

PROBABILISTIC CONSIDERATIONS OF THE CONSISTENCY INDEX IN RECIPROCAL MATRICES

Luis G. Vargas

The Joseph M. Katz Graduate School of Business
University of Pittsburgh
Pittsburgh, PA 15260
lgvargas@pitt.edu

Keywords: consistency, consistency ratio, random judgments, statistical consistency

Summary: *When checking the inconsistency level of a positive reciprocal matrix there are two parameters used, a benchmark, the average, and a consistency level, usually 10 percent. We show results from a simulation experiment with 100,000 positive random reciprocal matrices of size varying from 3 to 15 that support the use of both the average and the 10 percent consistency level. In addition, a statistical hypothesis test is developed.*

1. Introduction

As in physical problems where there are two ways to investigate metric properties of a system: deterministic and probabilistic, one would like to assume that those very two ways of thinking also apply to order properties. Some investigators (e.g., Crawford, 1987) have used the statistical approach to derive the best fit vector w for a positive reciprocal matrix of paired comparisons A using the logarithmic least squares metric which yields the maximum likelihood estimator of w . The outcome is a vector whose components are the geometric mean of the rows of A . This result coincides with the principal eigenvector for $n \leq 3$ but deviates from it for $n > 3$. The geometric mean is based solely on metric closeness without regard to dominance, which is the essential property one is attempting to capture by taking powers of the matrix. When the matrix A is consistent, the two outcomes coincide for all n because $A^n = n^{n-1}A$ and powers of A coincide with A to within a multiplicative constant. However, it is when A is inconsistent that the two differ, and the greater the inconsistency the less the geometric mean captures order. The strength of the statistical approach is that given partial information (the sample), and under some specific assumptions about the population from which a sample is selected, we can infer something about the parameters of the statistical population through hypothesis testing. These inferences assume that either we know the probability distribution of the population of interest, or the behavior of the parameters that need to be estimated is well defined. For example, when doing hypothesis tests about means, either we assume that the population is normal, or that the sample size is sufficiently large so that the central limit theorem can be invoked. The same problem arises when studying consistency. To study it from a probabilistic point of view we first need to characterize the population from which consistency is obtained. To study it from the deterministic point of view we need to set what we consider to be an acceptable level of inconsistency. This level need not be closely related to the level one would obtain from the statistical approach as we shall see below. One thing seems to be clear, what one learns from probability is less cogent as one moves away from consistency as we shall also see.

In the Analytic Hierarchy Process scales are built from priorities derived from judgments represented with pairwise comparisons. These pairwise comparisons are obtained by comparing two elements from a set of elements according to a criterion. Each criterion yields a set of pairwise comparisons arranged in a matrix. It is assumed that the judgments satisfy the reciprocal property, i.e., if a_{ij} represents how strongly

element i is preferred to element j , then $a_{ji} = 1/a_{ij}$. If in addition to the reciprocal property, the judgments satisfy the relationship:

$$a_{ij}a_{jk} = a_{ik}, \text{ for all } i, j \text{ and } k \quad (1)$$

then we say that the judgments are consistent. Consistency is necessary but not sufficient to make the priorities derived from the judgments valid in practice. To measure inconsistency, Saaty [1980] developed an index based on the principal eigenvalue of the matrix of pairwise comparisons.

Let $\mathbf{A} = \{a_{ij}\}$ be a reciprocal matrix of order n . Let λ_{\max} be the principal eigenvalue of \mathbf{A} . It is known that if (1) holds, that is, \mathbf{A} is a consistent matrix, then $\lambda_{\max} = n$, and that in general $\lambda_{\max} \geq n$. Since the trace of the matrix is equal to the sum of its eigenvalues, we have

$$\mu \equiv \frac{\lambda_{\max} - n}{n - 1} = \frac{1}{n - 1} \sum_{\lambda_i \neq \lambda_{\max}} \lambda_i \quad (2)$$

Thus, the index μ measures the average inconsistency in a reciprocal matrix of judgments. This index on its own has no meaning unless we compare it with some benchmark to determine the magnitude of the deviation from consistency.

A simple rule of thumb proposed by Saaty [1980], and justified in terms of orders of magnitude, is that the ratio of μ to the average of the same index from randomly generated reciprocal matrices (called the consistency ratio, CR) should be less than 10 percent. For example, consider the following 3-by-3 reciprocal matrix:

$$\begin{bmatrix} 1 & 2 & 2 \\ 1/2 & 1 & 2 \\ 1/2 & 1/2 & 1 \end{bmatrix}.$$

that has a CR = 0.052. This matrix is not consistent. However, the consistency ratio CR indicates that the inconsistency is acceptable (CR < 0.10). According to Golden and Wang (1989), the problem with the 10 percent rule is that it is too restrictive for values of $n > 3$ (see Table 1, column 4). Because of this, they state that the consistency index must be a function of the matrix size.

In this paper we show that because individuals do not provide judgments at random, even when they do not have enough information, that the 10 percent rule is appropriate and that the average benchmark is acceptable. The 10 percent rule does not refer to the randomness of the judgments. It is a deterministic approach. From a probabilistic point of view, the question is: Is there a value of μ to be considered natural and beyond which the judgments can be considered random? Note that randomness is not the same as inconsistency, and hence, the 10 percent rule cannot be used to test randomness. For this purpose we need a new statistical test also based on the concept of consistency ratio.

Table 1: Sample results (m=100,000)

n	ϖ	s	$P[\mathfrak{R} < 0.10]$
3	0.493583	0.66005	0.22430
4	0.825265	0.600793	0.03596
5	1.03846	0.494065	0.00300
6	1.17146	0.397672	0.00010
7	1.25865	0.327452	0.00001
8	1.31826	0.27803	0.00000
9	1.3641	0.236991	0.00000
10	1.41009	0.209089	0.00000
11	1.4349	0.186182	0.00000
12	1.45621	0.168613	0.00000
13	1.47205	0.15355	0.00000
14	1.48551	0.140442	0.00000
15	1.49788	0.129745	0.00000

2. Simulation Results

Let $\{X_{ij}, 1 \leq i < j \leq n\}$ be a set of independent identically distributed random variables with probability mass function given by:

$$P[X_{ij} = k] = \frac{1}{17}, k = \frac{1}{9}, \dots, \frac{1}{2}, 1, 2, \dots, 9, \\ 1 \leq i < j \leq n$$

Let $X_{ji} = X_{ij}^{-1}$ for all i and j . Let Λ_{\max} be the random variable corresponding to the principal eigenvalue of random reciprocal matrices whose entries are the variables $\{X_{ij}, 1 \leq i < j \leq n\}$, let

$$\Omega_n \equiv \frac{\Lambda_{\max} - n}{n - 1}, \text{ and let } E[\Omega_n] = \mu_n. \text{ Let } \mathfrak{R}_n \equiv \frac{\Omega_n}{E[\Omega_n]}$$

be the random variable corresponding to the consistency ratio CR. In Table 2 we give the empirical values of \mathfrak{R}_n for percentiles ranging from 1 to 99.

In Figure 1 we plot the critical values of \mathfrak{R}_n versus the matrix size.

Table 2: Empirical Percentiles of the Consistency Ratio \mathfrak{R}_n for Different Size Matrices

Percentile	MATRIX SIZE												
	3	4	5	6	7	8	9	10	11	12	13	14	15
0.01	0.000	0.043	0.150	0.277	0.401	0.497	0.583	0.642	0.691	0.724	0.754	0.777	0.794
0.05	0.004	0.125	0.275	0.432	0.559	0.642	0.708	0.751	0.782	0.807	0.826	0.843	0.855
0.1	0.019	0.195	0.375	0.542	0.655	0.724	0.775	0.807	0.833	0.851	0.866	0.878	0.888
0.25	0.136	0.388	0.619	0.758	0.822	0.859	0.884	0.900	0.913	0.923	0.930	0.936	0.943
0.5	0.441	0.824	0.983	1.004	1.006	1.005	1.004	1.002	1.002	1.001	1.000	1.001	1.001
0.75	1.256	1.501	1.346	1.242	1.182	1.146	1.119	1.103	1.089	1.079	1.071	1.064	1.059
0.9	3.021	2.089	1.644	1.442	1.334	1.269	1.221	1.189	1.165	1.148	1.134	1.121	1.111
0.95	4.207	2.386	1.812	1.554	1.418	1.340	1.280	1.240	1.211	1.188	1.172	1.155	1.142
0.99	5.757	2.886	2.100	1.753	1.574	1.464	1.390	1.335	1.294	1.261	1.239	1.216	1.198

Figure 1 shows that as the matrix size increases, randomly generated reciprocal matrices become more naturally inconsistent until the inconsistency seems indistinguishable from the average random inconsistency. Thus, if we make no effort to think about the judgments, the inconsistency of the matrix should be around the average inconsistency.

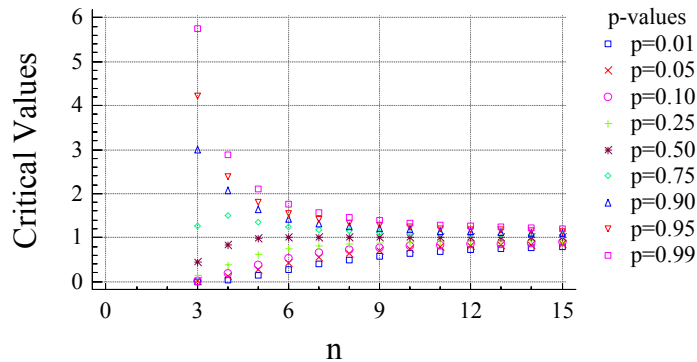


Figure 1. Critical Values of \mathfrak{R}_n vs. Matrix Size (n) for Different p-values

These results would suggest that when individuals compare larger and larger sets of elements, their natural inconsistency level is almost the same as the random inconsistency level. However, we have observed, quite the opposite effect in practice. Considered the following reciprocal matrix:

$$\begin{bmatrix}
 1 & 5 & 3 & 7 & 6 & 6 & 1/3 & 1/4 \\
 1/5 & 1 & 1/3 & 5 & 3 & 3 & 1/5 & 1/7 \\
 1/3 & 3 & 1 & 6 & 3 & 4 & 6 & 1/5 \\
 1/7 & 1/5 & 1/6 & 1 & 1/3 & 1/4 & 1/7 & 1/8 \\
 1/6 & 1/3 & 1/3 & 3 & 1 & 1/2 & 1/5 & 1/6 \\
 1/6 & 1/3 & 1/4 & 4 & 2 & 1 & 9 & 1/6 \\
 3 & 5 & 1/6 & 7 & 5 & 5 & 1 & 1/2 \\
 4 & 7 & 5 & 8 & 6 & 6 & 2 & 1
 \end{bmatrix}$$

whose consistency ratio CR=0.379. According to the results in Table 2, this matrix should be considered consistent because the chances that this matrix is obtained at random is less than 1 percent. According to Saaty's rule this is not the case.

Definition (Saaty): A positive reciprocal matrix \mathbf{A} is almost consistent if $\frac{\mu(\mathbf{A})}{\mu_n} < 0.1$

Following a statistical approach, the matrix of judgments is now the sample, and we must decide if its consistency level (the statistic) is statistically acceptable. Thus, given a positive reciprocal matrix \mathbf{A} with consistency index $\mu(\mathbf{A})$, to test if $\mu(\mathbf{A})$ is statistically significantly different that $0.1\mu_n$, we need to test the hypothesis $\mathbf{H}_0 : \mu \leq \mu_n \varepsilon$, where $\varepsilon = 0.1$. The p-value corresponding to this hypothesis is given by

$$\text{p-value} = P \left[\mathfrak{R}_n > \frac{\mu(\mathbf{A})}{\mu_n \varepsilon} \right].$$

If the p-value is small, then we would fail to accept the null hypothesis, $\mathbf{H}_0 : \mu \leq \mu_n \varepsilon$, and we could infer that the judgments provided lead to an inconsistency larger than $\mu_n \varepsilon$, the expected level. This formulation suggests that instead of comparing the consistency of a matrix with the average consistency we could also compare it with some other percentile of the distribution of $\Omega_n \equiv \frac{\Lambda_{\max} - n}{n - 1}$, e.g., the 95th percentile? This question is meaningful because as shown in Table 1, the consistency ratio is significantly larger than 0.10 for values of $n > 4$.

Let $m_{n,\alpha}$ be a value of $\Omega_n \equiv \frac{\Lambda_{\max} - n}{n - 1}$ such that $P[\Omega_n > m_{n,\alpha}] = \alpha$. Let $\mathbf{M}_{n,\alpha} \equiv \frac{\Omega_n}{m_{n,\alpha}}$.

Definition: A positive reciprocal matrix \mathbf{A} is almost consistent at the α level if $\frac{\mu(\mathbf{A})}{m_{n,\alpha}} < \varepsilon$.

Definition: A positive reciprocal matrix \mathbf{A} is said to be statistically consistent at the α level if the null hypothesis $\mathbf{H}_0 : \mu \leq m_{n,\alpha}\varepsilon$ cannot be rejected.

Lemma: *An almost consistent matrix at the α level is always statistically consistent at the α level.*

Proof: If \mathbf{A} is almost consistent then $\frac{\mu(\mathbf{A})}{m_{n,\alpha}} < \varepsilon$ or $\frac{\mu(\mathbf{A})}{m_{n,\alpha}\varepsilon} < 1$ and hence,

$$\begin{aligned} P\left[M_{n,\alpha} > \frac{\mu(\mathbf{A})}{m_{n,\alpha}\varepsilon}\right] &= 1 - P\left[M_{n,\alpha} \leq \frac{\mu(\mathbf{A})}{m_{n,\alpha}\varepsilon}\right] \\ &> 1 - P[M_{n,\alpha} \leq 1] \\ &= P[\Omega_n \leq m_{n,\alpha}] \\ &= 1 - \alpha \end{aligned}$$

and we would accept the hypothesis $\mathbf{H}_0 : \mu \leq m_{n,\alpha}\varepsilon$ if α is sufficiently large.

Recall that $P[\Omega_n > m_{n,\alpha}] = \alpha$, and if we use of the average as the benchmark, then the values of α are large,

Matrix Size	Mean	$\alpha = P[\Omega_n > \mu_n]^*$
3	0.493583	0.3412
4	0.825265	0.3941
5	1.03846	0.4299
6	1.17146	0.4510
7	1.25865	0.4633
8	1.31826	0.4708
9	1.3641	0.4763
10	1.39723	0.4798
11	1.42524	0.4825
12	1.44724	0.4845
13	1.46484	0.4860
14	1.47906	0.4873
15	1.49234	0.4884

* The probabilities $P[\Omega_n > \mu_n]$ have been estimated using the distributions derived in the next section.

3. The Theoretical Probability Distribution of Ω_n

Because numerical goodness-of-fit tests tend to be very restrictive as the sample size increases, we use a quantile plot to test for the distribution of Ω_n . We empirically observe that the gamma random variable provides close estimates of both the mean and the standard deviation of Ω_n . Figure 2 shows the comparisons for $n=3$. Table 3 gives the parameters of the gamma random variables, the theoretical mean and standard deviation, and the sample mean and standard deviation.

Table 3. Probability Distribution of Ω_n .

n	Gamma alpha	beta	E[X]	stddev(X)	sample mean	sample std.dev.
3	0.586355	1.15971	0.505605	0.660284	0.493583	0.66005
4	1.56954	1.90178	0.825301	0.658758	0.825265	0.600793
5	3.59988	3.46654	1.038465	0.547328	1.03846	0.494065
6	7.37915	6.29909	1.171463	0.431246	1.17146	0.397672
7	13.1702	10.4638	1.258644	0.346822	1.25865	0.327452
8	20.6816	15.6885	1.318265	0.289875	1.31826	0.27803
9	31.3003	22.9457	1.364103	0.243822	1.3641	0.236991
10	43.2073	30.9236	1.397227	0.212563	1.39723	0.208168
11	57.4517	40.3101	1.425243	0.188035	1.42524	0.185151
12	72.8596	50.3427	1.447272	0.169554	1.44724	0.167415
13	89.8307	61.3248	1.464835	0.154553	1.46484	0.153178
14	109.878	74.2892	1.479058	0.141101	1.47906	0.13999
15	131.091	87.8428	1.492336	0.130341	1.49234	0.129487

As can be seen in Figure 2, other distributions such as the normal and the lognormal do not seem to fit as well as the gamma distribution for values of $n < 7$. However, the Weibull distribution seems to fit just as well. Tables 4-10 give estimates of the percentiles obtained using Weibull and Gamma distributions as well as the average absolute percent deviation for $n=3,5,7,9, 11, 13$ and 15 .

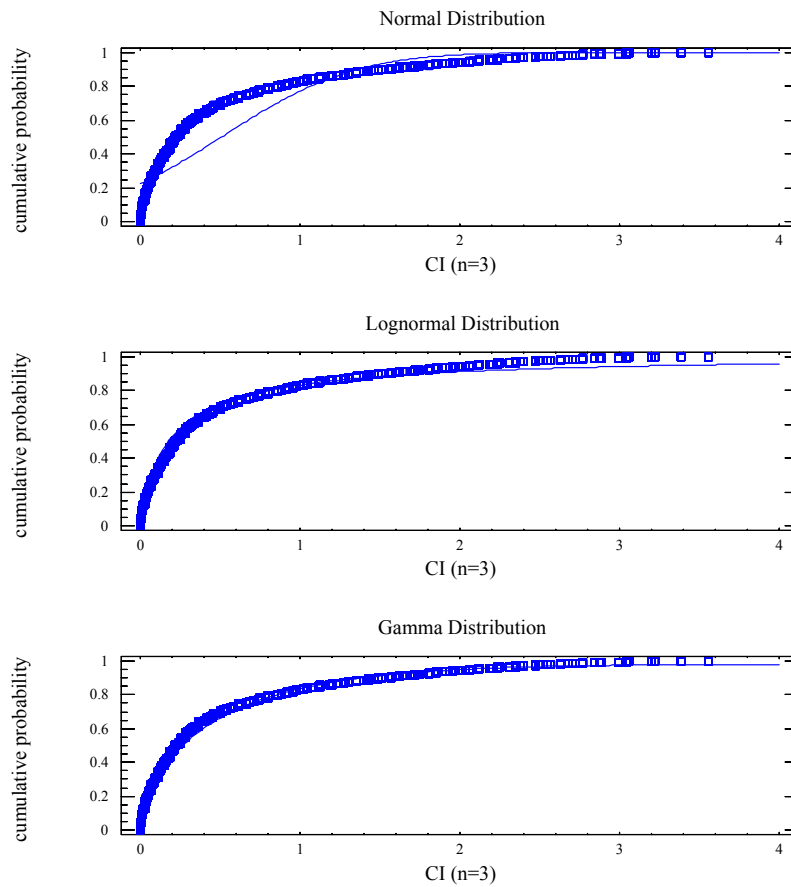


Figure 2. Normal, lognormal and gamma approximations of Ω_n

Table 4. Approximation by Weibull and Gamma for $n=3$

CDF	Weibull	Gamma	Empirical
0.01	0.000564	0.000276	0
0.05	0.005784	0.004301	0.001847
0.1	0.016167	0.014128	0.009147
0.25	0.06784	0.070198	0.067805
0.5	0.238117	0.261633	0.217845
0.75	0.640634	0.683937	0.61559
0.9	1.32205	1.32177	1.4861
0.95	1.92499	1.83445	2.08333
0.99	3.55685	3.07693	2.8459

Absolute % Deviation

Weibull	Gamma
N/A	N/A
213.12%	132.82%
76.74%	54.44%
0.05%	3.53%
9.31%	20.10%
4.07%	11.10%
11.04%	11.06%
7.60%	11.95%
24.98%	8.12%

Average 43.36% 31.64%

Table 5. Approximation by Weibull and Gamma for n=5

CDF				Absolute % Deviation	
	Weibull	Gamma	Empirical	Weibull	Gamma
0.01	0.149248	0.190021	0.155464	4.00%	22.23%
0.05	0.309825	0.328546	0.285502	8.52%	15.08%
0.1	0.427763	0.427261	0.389327	9.87%	9.74%
0.25	0.67094	0.637046	0.642998	4.35%	0.93%
0.5	0.994997	0.944053	1.02066	2.51%	7.51%
0.75	1.35743	1.33768	1.39773	2.88%	4.30%
0.9	1.70397	1.77232	1.70751	0.21%	3.80%
0.95	1.91723	2.07086	1.88131	1.91%	10.08%
0.99	2.32464	2.71184	2.18171	6.55%	24.30%

Average 4.60% 9.46%

Table 6. Approximation by Weibull and Gamma for n=7

CDF				Absolute % Deviation	
	Weibull	Gamma	Empirical	Weibull	Gamma
0.01	0.475979	0.593908	0.504175	5.59%	17.80%
0.05	0.69449	0.747413	0.703095	1.22%	6.30%
0.1	0.820592	0.839624	0.824073	0.42%	1.89%
0.25	1.03572	1.0107	1.03449	0.12%	2.30%
0.5	1.26988	1.22694	1.26661	0.26%	3.13%
0.75	1.49121	1.47208	1.48818	0.20%	1.08%
0.9	1.67732	1.71852	1.67863	0.08%	2.38%
0.95	1.78282	1.87808	1.78526	0.14%	5.20%
0.99	1.96966	2.20246	1.98082	0.56%	11.19%

Average 0.38% 4.18%

Table 7. Approximation by Weibull and Gamma for n=9

CDF				Absolute % Deviation	
	Weibull	Gamma	Empirical	Weibull	Gamma
0.01	0.712058	0.861571	0.795691	10.51%	6.59%
0.05	0.918979	0.98931	0.966498	4.92%	2.28%
0.1	1.02857	1.06244	1.05701	2.69%	0.54%
0.25	1.20367	1.19271	1.20631	0.22%	1.36%
0.5	1.38128	1.3496	1.36908	0.89%	1.95%
0.75	1.53956	1.5197	1.52652	0.85%	0.68%
0.9	1.6668	1.68441	1.66584	0.06%	1.86%
0.95	1.73689	1.78836	1.74618	0.53%	4.22%
0.99	1.8578	1.99444	1.89576	2.00%	9.87%

Average 1.52% 2.84%

Table 8. Approximation by Weibull and Gamma for n=11

CDF				Absolute % Deviation	
	Weibull	Gamma	Empirical	Weibull	Gamma
0.01	0.87032	1.02438	0.984301	11.58%	4.07%
0.05	1.05698	1.13068	1.11517	5.22%	1.39%
0.1	1.15169	1.19024	1.18702	2.98%	0.27%
0.25	1.2982	1.29437	1.30145	0.25%	0.54%
0.5	1.44169	1.41699	1.42817	0.95%	0.78%
0.75	1.56588	1.54712	1.55153	0.92%	0.28%
0.9	1.66352	1.67084	1.66083	0.16%	0.60%
0.95	1.71654	1.74795	1.72548	0.52%	1.30%
0.99	1.80682	1.89892	1.84455	2.05%	2.95%

Average 2.74% 1.36%

Table 9. Approximation by Weibull and Gamma for n=13

CDF				Absolute % Deviation	
	Weibull	Gamma	Empirical	Weibull	Gamma
0.01	0.980137	1.12933	0.984301	0.42%	14.73%
0.05	1.14864	1.22022	1.11517	3.00%	9.42%
0.1	1.23199	1.27061	1.18702	3.79%	7.04%
0.25	1.35852	1.35787	1.30145	4.39%	4.34%
0.5	1.47991	1.45941	1.42817	3.62%	2.19%
0.75	1.58319	1.56588	1.55153	2.04%	0.92%
0.9	1.66334	1.66603	1.66083	0.15%	0.31%
0.95	1.70649	1.72797	1.72548	1.10%	0.14%
0.99	1.77942	1.84824	1.84455	3.53%	0.20%
Average				2.45%	4.37%

Table 10. Approximation by Weibull and Gamma for n=15

CDF				Absolute % Deviation	
	Weibull	Gamma	Empirical	Weibull	Gamma
0.01	1.06706	1.2059	1.18537	9.98%	1.73%
0.05	1.2184	1.28462	1.27647	4.55%	0.64%
0.1	1.29189	1.32794	1.32517	2.51%	0.21%
0.25	1.40192	1.4025	1.40657	0.33%	0.29%
0.5	1.50591	1.48854	1.49308	0.86%	0.30%
0.75	1.59329	1.57805	1.58016	0.83%	0.13%
0.9	1.66045	1.66161	1.65769	0.17%	0.24%
0.95	1.69639	1.713	1.70348	0.42%	0.56%
0.99	1.75679	1.81223	1.78741	1.71%	1.39%
Average				2.37%	0.61%

For $n \geq 7$, Ω_n also fits a normal random variable. Let $\mathbf{A} = (a_{ij} \equiv \frac{w_i}{w_j} \varepsilon_{ij})$ where $\mathbf{w} = (w_i)$ is the principal right eigenvector of \mathbf{A} . It can be easily shown that $\mu_n \equiv \frac{\lambda_{\max} - n}{n - 1} = \frac{1}{\frac{n(n-1)}{2}} \sum_{1 \leq i < j \leq n} (\varepsilon_{ij} + \varepsilon_{ji})$, and hence, the consistency index is the average of a set of random variables. By the Central Limit Theorem, for n sufficiently large, in this case, $n(n-1)/2$, the distribution of Ω_n approaches a normal.

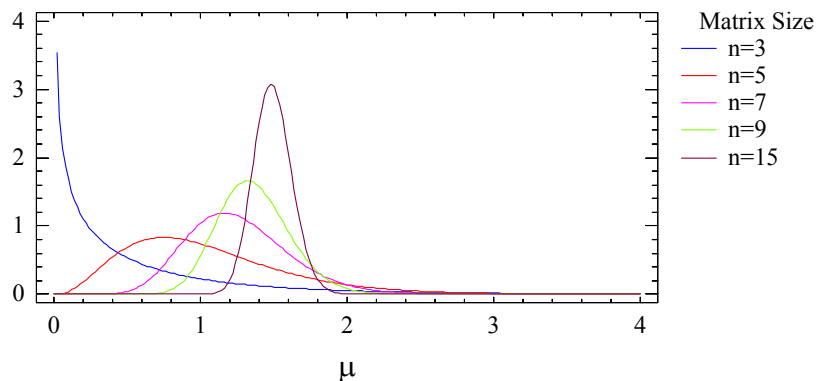


Figure 3. Distribution of the Consistency Index Ω_n

In Figure 4 we show some of the graphical tests of normality conducted.

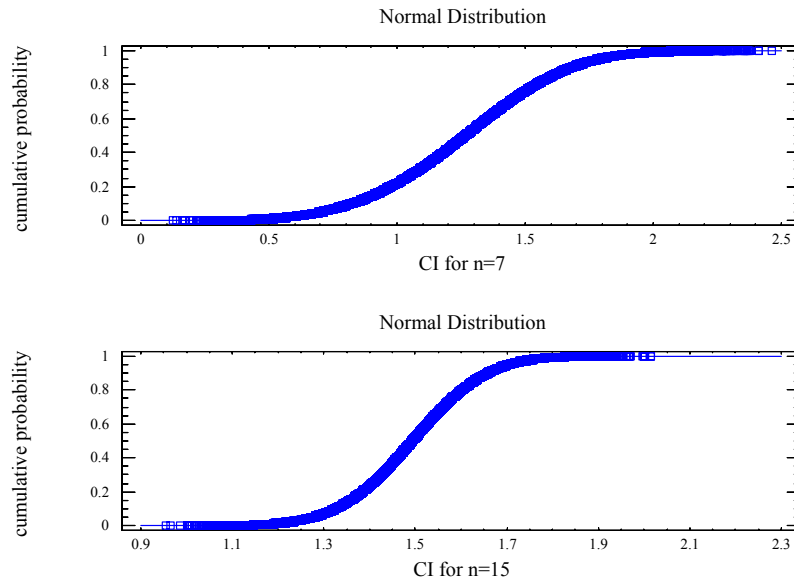


Figure 4. Quantile Plot of the consistency index for $n=7$ and $n=15$.

4. Consistency Ratios Based on Percentiles

Testing the hypothesis $H_0 : \mu \leq m_{n,\alpha} \varepsilon$ requires the computation of

$$\text{p-value} = P \left[M_{n,\alpha} > \frac{\mu(\mathbf{A})}{m_{n,\alpha} \varepsilon} \right].$$

The p-value, for the distributions at hand, is a monotonically decreasing function of the level α . Thus, if $m_{n,\alpha} \geq \mu_n$, then $P[\Omega_n > m_{n,\alpha}] \leq P[\Omega_n > \mu_n]$, and using a percentile greater than the average means that we are willing to accept more inconsistency. So, the question is: Which percentile is more appropriate? Note that for $n \geq 7$ the distribution of the consistency index is a Normal. Figure 3 shows the density functions of the gamma random variables estimated from the data. For $n \geq 7$, Ω_n has the shape of a Normal. Quantile plots confirm this fact (see Figure 4 for $n=7$ and $n=15$). As n increases it is clear that the standard deviation of the consistency index decreases (see Figure 5), thus, for $n \geq 7$ the best and easier percentile to use is the average. For $n \leq 7$, there does not seem to any particular reason why the average does not give good results. To illustrate the behavior of other percentiles, and in particular the 95th percentile of Ω_n we use some examples.

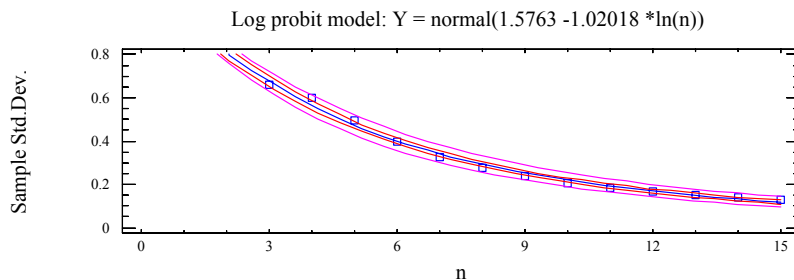


Figure 5. $\sigma(\Omega_n)$ as a function of the matrix size n

5. Examples

Consider the following reciprocal matrices:

(1)

$$\begin{bmatrix} 1 & 2 & 5 \\ 1/2 & 1 & 2 \\ 1/5 & 1/2 & 1 \end{bmatrix} \quad , CR = 0.005 \text{ (consistent),} \\ \text{p-value } (\mu_3) = 0.794 \text{ (statistically consistent),} \\ \text{p-value(95}^{\text{th}} \text{ percentile) = 0.864 (stat. consistent)}$$

(2)

$$\begin{bmatrix} 1 & 2 & 2 \\ 1/2 & 1 & 2 \\ 1/2 & 1/2 & 1 \end{bmatrix} \quad , CR = 0.052 \text{ (consistent),} \\ \text{p-value } (\mu_3) = 0.790 \text{ (statistically consistent)} \\ \text{p-value (95}^{\text{th}} \text{ percentile) = 0.514 (stat. consistent)}$$

(3)

$$\begin{bmatrix} 1 & 2 & 1/2 \\ 1/2 & 1 & 2 \\ 2 & 1/2 & 1 \end{bmatrix} \quad , CR = 0.481 \text{ (inconsistent),} \\ \text{p-value } (\mu_3) = 0.019 \text{ (statistically inconsistent)} \\ \text{p-value (95}^{\text{th}} \text{ percentile) = 0.038 (stat. inconsistent)}$$

(4)

$$\begin{bmatrix} 1 & 2 & 4 & 8 \\ 1/2 & 1 & 2 & 4 \\ 1/4 & 1/2 & 1 & 1 \\ 1/8 & 1/4 & 1 & 1 \end{bmatrix} \quad , CR = 0.023 \text{ (consistent) ,} \\ \text{p-value } (\mu_4) = 0.894 \text{ (statistically consistent)} \\ \text{p-value (95}^{\text{th}} \text{ percentile) = 0.941 (stat. consistent)}$$

(5)

$$\begin{bmatrix} 1 & 2 & 3 & 1/2 \\ 1/2 & 1 & 2 & 4 \\ 1/3 & 1/2 & 1 & 5 \\ 2 & 1/4 & 1/5 & 1 \end{bmatrix} \quad , CR = 0.551 \text{ (inconsistent),} \\ \text{p-value } (\mu_5) = 0.000003 \text{ (statistically inconsistent)} \\ \text{p-value (95}^{\text{th}} \text{ percentile) = 0.006 (stat. inconsistent)}$$

(6)

$$\begin{bmatrix} 1 & 5 & 5 & 6 & 3 \\ 1/5 & 1 & 1/4 & 1/3 & 2 \\ 1/5 & 4 & 1 & 4 & 1/5 \\ 1/6 & 3 & 1/4 & 1 & 1/7 \\ 1/3 & 1/2 & 5 & 7 & 1 \end{bmatrix} \quad , CR = 0.440 \text{ (inconsistent),} \\ \text{p-value } (\mu_5) = 0.0001 \text{ (statistically inconsistent)} \\ \text{p-value (95}^{\text{th}} \text{ percentile) = 0.026 (stat. inconsistent)}$$

(7)

$$\begin{bmatrix} 1 & 5 & 3 & 7 & 6 & 6 & 1/3 & 1/4 \\ 1/5 & 1 & 1/3 & 5 & 3 & 3 & 1/5 & 1/7 \\ 1/3 & 3 & 1 & 6 & 3 & 4 & 6 & 1/5 \\ 1/7 & 1/5 & 1/6 & 1 & 1/3 & 1/4 & 1/7 & 1/8 \\ 1/6 & 1/3 & 1/3 & 3 & 1 & 1/2 & 1/5 & 1/6 \\ 1/6 & 1/3 & 1/4 & 4 & 2 & 1 & 9 & 1/6 \\ 3 & 5 & 1/6 & 7 & 5 & 5 & 1 & 1/2 \\ 4 & 7 & 5 & 8 & 6 & 6 & 2 & 1 \end{bmatrix} \quad , CR = 0.379 \text{ (inconsistent),} \\ \text{p-value } (\mu_8) = 10^{-7} \text{ (statistically inconsistent)} \\ \text{p-value (95}^{\text{th}} \text{ percentile) = 0.015} \\ \text{(statistically inconsistent)}$$

6. Conclusions

In this paper we provide a statistical test of consistency to decide if the judgments of a positive reciprocal could be considered statistically consistent. The test is based on two parameters, a percentile ($m_{n,\alpha}$) and a consistency level (ε). This test is equivalent to the 10 percent rule. Because the level of inconsistency of randomly generated positive reciprocal matrices tends to stabilize around the average inconsistency (see Figure 3), it is appropriate to use the average as the benchmark of the test. The choice of consistency level, ε , is equivalent to the choice of a significant level in a statistical test, and hence, 10 percent is neither too strict nor too relaxed. It is just right!

REFERENCES

- Crawford, G.B., "The Geometric Mean Procedure for Estimating the Scale of a Judgment Matrix," *Mathematical Modelling* 9, 3-5 (1987) pp.327-334.
- Golden, B.L. and Q. Wang, "An Alternate Measure of Consistency," in *The Analytic Hierarchy Process: Applications and Studies*. B.L. Golden, E.A. Wasil and P.T. Harker (Eds.), Springer-Verlag, 1989.
- Saaty, T.L., *The Analytic Hierarchy Process*, McGraw-Hill, New York, 1980.

