571	0.625	0.661	0.721	0.757	0.800	0.826
33	0.082	0.195	0.275	0.385	0.454	0.541	0.596	0.635	0.699	0.735	0.783	0.810
34	0.058	0.161	0.238	0.347	0.417	0.508	0.565	0.606	0.673	0.714	0.762	0.794
35	0.038	0.126	0.193	0.302	0.371	0.469	0.527	0.573	0.643	0.686	0.738	0.770

Table 8
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 40

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	3.263	2.741	2.466	2.157	1.976	1.766	1.659	1.573	1.465	1.397	1.320	1.275
2	2.928	2.512	2.279	2.005	1.864	1.671	1.582	1.517	1.415	1.357	1.289	1.247
3	2.699	2.341	2.132	1.903	1.768	1.608	1.529	1.464	1.376	1.323	1.261	1.225
4	2.511	2.184	2.014	1.810	1.693	1.549	1.482	1.424	1.344	1.295	1.241	1.206
5	2.318	2.055	1.905	1.731	1.630	1.504	1.436	1.390	1.314	1.269	1.220	1.189
6	2.172	1.939	1.807	1.658	1.568	1.460	1.393	1.352	1.285	1.245	1.200	1.174
7	2.023	1.833	1.715	1.585	1.505	1.411	1.357	1.318	1.257	1.224	1.182	1.157
8	1.890	1.732	1.630	1.521	1.450	1.371	1.316	1.288	1.235	1.203	1.167	1.144
9	1.761	1.638	1.554	1.460	1.405	1.330	1.286	1.256	1.211	1.181	1.149	1.131
10	1.636	1.548	1.479	1.399	1.356	1.292	1.254	1.227	1.190	1.163	1.135	1.117
11	1.526	1.461	1.404	1.346	1.310	1.255	1.222	1.201	1.168	1.142	1.120	1.104
12	1.429	1.389	1.341	1.292	1.268	1.221	1.191	1.172	1.145	1.126	1.106	1.091
13	1.321	1.311	1.279	1.240	1.221	1.187	1.167	1.150	1.124	1.109	1.090	1.079
14	1.227	1.237	1.219	1.194	1.181	1.153	1.138	1.124	1.104	1.091	1.076	1.067
15	1.137	1.161	1.165	1.150	1.137	1.123	1.108	1.099	1.085	1.074	1.062	1.056
16	1.053	1.092	1.102	1.105	1.095	1.093	1.081	1.077	1.067	1.057	1.049	1.045
17	0.982	1.034	1.044	1.058	1.061	1.060	1.055	1.052	1.048	1.041	1.036	1.033
18	0.904	0.972	0.995	1.019	1.028	1.030	1.029	1.031	1.030	1.026	1.024	1.021
19	0.833	0.912	0.946	0.972	0.989	1.001	1.003	1.008	1.012	1.010	1.010	1.010
20	0.762	0.852	0.900	0.936	0.954	0.971	0.978	0.985	0.993	0.995	0.998	0.998
21	0.693	0.799	0.851	0.899	0.921	0.946	0.956	0.964	0.976	0.981	0.986	0.986
22	0.637	0.746	0.807	0.859	0.882	0.917	0.932	0.942	0.957	0.966	0.972	0.976
23	0.583	0.696	0.760	0.819	0.849	0.893	0.909	0.918	0.940	0.951	0.960	0.966
24	0.526	0.647	0.715	0.783	0.816	0.864	0.886	0.899	0.922	0.935	0.947	0.955
25	0.475	0.600	0.671	0.749	0.786	0.837	0.863	0.879	0.905	0.921	0.935	0.944
26	0.427	0.553	0.631	0.715	0.756	0.810	0.839	0.859	0.887	0.905	0.924	0.934
27	0.378	0.514	0.592	0.683	0.728	0.786	0.818	0.840	0.870	0.890	0.912	0.923
28	0.334	0.473	0.556	0.648	0.696	0.760	0.796	0.820	0.854	0.876	0.898	0.913
29	0.295	0.435	0.517	0.615	0.669	0.738	0.772	0.799	0.838	0.861	0.887	0.902
30	0.258	0.399	0.482	0.586	0.642	0.712	0.750	0.779	0.820	0.847	0.875	0.892
31	0.220	0.360	0.448	0.552	0.613	0.686	0.728	0.757	0.803	0.830	0.861	0.880
32	0.187	0.327	0.415	0.518	0.584	0.662	0.707	0.737	0.786	0.814	0.848	0.870
33	0.156	0.293	0.383	0.488	0.553	0.637	0.683	0.718	0.767	0.799	0.835	0.858
34	0.128	0.263	0.351	0.456	0.525	0.612	0.659	0.697	0.748	0.783	0.821	0.846
35	0.102	0.227	0.320	0.427	0.495	0.587	0.638	0.674	0.729	0.767	0.807	0.833
36	0.080	0.201	0.287	0.398	0.467	0.560	0.615	0.653	0.710	0.750	0.794	0.820
37	0.061	0.174	0.257	0.368	0.438	0.532	0.591	0.631	0.690	0.731	0.778	0.808
38	0.044	0.146	0.227	0.337	0.409	0.503	0.563	0.606	0.670	0.712	0.761	0.793
39	0.028	0.118	0.197	0.302	0.375	0.469	0.533	0.579	0.645	0.692	0.744	0.777
40	0.016	0.088	0.158	0.263	0.336	0.433	0.495	0.542	0.614	0.663	0.720	0.755

Table 9
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 45

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	3.474	2.908	2.584	2.239	2.055	1.828	1.699	1.628	1.509	1.430	1.343	1.295
2	3.121	2.660	2.390	2.100	1.932	1.745	1.628	1.561	1.452	1.386	1.309	1.267
3	2.880	2.489	2.249	1.992	1.842	1.674	1.576	1.512	1.414	1.353	1.286	1.246
4	2.664	2.338	2.128	1.909	1.765	1.619	1.531	1.469	1.378	1.326	1.265	1.228
5	2.480	2.193	2.021	1.811	1.700	1.563	1.488	1.431	1.352	1.300	1.245	1.209
6	2.329	2.073	1.917	1.740	1.637	1.519	1.444	1.396	1.323	1.276	1.224	1.193
7	2.187	1.961	1.832	1.671	1.582	1.473	1.407	1.363	1.297	1.254	1.206	1.179
8	2.048	1.858	1.747	1.606	1.527	1.432	1.375	1.334	1.270	1.234	1.190	1.164
9	1.927	1.765	1.663	1.547	1.474	1.393	1.340	1.305	1.244	1.215	1.175	1.151
10	1.801	1.669	1.586	1.495	1.425	1.356	1.307	1.275	1.224	1.196	1.159	1.139
11	1.690	1.581	1.522	1.439	1.379	1.318	1.277	1.245	1.202	1.177	1.145	1.126
12	1.582	1.503	1.455	1.387	1.335	1.284	1.248	1.218	1.182	1.158	1.131	1.113
13	1.474	1.427	1.387	1.330	1.293	1.248	1.218	1.195	1.161	1.141	1.117	1.102
14	1.381	1.351	1.326	1.284	1.251	1.216	1.189	1.171	1.142	1.124	1.103	1.090
15	1.283	1.278	1.264	1.235	1.211	1.182	1.162	1.149	1.123	1.107	1.090	1.078
16	1.206	1.209	1.207	1.188	1.176	1.153	1.136	1.124	1.103	1.091	1.077	1.068
17	1.122	1.146	1.149	1.142	1.137	1.120	1.108	1.099	1.085	1.076	1.063	1.057
18	1.047	1.083	1.098	1.101	1.101	1.089	1.082	1.078	1.067	1.060	1.053	1.046
19	0.971	1.022	1.040	1.060	1.065	1.060	1.060	1.055	1.050	1.045	1.039	1.035
20	0.901	0.966	0.998	1.019	1.031	1.030	1.034	1.032	1.031	1.030	1.027	1.024
21	0.832	0.908	0.948	0.981	0.997	1.005	1.010	1.012	1.014	1.016	1.015	1.014
22	0.759	0.853	0.901	0.944	0.961	0.976	0.988	0.990	0.998	1.000	1.003	1.002
23	0.703	0.801	0.855	0.907	0.929	0.950	0.964	0.971	0.981	0.986	0.991	0.992
24	0.642	0.754	0.814	0.870	0.899	0.926	0.940	0.950	0.964	0.972	0.978	0.982
25	0.581	0.707	0.770	0.834	0.865	0.900	0.919	0.930	0.949	0.957	0.968	0.971
26	0.528	0.656	0.729	0.801	0.834	0.873	0.897	0.912	0.933	0.943	0.956	0.962
27	0.478	0.614	0.689	0.769	0.803	0.850	0.874	0.892	0.917	0.929	0.945	0.951
28	0.424	0.572	0.652	0.734	0.775	0.825	0.853	0.871	0.900	0.914	0.932	0.941
29	0.385	0.532	0.613	0.699	0.747	0.800	0.832	0.852	0.884	0.900	0.919	0.931
30	0.339	0.492	0.575	0.667	0.718	0.777	0.810	0.834	0.868	0.885	0.908	0.922
31	0.299	0.456	0.541	0.636	0.690	0.754	0.789	0.815	0.851	0.872	0.897	0.912
32	0.262	0.418	0.509	0.605	0.663	0.730	0.770	0.796	0.837	0.858	0.887	0.902
33	0.232	0.384	0.475	0.576	0.635	0.706	0.749	0.777	0.820	0.844	0.874	0.892
34	0.198	0.348	0.444	0.543	0.610	0.683	0.726	0.756	0.804	0.830	0.862	0.881
35	0.168	0.319	0.409	0.516	0.584	0.660	0.706	0.738	0.788	0.816	0.849	0.872
36	0.143	0.286	0.380	0.487	0.558	0.637	0.684	0.720	0.772	0.803	0.839	0.861
37	0.118	0.260	0.351	0.459	0.531	0.614	0.665	0.701	0.753	0.789	0.827	0.849
38	0.096	0.232	0.319	0.432	0.505	0.590	0.645	0.682	0.736	0.773	0.813	0.838
39	0.076	0.206	0.292	0.406	0.478	0.569	0.623	0.662	0.720	0.758	0.801	0.827
40	0.059	0.180	0.261	0.380	0.453	0.545	0.602	0.641	0.704	0.742	0.788	0.815

Table 9 (cont.)

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
41	0.043	0.153	0.233	0.351	0.427	0.522	0.578	0.622	0.685	0.726	0.774	0.804
42	0.030	0.132	0.210	0.324	0.400	0.497	0.556	0.599	0.664	0.707	0.759	0.791
43	0.019	0.109	0.185	0.293	0.373	0.469	0.532	0.576	0.643	0.690	0.741	0.776
44	0.010	0.086	0.158	0.264	0.343	0.439	0.504	0.547	0.619	0.668	0.724	0.760
45	0.004	0.064	0.126	0.227	0.304	0.399	0.471	0.514	0.588	0.643	0.701	0.739

Table 10
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 50

	N =											
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
1	3.675	3.044	2.712	2.331	2.124	1.885	1.759	1.678	1.533	1.457	1.368	1.313
2	3.336	2.804	2.500	2.188	2.015	1.802	1.685	1.612	1.483	1.415	1.332	1.284
3	3.073	2.621	2.372	2.079	1.928	1.730	1.628	1.558	1.445	1.382	1.308	1.263
4	2.861	2.474	2.247	1.985	1.849	1.675	1.579	1.511	1.412	1.352	1.286	1.245
5	2.665	2.334	2.137	1.905	1.776	1.623	1.531	1.476	1.382	1.327	1.267	1.228
6	2.493	2.216	2.036	1.829	1.714	1.573	1.492	1.442	1.355	1.305	1.249	1.212
7	2.355	2.092	1.941	1.755	1.655	1.526	1.455	1.408	1.330	1.283	1.232	1.199
8	2.212	1.997	1.852	1.690	1.607	1.482	1.420	1.373	1.305	1.262	1.216	1.185
9	2.083	1.888	1.771	1.627	1.552	1.445	1.385	1.345	1.281	1.241	1.201	1.173
10	1.968	1.794	1.695	1.570	1.501	1.406	1.353	1.316	1.257	1.222	1.186	1.160
11	1.844	1.703	1.622	1.520	1.455	1.370	1.324	1.291	1.237	1.205	1.172	1.146
12	1.737	1.621	1.556	1.465	1.409	1.339	1.295	1.263	1.215	1.187	1.155	1.134
13	1.639	1.546	1.490	1.416	1.368	1.304	1.268	1.238	1.197	1.171	1.142	1.123
14	1.528	1.465	1.428	1.368	1.322	1.273	1.236	1.216	1.178	1.154	1.128	1.111
15	1.433	1.394	1.365	1.319	1.279	1.239	1.209	1.191	1.157	1.139	1.116	1.100
16	1.347	1.322	1.304	1.270	1.241	1.209	1.185	1.168	1.139	1.122	1.103	1.088
17	1.252	1.263	1.249	1.226	1.204	1.178	1.160	1.146	1.121	1.107	1.091	1.078
18	1.169	1.197	1.196	1.185	1.165	1.150	1.134	1.125	1.103	1.091	1.077	1.067
19	1.081	1.134	1.146	1.139	1.130	1.121	1.110	1.101	1.087	1.077	1.065	1.057
20	1.006	1.074	1.096	1.099	1.090	1.091	1.086	1.079	1.068	1.062	1.053	1.046
21	0.938	1.015	1.045	1.060	1.057	1.067	1.063	1.060	1.051	1.048	1.040	1.036
22	0.873	0.962	0.998	1.020	1.025	1.039	1.041	1.039	1.037	1.033	1.029	1.026
23	0.797	0.907	0.948	0.984	0.997	1.012	1.018	1.017	1.021	1.020	1.017	1.016
24	0.740	0.852	0.899	0.949	0.966	0.988	0.994	0.997	1.005	1.006	1.006	1.006
25	0.680	0.803	0.854	0.912	0.936	0.961	0.971	0.977	0.988	0.993	0.994	0.997
26	0.622	0.757	0.813	0.877	0.903	0.935	0.949	0.957	0.972	0.979	0.982	0.987
27	0.571	0.711	0.774	0.843	0.874	0.912	0.928	0.938	0.956	0.966	0.971	0.977
28	0.521	0.668	0.736	0.808	0.846	0.887	0.910	0.918	0.941	0.952	0.960	0.968

Table 10 (cont.)

	N =										
Root	50	75	100	150	200	300	400	500	750	1000	1500	2000
29	0.474	0.623	0.697	0.776	0.815	0.863	0.886	0.900	0.926	0.937	0.949	0.959
30	0.432	0.578	0.660	0.744	0.787	0.839	0.867	0.881	0.909	0.923	0.939	0.948
31	0.385	0.539	0.620	0.714	0.761	0.816	0.845	0.863	0.894	0.910	0.929	0.939
32	0.343	0.501	0.584	0.683	0.732	0.793	0.826	0.844	0.879	0.897	0.918	0.930
33	0.307	0.465	0.552	0.655	0.707	0.770	0.806	0.827	0.864	0.883	0.907	0.920
34	0.273	0.433	0.519	0.624	0.681	0.747	0.785	0.809	0.848	0.869	0.896	0.910
35	0.237	0.398	0.490	0.596	0.654	0.725	0.765	0.791	0.834	0.856	0.885	0.901
36	0.207	0.365	0.458	0.567	0.631	0.705	0.746	0.773	0.819	0.843	0.873	0.892
37	0.176	0.337	0.430	0.539	0.606	0.683	0.726	0.756	0.804	0.831	0.861	0.882
38	0.149	0.307	0.401	0.513	0.582	0.660	0.706	0.738	0.790	0.817	0.851	0.873
39	0.125	0.278	0.374	0.486	0.557	0.639	0.687	0.721	0.774	0.804	0.839	0.863
40	0.106	0.248	0.345	0.463	0.531	0.617	0.667	0.702	0.757	0.792	0.829	0.853
41	0.085	0.223	0.320	0.437	0.508	0.594	0.648	0.684	0.742	0.778	0.818	0.842
42	0.066	0.198	0.292	0.412	0.483	0.572	0.628	0.666	0.726	0.763	0.806	0.832
43	0.050	0.175	0.269	0.388	0.457	0.551	0.608	0.647	0.709	0.749	0.794	0.822
44	0.036	0.153	0.244	0.361	0.435	0.529	0.588	0.628	0.694	0.734	0.781	0.811
45	0.024	0.132	0.219	0.334	0.410	0.507	0.568	0.609	0.678	0.720	0.768	0.800
46	0.015	0.114	0.195	0.311	0.387	0.485	0.547	0.589	0.660	0.704	0.754	0.788
47	0.008	0.095	0.174	0.286	0.364	0.462	0.523	0.570	0.642	0.687	0.741	0.776
48	0.003	0.077	0.149	0.262	0.337	0.437	0.500	0.549	0.622	0.670	0.725	0.761
49	0.001	0.059	0.127	0.232	0.309	0.409	0.475	0.524	0.601	0.650	0.708	0.748
50	0.000	0.043	0.100	0.197	0.275	0.374	0.442	0.491	0.576	0.625	0.685	0.727

Table 11
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 60

	N =									
Root	100	150	200	300	400	500	750	1000	1500	2000
1	2.939	2.508	2.274	2.006	1.845	1.743	1.599	1.508	1.410	1.353
2	2.722	2.359	2.149	1.918	1.772	1.676	1.547	1.467	1.378	1.323
3	2.579	2.245	2.062	1.841	1.714	1.629	1.507	1.434	1.353	1.302
4	2.456	2.154	1.982	1.779	1.668	1.589	1.474	1.406	1.330	1.284
5	2.349	2.061	1.910	1.727	1.623	1.551	1.446	1.382	1.310	1.267
6	2.249	1.985	1.849	1.682	1.583	1.518	1.418	1.359	1.292	1.251
7	2.146	1.920	1.792	1.635	1.546	1.487	1.392	1.338	1.274	1.238
8	2.061	1.859	1.736	1.594	1.510	1.455	1.371	1.318	1.259	1.224
9	1.973	1.797	1.682	1.552	1.478	1.425	1.347	1.299	1.244	1.210
10	1.896	1.739	1.634	1.512	1.449	1.399	1.326	1.280	1.228	1.196
11	1.822	1.678	1.589	1.476	1.417	1.372	1.305	1.264	1.214	1.184

Table 11 (cont.)

					N =					
Root	100	150	200	300	400	500	750	1000	1500	2000
12	1.756	1.626	1.541	1.445	1.387	1.346	1.282	1.246	1.201	1.173
13	1.685	1.568	1.497	1.411	1.358	1.319	1.262	1.230	1.187	1.163
14	1.617	1.517	1.453	1.376	1.330	1.296	1.244	1.214	1.174	1.150
15	1.552	1.473	1.411	1.343	1.303	1.273	1.225	1.198	1.160	1.141
16	1.491	1.426	1.372	1.313	1.277	1.249	1.206	1.181	1.148	1.130
17	1.434	1.372	1.334	1.285	1.249	1.226	1.187	1.165	1.136	1.118
18	1.378	1.330	1.298	1.255	1.226	1.203	1.171	1.149	1.125	1.109
19	1.319	1.285	1.261	1.228	1.203	1.182	1.153	1.135	1.112	1.098
20	1.266	1.250	1.225	1.197	1.178	1.160	1.137	1.121	1.100	1.088
21	1.216	1.208	1.190	1.168	1.155	1.142	1.120	1.107	1.089	1.079
22	1.167	1.167	1.157	1.142	1.131	1.123	1.104	1.093	1.078	1.068
23	1.121	1.129	1.122	1.118	1.109	1.103	1.088	1.079	1.067	1.060
24	1.074	1.091	1.093	1.093	1.087	1.083	1.071	1.065	1.057	1.050
25	1.028	1.055	1.063	1.068	1.064	1.064	1.055	1.051	1.046	1.041
26	0.982	1.020	1.032	1.040	1.043	1.044	1.040	1.039	1.035	1.031
27	0.939	0.983	1.002	1.016	1.023	1.026	1.026	1.025	1.024	1.022
28	0.900	0.949	0.972	0.994	1.004	1.005	1.011	1.012	1.013	1.013
29	0.862	0.917	0.943	0.970	0.983	0.988	0.996	1.000	1.002	1.005
30	0.823	0.885	0.917	0.946	0.964	0.970	0.981	0.988	0.992	0.995
31	0.783	0.853	0.888	0.923	0.942	0.952	0.968	0.975	0.982	0.986
32	0.746	0.824	0.861	0.902	0.923	0.935	0.953	0.963	0.972	0.977
33	0.713	0.792	0.835	0.879	0.904	0.919	0.938	0.950	0.961	0.969
34	0.679	0.763	0.808	0.857	0.884	0.901	0.924	0.938	0.951	0.960
35	0.645	0.733	0.785	0.837	0.865	0.885	0.910	0.926	0.942	0.952
36	0.613	0.705	0.759	0.815	0.847	0.868	0.896	0.914	0.932	0.943
37	0.579	0.679	0.735	0.795	0.828	0.851	0.882	0.901	0.921	0.934
38	0.550	0.652	0.710	0.775	0.810	0.834	0.869	0.889	0.912	0.925
39	0.519	0.625	0.686	0.752	0.792	0.818	0.856	0.877	0.901	0.916
40	0.492	0.599	0.665	0.734	0.775	0.801	0.841	0.866	0.892	0.907
41	0.464	0.574	0.641	0.713	0.758	0.786	0.828	0.853	0.881	0.898
42	0.436	0.549	0.619	0.694	0.740	0.769	0.815	0.841	0.871	0.889
43	0.407	0.524	0.599	0.675	0.722	0.754	0.801	0.829	0.862	0.880
44	0.382	0.501	0.574	0.654	0.704	0.736	0.789	0.816	0.852	0.872
45	0.355	0.478	0.552	0.636	0.686	0.722	0.774	0.804	0.841	0.863
46	0.332	0.455	0.531	0.618	0.669	0.706	0.760	0.792	0.831	0.854
47	0.307	0.433	0.510	0.599	0.652	0.689	0.748	0.780	0.821	0.844
48	0.285	0.411	0.490	0.579	0.634	0.674	0.733	0.768	0.811	0.836
49	0.264	0.389	0.469	0.561	0.617	0.658	0.720	0.755	0.800	0.827
50	0.243	0.369	0.449	0.541	0.599	0.640	0.706	0.743	0.789	0.818
51	0.222	0.347	0.427	0.523	0.582	0.626	0.691	0.731	0.779	0.808
52	0.202	0.327	0.404	0.504	0.563	0.609	0.678	0.718	0.767	0.797
53	0.185	0.306	0.385	0.486	0.546	0.594	0.662	0.705	0.756	0.787
54	0.163	0.285	0.362	0.466	0.528	0.576	0.648	0.692	0.745	0.778

Table 11 (cont.)

Root	N = 100	150	200	300	400	500	750	1000	1500	2000
55	0.147	0.265	0.340	0.448	0.510	0.559	0.631	0.677	0.733	0.767
56	0.131	0.244	0.320	0.427	0.490	0.540	0.615	0.663	0.721	0.755
57	0.115	0.224	0.300	0.408	0.471	0.522	0.599	0.647	0.708	0.743
58	0.098	0.203	0.276	0.384	0.451	0.500	0.582	0.632	0.694	0.731
59	0.080	0.182	0.254	0.361	0.431	0.480	0.561	0.612	0.678	0.718
60	0.062	0.153	0.226	0.327	0.398	0.454	0.535	0.589	0.657	0.699

Table 12
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 70

Root	N = 100	150	200	300	400	500	750	1000	1500	2000
1	3.159	2.684	2.406	2.117	1.941	1.829	1.650	1.560	1.450	1.382
2	2.968	2.527	2.283	2.016	1.864	1.763	1.605	1.519	1.415	1.356
3	2.798	2.413	2.191	1.948	1.805	1.712	1.567	1.486	1.389	1.336
4	2.671	2.322	2.114	1.888	1.754	1.670	1.538	1.457	1.369	1.317
5	2.545	2.228	2.044	1.837	1.712	1.631	1.503	1.432	1.349	1.301
6	2.437	2.152	1.980	1.789	1.668	1.594	1.478	1.410	1.333	1.284
7	2.342	2.082	1.916	1.742	1.633	1.561	1.454	1.388	1.316	1.270
8	2.260	2.010	1.860	1.700	1.594	1.533	1.429	1.369	1.299	1.257
9	2.168	1.944	1.804	1.659	1.560	1.507	1.405	1.349	1.283	1.245
10	2.089	1.883	1.754	1.616	1.529	1.476	1.385	1.330	1.268	1.232
11	2.018	1.830	1.710	1.579	1.496	1.449	1.365	1.314	1.254	1.221
12	1.940	1.770	1.664	1.541	1.468	1.421	1.344	1.296	1.240	1.210
13	1.871	1.714	1.615	1.511	1.441	1.396	1.324	1.279	1.228	1.198
14	1.801	1.666	1.575	1.474	1.414	1.373	1.305	1.263	1.217	1.186
15	1.731	1.615	1.536	1.443	1.389	1.349	1.286	1.246	1.203	1.176
16	1.672	1.570	1.500	1.412	1.363	1.326	1.267	1.231	1.191	1.166
17	1.610	1.522	1.462	1.379	1.336	1.303	1.249	1.215	1.178	1.155
18	1.551	1.475	1.423	1.348	1.311	1.281	1.233	1.200	1.168	1.145
19	1.496	1.431	1.386	1.323	1.286	1.259	1.216	1.186	1.157	1.135
20	1.439	1.388	1.350	1.294	1.263	1.238	1.198	1.173	1.145	1.126
21	1.385	1.345	1.315	1.266	1.240	1.217	1.183	1.159	1.134	1.116
22	1.337	1.306	1.281	1.237	1.215	1.197	1.167	1.145	1.123	1.106
23	1.286	1.268	1.249	1.212	1.193	1.176	1.153	1.132	1.112	1.097
24	1.236	1.227	1.213	1.188	1.171	1.157	1.136	1.119	1.101	1.088
25	1.190	1.189	1.181	1.165	1.149	1.139	1.121	1.106	1.090	1.079
26	1.140	1.151	1.151	1.139	1.128	1.122	1.105	1.093	1.080	1.070

Table 12 (cont.)

Root	N = 100	150	200	300	400	500	750	1000	1500	2000
27	1.094	1.117	1.117	1.115	1.108	1.102	1.090	1.081	1.069	1.061
28	1.051	1.081	1.089	1.092	1.087	1.083	1.074	1.068	1.059	1.053
29	1.013	1.047	1.059	1.069	1.067	1.067	1.062	1.055	1.049	1.044
30	0.971	1.014	1.032	1.046	1.049	1.048	1.046	1.043	1.040	1.035
31	0.932	0.984	1.004	1.021	1.028	1.031	1.032	1.031	1.030	1.027
32	0.894	0.954	0.981	0.999	1.009	1.015	1.019	1.019	1.019	1.019
33	0.856	0.924	0.952	0.979	0.989	0.997	1.005	1.007	1.009	1.010
34	0.820	0.891	0.927	0.956	0.970	0.979	0.991	0.995	1.000	1.002
35	0.783	0.861	0.903	0.934	0.952	0.963	0.976	0.984	0.991	0.994
36	0.752	0.833	0.875	0.915	0.934	0.947	0.964	0.971	0.980	0.985
37	0.718	0.805	0.850	0.894	0.917	0.931	0.950	0.960	0.970	0.977
38	0.682	0.777	0.826	0.874	0.900	0.916	0.936	0.949	0.961	0.969
39	0.650	0.750	0.801	0.854	0.882	0.898	0.925	0.938	0.951	0.961
40	0.620	0.721	0.777	0.833	0.863	0.881	0.911	0.927	0.943	0.953
41	0.587	0.696	0.753	0.814	0.845	0.866	0.899	0.915	0.933	0.944
42	0.561	0.669	0.732	0.795	0.829	0.851	0.884	0.904	0.924	0.936
43	0.531	0.644	0.708	0.776	0.812	0.836	0.871	0.893	0.914	0.928
44	0.505	0.620	0.686	0.757	0.795	0.820	0.859	0.881	0.906	0.919
45	0.478	0.597	0.663	0.737	0.779	0.804	0.847	0.871	0.896	0.911
46	0.453	0.574	0.642	0.719	0.762	0.790	0.833	0.860	0.887	0.903
47	0.429	0.549	0.620	0.700	0.746	0.775	0.821	0.849	0.877	0.895
48	0.405	0.527	0.600	0.683	0.730	0.760	0.809	0.838	0.868	0.887
49	0.380	0.506	0.579	0.665	0.714	0.746	0.797	0.828	0.859	0.880
50	0.357	0.483	0.559	0.646	0.699	0.731	0.784	0.816	0.849	0.872
51	0.335	0.462	0.539	0.629	0.683	0.716	0.772	0.805	0.840	0.863
52	0.313	0.441	0.520	0.611	0.667	0.701	0.760	0.794	0.831	0.855
53	0.291	0.420	0.498	0.594	0.652	0.686	0.747	0.783	0.823	0.846
54	0.272	0.399	0.481	0.577	0.637	0.673	0.734	0.772	0.813	0.839
55	0.251	0.381	0.463	0.560	0.622	0.659	0.722	0.761	0.803	0.831
56	0.234	0.361	0.446	0.544	0.605	0.644	0.709	0.750	0.794	0.822
57	0.216	0.341	0.427	0.527	0.589	0.629	0.696	0.738	0.785	0.813
58	0.197	0.324	0.410	0.509	0.573	0.614	0.684	0.727	0.776	0.804
59	0.180	0.304	0.393	0.492	0.556	0.600	0.671	0.715	0.766	0.796
60	0.163	0.287	0.376	0.475	0.541	0.585	0.659	0.704	0.756	0.788
61	0.148	0.270	0.357	0.460	0.526	0.570	0.646	0.692	0.746	0.779
62	0.134	0.252	0.340	0.442	0.510	0.555	0.632	0.679	0.736	0.770
63	0.121	0.236	0.321	0.424	0.495	0.541	0.618	0.667	0.725	0.760
64	0.106	0.220	0.304	0.408	0.478	0.526	0.604	0.655	0.714	0.750
65	0.093	0.203	0.287	0.391	0.459	0.510	0.589	0.642	0.703	0.741
66	0.080	0.187	0.268	0.374	0.442	0.492	0.574	0.628	0.691	0.730
67	0.068	0.171	0.250	0.353	0.425	0.475	0.559	0.614	0.679	0.719
68	0.057	0.155	0.231	0.335	0.406	0.457	0.544	0.597	0.666	0.706
69	0.046	0.135	0.209	0.313	0.384	0.436	0.524	0.581	0.649	0.693
70	0.034	0.114	0.181	0.287	0.357	0.412	0.503	0.557	0.631	0.677

Table 13
Average Eigenvalues of Random Correlation Matrices Assuming Independence for P = 80

Root	N = 100	150	200	300	400	500	750	1000	1500	2000
1	3.386	2.839	2.533	2.215	2.021	1.887	1.713	1.610	1.484	1.415
2	3.165	2.688	2.396	2.118	1.945	1.835	1.661	1.565	1.451	1.389
3	3.009	2.561	2.309	2.045	1.887	1.782	1.622	1.531	1.425	1.369
4	2.873	2.462	2.237	1.981	1.833	1.742	1.585	1.506	1.405	1.349
5	2.756	2.376	2.168	1.928	1.791	1.703	1.558	1.479	1.386	1.332
6	2.645	2.300	2.104	1.880	1.747	1.665	1.534	1.458	1.368	1.318
7	2.543	2.221	2.046	1.834	1.711	1.629	1.508	1.435	1.351	1.303
8	2.443	2.158	1.988	1.790	1.679	1.600	1.484	1.416	1.337	1.290
9	2.356	2.088	1.935	1.750	1.643	1.572	1.461	1.397	1.323	1.278
10	2.277	2.026	1.886	1.710	1.612	1.543	1.439	1.379	1.307	1.266
11	2.200	1.967	1.836	1.672	1.579	1.515	1.418	1.361	1.295	1.255
12	2.120	1.908	1.784	1.637	1.550	1.488	1.399	1.345	1.281	1.243
13	2.042	1.858	1.743	1.603	1.523	1.465	1.378	1.328	1.267	1.232
14	1.972	1.808	1.696	1.572	1.496	1.441	1.361	1.312	1.255	1.221
15	1.908	1.754	1.653	1.536	1.469	1.420	1.342	1.297	1.242	1.210
16	1.835	1.701	1.612	1.508	1.441	1.395	1.325	1.281	1.229	1.200
17	1.775	1.654	1.573	1.474	1.417	1.374	1.307	1.266	1.218	1.190
18	1.720	1.608	1.533	1.448	1.395	1.351	1.289	1.251	1.206	1.180
19	1.658	1.562	1.498	1.417	1.369	1.330	1.272	1.237	1.195	1.170
20	1.597	1.515	1.462	1.390	1.345	1.308	1.256	1.223	1.184	1.160
21	1.543	1.474	1.427	1.363	1.320	1.290	1.242	1.211	1.174	1.151
22	1.489	1.435	1.394	1.334	1.298	1.269	1.225	1.197	1.162	1.142
23	1.438	1.396	1.359	1.308	1.275	1.250	1.209	1.183	1.151	1.133
24	1.388	1.359	1.326	1.279	1.253	1.230	1.192	1.170	1.142	1.124
25	1.339	1.320	1.293	1.254	1.233	1.210	1.177	1.156	1.131	1.114
26	1.297	1.282	1.265	1.229	1.210	1.192	1.161	1.143	1.120	1.106
27	1.249	1.244	1.232	1.202	1.188	1.174	1.146	1.130	1.110	1.097
28	1.200	1.209	1.201	1.179	1.168	1.154	1.132	1.118	1.101	1.088
29	1.157	1.175	1.170	1.155	1.149	1.136	1.119	1.106	1.090	1.080
30	1.114	1.140	1.142	1.133	1.130	1.118	1.105	1.094	1.080	1.071
31	1.069	1.104	1.114	1.111	1.110	1.102	1.092	1.082	1.070	1.063
32	1.033	1.073	1.086	1.087	1.091	1.084	1.078	1.070	1.060	1.055
33	0.990	1.042	1.060	1.066	1.071	1.068	1.062	1.058	1.051	1.047
34	0.956	1.012	1.032	1.046	1.051	1.052	1.049	1.046	1.042	1.039
35	0.920	0.979	1.008	1.025	1.033	1.036	1.036	1.035	1.033	1.030
36	0.884	0.949	0.980	1.003	1.014	1.020	1.023	1.024	1.023	1.023
37	0.846	0.922	0.956	0.983	0.996	1.004	1.010	1.012	1.014	1.014
38	0.814	0.893	0.930	0.964	0.977	0.986	0.998	1.002	1.005	1.007
39	0.780	0.866	0.904	0.945	0.961	0.970	0.985	0.991	0.996	0.999
40	0.747	0.838	0.881	0.923	0.944	0.956	0.971	0.980	0.987	0.991
41	0.715	0.811	0.858	0.903	0.926	0.939	0.959	0.968	0.979	0.983

Table 13 (cont.)

					N =					
Root	100	150	200	300	400	500	750	1000	1500	2000
42	0.685	0.786	0.835	0.885	0.911	0.925	0.947	0.957	0.969	0.976
43	0.656	0.758	0.810	0.867	0.893	0.910	0.935	0.947	0.961	0.967
44	0.625	0.735	0.789	0.847	0.877	0.895	0.923	0.936	0.953	0.959
45	0.597	0.708	0.766	0.826	0.860	0.880	0.911	0.926	0.943	0.952
46	0.568	0.683	0.745	0.809	0.845	0.866	0.899	0.916	0.935	0.945
47	0.542	0.659	0.724	0.790	0.828	0.852	0.887	0.906	0.925	0.937
48	0.516	0.637	0.702	0.771	0.812	0.837	0.875	0.895	0.917	0.929
49	0.490	0.613	0.684	0.753	0.796	0.823	0.863	0.885	0.909	0.922
50	0.465	0.590	0.664	0.737	0.780	0.809	0.851	0.874	0.900	0.915
51	0.439	0.569	0.643	0.721	0.764	0.796	0.839	0.864	0.892	0.907
52	0.416	0.548	0.622	0.704	0.749	0.781	0.828	0.853	0.883	0.899
53	0.393	0.528	0.602	0.688	0.734	0.766	0.815	0.842	0.875	0.893
54	0.371	0.506	0.583	0.671	0.718	0.754	0.804	0.832	0.866	0.885
55	0.350	0.487	0.564	0.655	0.703	0.740	0.793	0.822	0.857	0.877
56	0.329	0.467	0.546	0.638	0.688	0.727	0.781	0.812	0.849	0.869
57	0.309	0.446	0.529	0.621	0.674	0.714	0.769	0.801	0.839	0.862
58	0.289	0.428	0.511	0.607	0.661	0.700	0.758	0.791	0.831	0.854
59	0.270	0.410	0.492	0.590	0.645	0.687	0.746	0.782	0.823	0.847
60	0.252	0.391	0.474	0.574	0.632	0.673	0.736	0.771	0.814	0.839
61	0.234	0.373	0.457	0.560	0.618	0.660	0.725	0.760	0.806	0.832
62	0.217	0.356	0.439	0.544	0.603	0.646	0.712	0.751	0.797	0.824
63	0.202	0.338	0.424	0.529	0.588	0.633	0.701	0.740	0.788	0.816
64	0.187	0.322	0.408	0.514	0.574	0.619	0.689	0.730	0.780	0.808
65	0.173	0.305	0.391	0.497	0.560	0.605	0.678	0.719	0.771	0.801
66	0.159	0.289	0.376	0.481	0.546	0.593	0.666	0.709	0.762	0.793
67	0.145	0.273	0.359	0.468	0.533	0.581	0.654	0.699	0.753	0.785
68	0.130	0.258	0.343	0.452	0.518	0.566	0.642	0.688	0.744	0.777
69	0.118	0.242	0.328	0.437	0.504	0.554	0.631	0.677	0.735	0.769
70	0.106	0.226	0.312	0.423	0.488	0.540	0.619	0.666	0.726	0.760
71	0.094	0.212	0.296	0.407	0.475	0.526	0.606	0.656	0.715	0.752
72	0.084	0.197	0.281	0.392	0.460	0.512	0.593	0.644	0.706	0.743
73	0.073	0.183	0.266	0.377	0.447	0.499	0.581	0.633	0.696	0.734
74	0.063	0.169	0.252	0.362	0.432	0.484	0.568	0.622	0.686	0.724
75	0.054	0.156	0.235	0.346	0.416	0.469	0.555	0.610	0.675	0.715
76	0.045	0.143	0.221	0.329	0.399	0.454	0.541	0.597	0.664	0.705
77	0.038	0.130	0.205	0.312	0.383	0.437	0.527	0.584	0.652	0.694
78	0.030	0.116	0.189	0.293	0.366	0.419	0.511	0.569	0.639	0.683
79	0.022	0.102	0.170	0.273	0.348	0.401	0.491	0.552	0.625	0.670
80	0.015	0.085	0.150	0.250	0.321	0.378	0.470	0.532	0.607	0.654

Relative Accuracy of Interpolation and Regression Estimates

It was hypothesized that simple linear interpolation of non-tabled N, P combinations may prove sufficiently accurate for deriving PA criteria, and this hypothesis was tested in the sections below. The accuracy of the various estimation methods was assessed by computing RMSEs for each method based on deviations from averaged eigenvalues of the same order obtained from Monte Carlo simulation results to cover each of the interpolation cases described below. Several specific examples and a number of general tests of the accuracy of the interpolation method were provided relative to the *AH* and *LLF* regression equation estimates.

Examples of Linear Interpolation.

The first interpolation case concerned a value for P which was covered by one of the tables, for a specific sample size which fell within those covered though it specifically was not. For example, let $P = 25$ and $N = 890$. In this case the estimated first eigenvalue for random data was interpolated from the tables as 1.319. The *AH* estimated value in this case was 1.323 and the *LLF* estimated value was 1.341. Subsequent eigenvalues were interpolated and regression estimates were obtained as well. These values could have been used to serve as PA criteria. A Monte Carlo simulation of 100 replications was conducted for this case ($P = 25$ and $N = 890$) and the results served as a baseline for computing RMSE values for the sets of eigenvalues estimated using linear interpolation, *AH* estimates and *LLF* estimates. Since the regression methods could only estimate the first 23 eigenvalues, only the first 23 values were used in all RMSE calculations. The RMSEvalues obtained were 0.0019, 0.3069 and 0.1477, respectively. It should be noted in this case that both regression estimation methods produced 21 estimated eigenvalues greater than 1.0, and some of the latter *AH* estimates were greater than preceding estimates. The interpolation method was superior in this case.

A second interpolation case involved a value for P which was not covered by the tables, for a specific sample size that was covered. For example, when $P = 17$ and $N = 75$ the linear interpolation estimate for the first eigenvalue for random data was 1.937. This value was closer to the *LLF* regression estimated value of 1.914 than it was to the *AH* value of 2.003. Another Monte Carlo simulation involving 100 random samples for this particular case was used to develop expected eigenvalues. The RMSEs for the sets of eigenvalues obtained from linear interpolation, *AH* estimates and *LLF* estimates were 0.0109, 0.0381 and 0.0264, respectively. Once again linear interpolation was more accurate, though the advantage was not as great as in the previous case.

A third interpolation case involved values for both N and P which were not specifically covered by the tables, but did fall in the ranges of both parameters. For example, when $P = 37$ and $N = 177$ the interpolation of the first eigenvalue for random data yields 2.000. In this case the *AH* regression estimated value was 2.037 and the value was 1.938 for the *LLF* method. Based on an independent Monte Carlo simulation of 100 replications for this case, the RMSEs for the sets of eigenvalues obtained by linear interpolation, *AH* estimates and *LLF* estimates were 0.0055, 0.1682 and 0.2228, respectively. Again, linear interpolation proved more accurate than either of the regression estimation methods.

A More General Comparison of the Accuracy of Interpolations.

Although the foregoing examples suggested that linear interpolation was relatively more accurate in a limited scope, it was important that the accuracy of this method be more definitively examined across a wider range of conditions. To this end the following additional comparisons were made. Since each of the first ten tables presents 10 separate conditions (i.e., values of N covered by the regression estimation methods) for a fixed value of P, it was possible to "interpolate" eigenvalues for the middle column of each group of three adjacent columns. For example, in the table covering averaged eigenvalues where $P = 5$, it was possible to interpolate values for $N = 75$ by using the columns for $N = 50$ and $N = 100$. The criterion for accuracy was then the respective column that has been interpolated. This resulted in a total of eighty sets of interpolations on values of N produced by interpolating for each of the middle eight columns (for the first 10 values of N) in each of the first ten tables.

The benefit of interpolating for existing columns in the tables was that it did not require any further Monte Carlo simulations. A very significant drawback of this procedure was that the relative accuracy of the linear interpolation method was placed at a distinct disadvantage, as interpolations were made over ranges of values for N much broader than necessary if the tables were used in actual practice for interpolation. For the example given above where $P = 5$, there would be no need to interpolate for $N = 75$ in actual applications, and interpolations could be made between $N = 50$ and $N = 75$ and also between $N = 75$ and $N = 100$.

Regression estimates for both the *AH* and *LLF* equations were made for each of these conditions, and RMSEs were examined for the interpolations and regression estimates. In 71 out of the 80 cases the interpolated values produced the smallest RMSEs. A comparison of the relative sizes of the RMSEs for each method is presented in Table 14. Eighty-four percent of the interpolation RMSEs were less than or equal to .050, while only 34% or less of the regression estimates were as accurate. These results indicated that the interpolations were generally superior to either regression method.

Table 14
Ranges of RMSE Values for Interpolations on N Within Tables

Method of Eigenvalue Estimation	RMSE Range						
	≤.01	.01-.02	.02-.05	.05-.10	.10-.20	.20-.50	>.5
Linear Interpolation	34	20	13	12	1	0	0
LL Equations	4	8	15	19	17	10	7
AH Equations	0	1	16	16	15	17	15

For the 9 cases where a regression method performed better than the interpolation method, it was always the *LLF* method. In all but two of these cases $N \leq 100$, but there was no other discernable pattern to these nine cases. The largest discrepancy resulted from a case where the *LLF* RMSE was .027 and the interpolation RMSE was .089. The average improvement in accuracy over these 9 cases was .019. The *AH* method never did better than linear interpolation. Focusing on the two regression methods only, the *LLF* method was better in 54 cases.

It was also possible to conduct a series of interpolations for values of P across tables, for fixed values of N. For example, using tables for $P = 5$ and $P = 15$ it was possible to interpolate eigenvalues for $P = 10$ by using the same value (column) for N in each of the two tables. Accuracy was then determined by computing RMSEs using the entries in the respective column in the table for $P = 10$ as the criteria. Eighty sets of interpolations of this kind were produced for values of P. It should be noted that for these interpolations it was only possible to interpolate for eigenvalues up to the order of the smaller value of P used in the interpolation (e.g., when interpolating for $P = 10$ only 5 eigenvalues could be interpolated, as only 5 were available in the smallest set). This type of interpolation is hereafter referred to as across-tables interpolation.

Once again, regression estimates for both the *AH* and *LLF* equations were made for each of these conditions, and RMSEs were examined for the interpolations and regression estimates. In 69 out of the 80 cases the interpolated values produced the smallest RMSEs. A comparison of the relative sizes of the RMSEs for each method is presented in Table 15. All 80 RMSEs for the interpolations were less than .050, while less than one third of the values for either of the regression methods were as accurate. Again it would appear that the interpolations were superior to either regression method.

Table 15
Ranges of RMSE Values for Interpolations on P Across Tables

Method of Eigenvalue Estimation	RMSE Range						
	≤.01	.01-.02	.02-.05	.05-.10	.10-.20	.20-.50	>.5
Linear Interpolation	47	23	10	0	0	0	0
LL Equations	2	10	14	17	21	8	8
AH Equations	0	1	15	19	16	17	12

For the 11 cases where a regression method performed better than interpolation, it was always the *LLF* method. Eight of these eleven cases occurred for $P = 10$ involving all sample sizes except the two at either extreme. The largest discrepancy among those 11 cases resulted from an *LLF* RMSE of .005 and an interpolation RMSE of 0.030. The average improvement in accuracy over all 11 such cases was .008, which indicated little actual difference between methods. The *AH* method never did better than linear interpolation. Focusing on the two regression methods only, the *LLF* method was better than the *AH* method in 54 cases.

More Stringent Tests of Interpolation Accuracy.

In practice one would rarely want to examine more than half the eigenvalues obtained, and more likely only the first third of the eigenvalues or less. So, as a further check on the accuracy of the interpolations vis-à-vis the regression estimation methods, an additional series of comparisons were made limiting the focus to only the first third of the possible eigenvalues. These comparisons would tend to show the regression estimates in a much better light, however the linear interpolation method would still suffer the same disadvantage noted earlier in spanning at least twice the distance than would be required in actual use of the tables.

Under these circumstances, computing RMSEs for interpolations and regression estimates for only the first third of the possible eigenvalues, there was only a noticeable change for the within-table interpolations. Here the interpolation method was superior to either regression method in 53 out of the 80 possible

cases. Of the twenty-seven cases where a regression method was better, 17 were for sample sizes of 75 or 100, with the remaining ten cases scattered from sample sizes of 150 to 500. For 21 of these 27 cases the *LLF* regression method was best. Thus although the regression estimates improved when the proportion of eigenvalues estimated was reduced, and one of the regression methods proved more accurate for the smallest sample sizes, the interpolations were still generally better even for these stringent and somewhat biased tests.

It was interesting to note that for the across-tables (over values of P for fixed N) interpolations there was virtually no improvement over the results obtained earlier when all estimated eigenvalues were involved. The interpolation method was superior for 69 of the 80 possible cases. For the 11 cases where a regression method was better, eight of these cases involved $P = 10$ for the eight smallest sample sizes. The *LLF* method was best in 10 of the 11 cases. This might have been anticipated in that the maximum number of eigenvalues that could be interpolated (and hence compared) was necessarily limited to the smaller value of P involved in the interpolation. In effect, the number of eigenvalues that were compared was already somewhat reduced in the initial set of comparisons.

It is important to point out that the within-table interpolations must necessarily estimate a large number of the exact same sets of eigenvalues as did the across-tables interpolations. In fact, sixty-four of the 96 possible interpolation cases were common to the two modes of interpolating values used here. Yet, a somewhat different picture of the accuracy of interpolations vis-à-vis the regression estimates occurred for the two modes of interpolation. Some of this difference in accuracy can be explained in terms of the range of values for P or N that were involved in a particular interpolation. Interpolations over values of P across-tables covered much narrower intervals, and were generally more precise than were interpolations involving values of N within a given table.

Comparisons Involving Adjusted Interpolation RMSEs.

In all the preceding comparisons the interpolation method was at a distinct disadvantage. In practice one would interpolate between adjacent columns either within or across tables, but all the preceding interpolations spanned three columns by using the two extremes of the triad to "interpolate" for the middle. This had been done so as not to require further Monte Carlo simulations. Thus, in practice interpolation based on Tables 1 through 13 would be much more accurate than the regression estimation methods.

To attempt to control for this disadvantage a final series of comparisons were made. It was assumed that the lack of fit for linear interpolations was due to two sources, random error and nonlinearity. Given the number of replications used to generate the tabled values, the first source was presumed a minor

influence. In addition, it was assumed that even in the face of nonlinearity, the linear interpolations would have to improve when adding a point that fell between two points, where all three such points were expected to fall on a specific monotonic curve. For the example given earlier where $P = 5$, there would be no need to interpolate eigenvalues for $N = 75$ by using the columns for $N = 50$ and $N = 100$, but this is exactly what had been done in all the preceding interpolation cases. Thus, a considerable improvement in fit was expected since in practice there would be no need to skip over columns to make interpolations, as had been done above in the previous two sections (to avoid additional Monte Carlo simulations).

It was further assumed that the linear interpolation RMSEs would potentially be reduced by a factor of either 25%, 33% or 50% of their previously estimated values. In view of a monotonic nonlinear relation it is quite likely that even more than a 50% improvement in RMSEs could be obtained. Though the exact amount of improvement expected was not known, these percentage reductions in RMSE values probably range from conservative to realistic, and likely provide a fairer representation of the accuracy of the interpolation method versus the regression estimation methods.

Adjusted interpolation RMSEs, based on each of the percentage reduction factors stated above, were calculated based on the first third of the possible eigenvalues for the cases involving within-table interpolations. Under these conditions the linear interpolation method proved superior in 61 out of the 80 cases assuming only a 25% reduction in interpolation RMSEs could be expected. The 19 cases where one of the regression methods was more accurate included: values of $P = 10$ through 25, with sample sizes of 75 or 100; values of $P = 30$ through 40, for a sample size of 75; and values of $P = 40$ through 50, for a sample size of 100. The *LLF* regression method was best in 14 of these cases. The number of cases where the linear interpolation method was superior increased to 66 and 69 assuming a 33% and a 50% reduction in interpolation RMSEs, respectively.

When similar adjustments were made to RMSEs for across-table interpolations the number of cases, out of 80 possible, where the interpolation method was superior increased to 75, 76 and 77 assuming a 25%, a 33% and a 50% reduction in interpolation RMSEs, respectively. In each case where a regression method was the best it was always the *LLF* method.

Examining Redundant and Non-Redundant Interpolations.

Because of the way in which interpolations were made for specific values (columns) using the tables, sixty-four of the 80 possible interpolations conducted within-tables and across-tables were redundant with one another. As such, it

was useful to examine whether there were any cases where both types of interpolation for a specific combination of *N* and *P* were less accurate than the best regression method. This occurred in 3 out of 64 possible cases.

Still considering those cases where the two modes of interpolation were redundant with each other, and now limiting the RMSE calculations to only the first third of the eigenvalues led to 6 out of the 64 cases where a regression method was better than either mode of interpolation. Of the 32 cases that were unique to one particular mode of interpolation, a regression method was best in only 5 of these cases.

Discussion

The results clearly suggest the superiority of simple linear interpolation of the tabled eigenvalues covered by the conditions simulated over either of the two regression estimation methods. It also seems reasonable to assume that the linear interpolation method for non-tabled values (i.e. values of *N* and *P* not specifically included within the simulated data, but covered within their bounds), will generally be much better than either regression estimation method across the conditions covered by the tables. In all cases where either of the regression methods produced estimates of eigenvalues which were unreasonable (staying within the bounds of the *N*, *P* combinations used) the interpolation method produced reasonable and reasonably accurate estimates of Monte Carlo generated eigenvalues. Even where regression estimates were reasonable, the interpolated values were more accurate in most cases, and were nearly as accurate when they were not the best. This was true in spite of the fact that many of the comparisons made did not make allowance for the fact that the interpolations spanned greater distances over *N* and/or *P* than would be needed in practice. Clearly, continued use of regression estimation methods does not seem warranted for the ranges of *N* and *P* covered. And since it is never good methodological practice to extrapolate a regression equation beyond the extremes of the values used to generate model parameters, it is not recommended that PA criteria be derived from regresssion estimation methods for cases going beyond the bounds *N* and *P* used to generate the equations. On the other hand, Tables 1 through 13 can be used to provide relatively more accurate estimates of random data eigenvalues for use as parallel analysis criteria.

Researchers applying any method for determining PA criteria may want to consider the conditional nature of the decisions involved in determining the number of eigenvalues to retain when going beyond the first root. How is the test of a subsequent root to be influenced by having a given *real* root identified in a prior step? Once a first root has been tested, and its variance removed from consideration, does it still make sense to consider the test of the second root as

coming from a null case? One suggestion here would be to reduce the rank of the null matrix considered at each step by the number of roots previously removed. That is, if the first root is larger than the comparison random root, then the test of the second root should be conducted assuming P-1 and not P. Each subsequent root is evaluated against a null case assuming P reduced by the number of components previously accepted as nonrandom. Such a procedure would introduce a more conservative bias than a simple application of the procedure outlined by Horn (1965).[4] The results presented by Zwick and Velicer (1986) had indicated that there was a slight tendency for PA criteria to overestimate the number of components to retain. This suggests that it may be desirable to be slightly more conservative when applying the PA criterion. In general, it would also seem that the PA criterion, applied by either method just mentioned, may run somewhat contrary to recent findings by Cliff (1988) where he noted reliable principal components identified for eigenvalues less than one. It is clear that much work remains to be done regarding these matters.

Until mathematical statisticians develop more exacting analytical solutions for the principal components analysis of correlation matrices, researchers are advised against using current recursive regression estimation methods for determining PA criteria. In the meantime, the tables presented here should provide useful, relatively accurate and accessible estimates of PA criteria.

References

Allen, S. J. & Hubbard, R. (1986). Regression equations for the latent roots of random data correlation matrices with unities on the diagonal. *Multivariate Behavioral Research, 21*, 393-398.

Bartlett, M. S. (1933). On the theory of statistical regression. *Proceedings of the Royal Society at Edinburgh, 53*, 260-283.

Bartlett, M. S. (1951). The effect of standardization on a χ^2 approximation in factor analysis. *Biometrika, 38*, 337-344.

Clemm, D. S., Chattopadhyay, A. K. & Krishnaiah, P. R. (1973). Upper percentage points of the individual roots of the Wishart matrix. *Sankhya*, Series B, 325-328.

Clemm, D. S., Krishnaiah, P. R. & Waikar, V. B. (1973). Tables for the extreme roots of the Wishart matrix. *Journal of Statistical Computation and Simulation, 2*, 65-92.

Cliff, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin, 103*, 276-279.

Flaherty, V. L., Lautenschlager, G. J., & Lance, C. L. (1988, March). *Parallel analysis criteria: Application to the Minnesota Satisfaction Questionnaire.* Paper presented at the meeting of the Southeastern Psychological Association, New Orleans, LA.

---

[4] I want to thank Richard J. Harris and an anonymous reviewer for bringing this issue to my attention. Harris further suggested the step-down method as a means for the sequential testing of eigenvalues beyond the first.

Hays, R. D. (1987). PARALLEL: A program for performing parallel analysis. *Applied Psychological Measurement, 11*, 58.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Hsu, P. L. (1939). On the distribution of roots of certain determinantal equations. *Annals of Eugenics, 9*, 250-258.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kendall, M. G. & Stuart, A. (1969). *The advanced theory of statistics, Volume 1*. London: Hafner.

Krishnaiah, P. R. (1980). Computations of some multivariate distributions. In P. R. Krishnaiah (Ed.) *Handbook of statistics 1: analysis of variance*. Amsterdam: North-Holland.

Krishnaiah, P. R. & Chang, T. C. (1971). On the exact distributions of the extreme roots of the Wishart and MANOVA matrices. *Journal of Multivariate Analysis, 1*, 108-117.

Humphreys, L. G. & Montanelli, Jr., R. G. (1975). An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research, 10*, 193-205.

Lautenschlager, G. J. (1987). PARANAL: A program for estimating parallel analysis criteria for principal components analysis [Computer program]. Author.

Lautenschlager, G. J. (in press). PARANAL.TOK: A program for developing parallel analysis criteria. *Applied Psychological Measurement*.

Lautenschlager, G. J., Lance, C. E. & Flaherty, V. L. (1989). Parallel analysis criteria: revised regression equations for estimating the latent roots of random data correlation matrices. *Educational and Psychological Measurement, 49*, 339-345.

Lawley, D. N. (1956). Tests of significance for the latent roots of covariance and correlation matrices. *Biometrika, 43*, 128-136.

Montanelli, Jr., R. G. & Humphreys, L. G. (1976). Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. *Psychometrika, 41*, 341-348.

Myers, R. H. (1986). *Classical and modern regression*. Boston: Duxbury.

Roy, S. N. (1953). On a heuristic method of test construction and its use in multivariate analysis. *Annals of Mathematical Statistics, 24*, 220-238.

Roy, S. N. (1957). *Some aspects of multivariate analysis*. New York: Wiley.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*, 321-327.

Wilkinson, L. (1986). *SYSTAT: The system for statistics*. Evanston, IL: SYSTAT, Inc.

Zwick, W. R. & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.