START Selected Topics in Assurance Related Technologies

Volume 10, Number 4

The Chi-Square: a Large-Sample Goodness of Fit Test

Table of Contents

Introduction

Most statistical methods (of parametric statistics) assume an underlying distribution in the derivation of their results (methods that do not assume an underlying distribution are called non-parametric, or distribution free, and will be the topic of a separate paper).

When we assume that our data follow a specific distribution we are taking a serious risk. If the assumed distribution does not hold then the confidence levels of the confidence intervals (CI) or of the hypotheses tests implemented may be completely off [5]. Consequences of incorrectly identifying the underlying distribution may prove very costly. One way to deal with this problem is to check distribution assumptions carefully.

There are two approaches to checking distribution assumptions. One is via empirical procedures. These are easy to understand and implement and are based on intuitive and graphical properties of the distribution that we want to assess. Such empirical procedures can be used to check and validate distribution assumptions and have been discussed at length in several other RAC START sheets [6, 7, 8, and 9].

There are also other, more formal procedures to assess the underlying distribution of a data set. These are the Goodness of Fit (GoF) tests, based on statistical theory [3, 4]. They are numerically convoluted and usually require specific software to aid the user through their lengthy calculations. But their results are quantifiable and more reliable than the ones from the empirical procedures. This paper discusses one of such theoretical GoF procedures, for large samples: the Chi-Square GoF test.

In what follows, we review some issues associated with the implementation of the Chi-Square GoF test, especially when assessing distribution assumptions for the Exponential, Weibull, Normal, and Lognormal. For, these distributions are widely used in quality and reliability work. We first review some theoretical considerations that will help us better understand (and use) the underlying statistical theory behind the GoF tests. Then, we develop several numerical and graphical examples that illustrate how to implement and interpret the Chi-Square GoF test for fitting several distributions.

Some Statistical Background

Establishing the underlying distribution of a data set or random variable is crucial for the correct implementation of some statistical procedures. For example, deriving the test and CI for the population MTBF requires knowledge about the distribution of the lives of the device. If the lives are Exponential, things will be done one way; if they are Weibull, they will be done differently. Therefore, we first need to establish the life distribution from the data, before we can correctly implement the test procedures.

The GoF tests are the statistical procedures that allow us to establish whether an assumed distribution is correct. GoF tests are essentially based on either of two distribution basics: the cumulative distribution function, or CDF, and the probability density function or PDF. Procedures based on the CDF are called "distance tests" while those based on the PDF are called "area tests" [3, 4]. The Chi-Square GoF test, which is the topic of this paper, is an area test.

To assess data, we implement a well-defined scheme. First, assume that data follow a pre-specified distribution (e.g., Normal). Then, we either estimate the distribution parameters (e.g., mean and variance) from the data or obtained from prior experience. Such process yields the "composite" distribution hypothesis (which has more than one element that jointly must be true) called the null hypothesis (or H_0). The negation of the assumed distribution (null hypothesis) is called the alternative hypothesis (or H_1). We then test the assumed (hypothesized) distribution using the data set. Finally, H_0 is rejected whenever any one (or more) of the several elements in hypothesis H_0 is not supported by the data.

The Chi-Square test is conceptually based on the probability density function (PDF) of the assumed distribution. If this distribution is correct, its PDF (yielding an area of unity) should closely encompass the data range (of X). We thus select convenient values in this data range (Figure 1) that divide it into several subintervals. Then, we compute the number of data points in each subinterval. These are called "observed" values. Then, we compute the number that should have fallen in these same subintervals, according to the PDF of the assumed distribution. These are called the "expected" values and the Chi-Square test requires at least five of them in every subinterval. Finally, we compare these two results. If they agree (probabilistically) then the data supports the assumed distribution. If they do not, the assumption is rejected. The formula (statistic) that uses the differences between "expected" and "observed" values to test the GoF follows a Chi-Square distribution. Hence, the name Chi-Square test.

In what follows we proceed as in Figure 1, using several data sets to fit a Normal, an Exponential, and a Weibull distribution. We will work with the same data sets used in the START sheets that discussed these empirical GoF procedures [7, 8, and 9]. In this way, the reader can compare the results for these two approaches and verify that they agree.

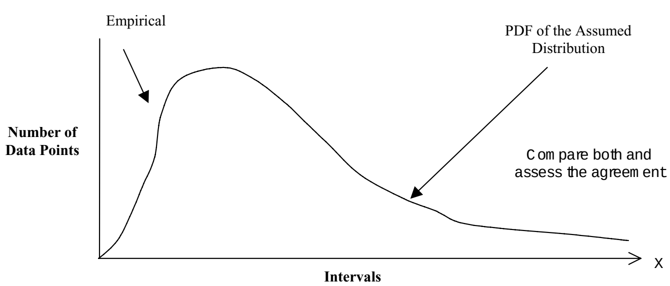

Figure 1. Area Goodness of Fit Test Conceptual Approach

The procedure is as follows:

1. Divide the data range of X into k subintervals.
2. Count the number of data points in each subinterval (histogram).
3. Superimpose the PDF of the assumed (theoretical) distribution.
4. Compare the empirical (histogram) with theoretical (PDF).
5. If they agree (probabilistically) the distribution assumption is supported by the data.
6. If they do not, the assumption is most likely incorrect.

The formula for the Chi-Square Statistic is:

$$\chi^2 = \Sigma_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} \sim \chi^2_{k\text{-}1\text{-}nep}$$

where

- e_i expected number of data points in cell i ($e_i \geq 5$)
- o_i actual (observed) number of data points in cell i;
- k total number of cells or subintervals in the range;
- n sample size for implementing the Chi-Square test ($n \geq 5*k$)
- k total number of cells or range subintervals
- k -1- No. Estimated Parameters (nep); Chi-Square degrees of freedom (DF>0)
- χ^2_γ is the Chi-Square distribution (table) with DF=γ

Fitting Normal and Lognormal Distribution

In the START sheet on empirically assessing the Normal and Lognormal distributions [8], we used the large data set shown in Table 1. We will now reassess it using the Chi-Square GoF test. We first obtain point estimations of the assumed Normal distribution parameters: mean and standard deviation shown in Table 2.

The point estimations allow us to define the composite distribution hypothesis: Normal (μ=19.5; σ=7.05). Since parameters mean and variance were estimated from the data (Table 2) the resulting Chi-Square statistic degrees of freedom are: DF=k-2-1 = No. of subintervals - No. of parameters estimated - 1 (with DF>0). We thus can safely select k = 5 subintervals.

Next, we select the following interval endpoints: 14, 17, 22, and 26 which, in turn, define five cells or subintervals, each of which contains more than the required five minimum expected observations (Figure 2).

In Table 3, we present the intermediate results for this Chi-Square GoF test example.

In the first column we show the endpoints of the intervals. In the second, we give the standardized endpoints: (endpoints-average) /Std-dev. In the third column, we give their cumulative values,obtained from the usual Normal tables. For example, for the first endpoint (14), then for the standardized (-0.78014) endpoint, we have:

$$P_{19.5,7.05}(14) = \text{Normal}\left(\frac{14 - 19.5}{7.05}\right) = \text{Normal}(-0.78014) = 0.2176$$

Table 1. Data for the Normal GoF Test (Sorted)

6.1448	6.6921	6.7158	7.7342	9.6818	12.3317	12.5535	13.0973	13.6704
14.0077	14.7975	15.3237	15.5832	15.7808	15.7851	16.2981	16.3317	16.8147
16.8860	17.5166	17.5449	17.9186	18.5573	18.8098	19.2541	19.5172	19.7322
21.9602	23.2046	23.2625	23.7064	23.9296	24.8702	25.2669	26.1908	26.9989
27.4122	27.7297	28.0116	28.2206	28.5598	29.5209	30.0080	31.2306	32.5446

Table 2. Descriptive Statistics

Variable	N	Mean	Median	StDev	Min	Max	Q1	Q3
Data	45	19.50	18.56	7.05	6.14	32.54	15.06	25.73

Table 3. Intermediate Values for the GoF Test for Normality

Row	IntEnd	StdEnd	CumProb	CellProb	Expect	Obsvd	(e-o)^2/e
1	14	-0.78014	0.217654	0.217654	9.7944	9	0.06443
2	17	-0.35461	0.361441	0.143787	6.4704	10	1.92535
3	22	0.35461	0.638559	0.277118	12.4703	9	0.96574
4	26	0.92199	0.821732	0.183173	8.2428	6	0.61024
5	Infin	Infin	0.999999	0.178300	8.0235	11	1.10420
			Totals	1.000000	45.0010	45	4.67000

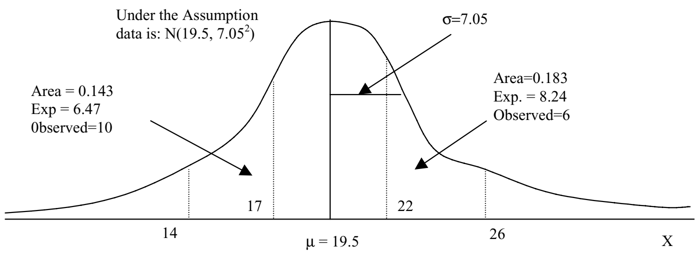

Figure 2. Representation of the Chi-Square GoF test for Normality

Then, we obtain in column four, the lagged differences of the Cumulative values, which constitute the individual cell "areas," under the assumed Normal (19.5, 7.05) PDF. We now multiply each cell "area" by the total sample size (n=45). Since each "area" is the probability that any sample element falls in the corresponding cell, this product yields the Expected number of elements (e) in each cell, according to the assumed distribution.

Then, we process the observed (o) and expected (e) values, of each cell, through the statistic:

$$\chi^2 = \sum_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} = 4.67 < \chi^2_{0.95,2} = 5.99$$

The Chi-Square statistic value (4.67) is smaller than the Chi-Square table value (5.99) for DF=5-2-1=2 and $1-\alpha=0.95$, so we can assume that the distribution of the population originating the data set is Normal (19.5, 7.05). Furthermore, we will be wrong less than 5% of the times. The process is summarized in Table 4.

Finally, if we want to fit a LogNormal distribution, we take the logarithm of the data and then implement the Table 4 procedure on these transformed data. If the original data is Lognormal, its logarithm is Normally distributed.

Table 4. Step-by-Step Summary of the Chi-Square GoF Test

1. Establish the Null Hypothesis H_0: Data is assumed Normal (μ; σ).
2. Estimate the Normal parameters from the data: μ=19.5; σ=7.05.
3. The test statistic is:

$$\chi^2 = \sum_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i}$$

4. Establish/Standardize the K=5 Subintervals (Figure 2).
5. Obtain Probability for the K=5 Subintervals (Table 3).
6. Test statistic distribution: Chi-Square; DF =5-2-1=2.
7. Establish test significance level (error): α=0.05.
8. Obtain Chi-Square critical value: 5.99.
9. Obtain Test Statistic value: 4.67.
10. As Critical Value > Test Statistic, assume Normality!

Fitting an Exponential Distribution

The large data set in Table 5 came from the same Exponential (θ=100) population that generated the sample in the START sheet on Exponential distribution assessment [7].

We will now assess the Exponentiality of the data via the Chi-Square GoF test, just like we did in the previous section for the Normal. We first obtain the descriptive statistics as shown in Table 6.

Analogously, this allows us to define the composite distribution hypothesis: Exponential (θ=100.2). Since we estimated the mean from the data (Table 6) the resulting Chi-Square has DF=k-1-1 and we can safely select k=5 and still have DF=5-2=3>0.

For endpoints we now select 30, 50, 95 and 160, which in turn, and just like before, define five subintervals. We also obtain the cumulative and individual cell probability values, as in the previous section. We illustrate it, for the first endpoint (30).

$$P_{100.2}(30) = 1.0 - \mathrm{Exp}\left(-\frac{30}{100.2}\right) = 1.0 - \mathrm{Exp}(-0.2994) = 0.25874$$

Table 5. Data for the Exponential GoF Test (Sorted)

5.142	16.344	17.150	18.325	22.473	25.789	25.928	26.230	29.153
32.264	35.138	35.387	41.743	42.374	43.388	46.975	47.246	51.309
53.628	56.689	60.392	74.860	76.610	77.456	93.350	94.216	95.831
103.956	111.403	117.269	118.441	121.334	122.694	128.675	136.434	168.727
172.222	213.474	213.889	215.220	221.943	229.777	235.789	281.492	351.505

Table 6. Descriptive Statistics

Variable	N	Mean	Median	StDev	Min	Max	Q1	Q3
Data	45	100.2	76.6	81.8	5.1	351.5	35.3	132.6

The resulting values, equivalent to those in Table 3, are shown in Table 7.

The result of the Chi-Square GoF test statistic for this data set and assumption is:

$$\chi^2 = \sum_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} = 1.46 < \chi^2_{0.95,3} = 7.81$$

Here, like before, the Chi-Square statistic value (1.46) is not larger than the Chi-Square table value (7.81) for DF=5-2=3 and 1–α=0.95. We can then assume H_0: that the distribution of the population originating the data set is Exponential with θ =100.2. Furthermore, we will be wrong less than 5% of the times.

The entire GoF process, for this case, is summarized in Table 8.

Table 7. Intermediate Values for the Exponential GoF t\Test

Row	IntEnd	CumProb	CellProb	Expected	Observed	(e-o)^2/e
1	30	0.25874	0.258738	11.6432	9	0.600055
2	50	0.39286	0.134126	6.0357	8	0.639309
3	95	0.61252	0.219661	9.8848	9	0.079192
4	160	0.79746	0.184933	8.3220	9	0.055242
5	Infin	1.00000	0.202500	9.1125	10	0.086437
		Totals	1.000000	49.9980	45	1.460200

Table 8. Step-by-Step Summary of the Chi-Square GoF Test

1. Establish the Null Hypothesis H_0: Data is assumed Exponential (θ).
2. Estimate the Exponential parameter from the data: θ = 100.2.
3. The test statistic is:

$$\chi^2 = \sum_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i}$$

4. Establish the K=5 Subintervals (no need to standardize).
5. Obtain Probability for the K=5 Subintervals (Table 7).
6. Test statistic distribution: Chi-Square; DF =5-1-1=3.
7. Establish significance level (error): α=0.05.
8. Obtain Chi-Square critical value: 7.81.
9. Obtain Test Statistic value: 1.46.
10. As Critical Value > Test Statistic, assume Exponentiality!

Fitting a Weibull Distribution

In the START sheet on empirically assessing the Weibull distribution [9], we used the large data set as shown in Table 9. We now assess whether the data is Weibull via the Chi-Square GoF test, just like we did for the Normal and Exponential data. We first obtain the descriptive statistics as shown in Table 10.

To obtain the Weibull parameter estimators, we can use Weibull paper [1, 2], or regress the following equation.

$$\ln\left\{\ln\left(\frac{1}{1-p_x}\right)\right\} \text{ versus } \ln(x); \text{ where } p_x = \frac{\text{Rank}(x) - 0.3}{n - 0.4}$$

The regression equation for the present case is:

Predictor	Coef	StDev	T	P
Constant	-3.40715	0.06856	-49.69	0.000
C1	1.35424	0.03008	45.02	0.000

C2 = - 3.41 + 1.35 C1

S = 0.1774 R-Sq = 97.9% R-Sq(adj) = 97.9%

The regression slope (1.35) is the Weibull Shape Parameter; the Weibull Characteristic Life (CharLf) is obtained by:

CharLf = Exp(-(Intercept/Slope)) = Exp(-(-3.41/1.35)) = 12.378

The parameter estimators allow us to define the composite distribution hypothesis H_0: Weibull (α=12.378; β=1.354). Since we estimated both of them from the data the resulting Chi-Square has DF=k-2-1. We can select k=5 and still have: DF=5-3=2>0.

For endpoints we now select 3.9, 7.8, 12.3, and 17.4, which, in turn and just like before, define five subintervals. We also obtain the cumulative and individual cell probability values, as in the previous section. We illustrate it, for the first endpoint (3.9):

$$P_{\alpha=12.38;\beta=1.35}(3.9) = 1 - \exp\left\{-\left(\frac{x_i}{\alpha}\right)^{\beta}\right\} =$$

$$1 - \exp\left\{-\left(\frac{3.9}{12.38}\right)^{1.35}\right\} = 1 - 0.8103 = 0.1896$$

Table 9. Data for the Weibull GoF Test (Sorted)

0.8997	1.2838	1.5766	1.8627	2.4193	2.4353	3.1520	3.3367	3.4850
3.9605	3.9921	3.9934	4.1013	4.8306	5.3545	5.6094	7.7829	7.8240
8.3431	9.0248	9.2627	9.2766	9.7943	11.4391	12.2847	12.4112	13.1651
13.4990	13.5532	14.1542	14.4694	14.5857	15.1603	15.6962	15.7833	17.4998
18.1497	18.6342	19.4354	19.7557	19.9496	22.5383	23.8066	29.9006	34.0658

Table 10. Descriptive Statistics of Data in Table 9

Variable	N	Mean	Median	StDev	Min	Max	Q1	Q3
WeibSamp	45	11.19	9.79	7.85	0.9	34.07	3.99	15.74

The resulting values, equivalent to those in Table 3, are shown in Table 11. The result of the Chi-Square GoF test statistic for this data set and assumption is:

$$\chi^2 = \Sigma_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} = 1.68 < \chi^2_{0.95,2} = 5.99$$

Here, like before, the Chi-Square statistic value (1.68) is not larger than the Chi-Square table value (5.99) for DF=2 and 1–α=0.95. We can then, as done before, assume H_0: that the distribution of the population is Weibull (α=12.37; β=1.35). Furthermore, we will be wrong less than 5% of the times. The entire process is summarized in Table 12.

Table 11. Intermediate Values for the Weibull GoF Test

Row	IntEnd	CumProb	CellProb	Expected	Observed	(e-o)^2/e
1	3.9	0.18963	0.189629	8.5333	9	0.02522
2	7.8	0.41491	0.225279	10.1376	8	0.450721
3	12.3	0.62890	0.213992	9.6296	8	0.275787
4	17.4	0.79471	0.165807	7.4613	10	0.863791
5	Infin	1.00000	0.205300	9.2385	10	0.062768
		Totals	1.000000	45.0000	45	1.678600

Table 12. Step-by-Step Summary of the Chi-Square GoF Test

1. Establish the Null Hypothesis H_0: Data is assumed Weibull (α, β).
2. Estimate the Weibull parameters from the data: α=12.378; β=1.354.
3. The test statistic is:

$$\chi^2 = \Sigma_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i}$$

4. Establish the K=5 Subintervals (no need to standardize).
5. Obtain Probability for the K=5 Subintervals (Table 11).
6. Test statistic distribution: Chi-Square; DF =5-1-2=2.
7. Establish significance level (error): α=0.05.
8. Obtain Chi-Square critical value: 5.99.
9. Obtain Test Statistic value: 1.68.
10. As Critical Value > Test Statistic, we assume Weibull!

A Counter Example

For completion, we now develop an example where the data does not fit the hypothesized distribution. We do that using the Exponential data (Table 5). We will now assume these data come from a Normal distribution and will use the descriptive statistics in Table 6, to establish the hypothesized Normal parameters: μ=100.2; σ=81.8.

For endpoints we select 30, 80, 120, and 170 which, in turn and just like before, define five subintervals. Since we estimated the two Normal parameters from the data, the resulting Chi-Square statistic has DF=5-3=2>0. We obtain cumulative and individual cell probability values, as done in the previous sections. Results are shown in Table 13.

$$\chi^2 = \Sigma_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} = 5.82 \approx \chi^2_{0.95,2} = 5.99$$

We observe: (1) a large discrepancy between Observed and Expected values (15 and 9.31) in Cell 2 and (2) test statistic value (5.82) is very close to the critical Chi-Square table value (5.99). This shows that the assumption of Normality is not well supported by the data. We repeat the GoF test for seven cells (expected value per cell 45÷7=6.3>5 and DF=7-2-1=4>0). Results, showing the data are not Normal(100.2,81.8) are in Table 14.

$$\chi^2 = \Sigma_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} = 27.97 > \chi^2_{0.95,4} = 9.487$$

Table 13. Intermediate Values for the GoF Test With Five Cells

Row	CounterEx	CumProb	CellProb	Expected	Observed	(e-o)^2/e
1	30	0.19539	0.195394	8.79271	9	0.00489
2	80	0.40248	0.207082	9.31871	15	3.46369
3	120	0.59563	0.193155	8.69196	7	0.32935
4	170	0.80325	0.207623	9.34304	5	2.01883
5	Infinit	1.00000	0.196746	8.85358	9	0.00242
		Totals	1.000000	45.00000	45	5.81920

Table 14. Intermediate Values for the GoF Test With Seven Cells

Row	CountEx	CumProb	CellProb	Expected	Observed	(e-o)^2/e
1	15	0.14881	0.148807	6.69631	1	4.8456
2	50	0.26971	0.120903	5.44062	16	20.4941
3	85	0.42629	0.156584	7.04628	7	0.0003
4	115	0.57179	0.145495	6.54727	6	0.0457
5	150	0.72867	0.156884	7.05977	6	0.1591
6	185	0.85006	0.121384	5.46228	2	2.1946
7	Infin	1.00000	0.149944	6.74747	8	0.2325
		Totals	1.000000	45.00000	45	27.9720

Summary

In this START sheet, we have discussed the important concept of Goodness of Fit assessment of statistical distributions, for large samples, via the Chi-Square test. We have provided several numerical and graphical examples for the Normal, Lognormal, Exponential, and Weibull distributions, relevant in reliability and maintainability studies. We also have discussed some related theoretical and practical issues, providing several references to background information and further readings.

The small sample Goodness of Fit problem cannot be dealt with via the Chi-Square test. For the number of observations per cell is too small for the GoF test statistic to converge to its Chi-Square underlying distribution. In such cases, we use other, CDF-based distance Goodness of Fit tests, such as the Anderson-Darling and Kolmogorov-Smirnov. Due to their complexity, these tests are treated in more detail in separate START sheets.

Bibliography

1. Practical Statistical Tools for Reliability Engineers, Coppola, A., RAC, 1999.
2. Mechanical Applications in Reliability Engineering, Sadlon, R.J., RAC, 1993.
3. A Practical Guide to Statistical Analysis of Material Property Data, Romeu, J.L. and C. Grethlein, AMPTIAC, 2000.
4. Probability and Statistics for Engineers and Scientists, Walpole, Myers, and Myers. Sixth Edition, Prentice Hall, NJ, 2001.
5. Statistical Confidence, Romeu, J.L., RAC START, Volume 9, Number 4, http://rac.alionscience.com/pdf/NLDIST.pdf.
6. Graphical Comparison of Two Populations, Romeu, J.L., RAC START, Volume 9, Number 5, http://rac.alionscience.com/pdf/2POP.pdf.
7. Statistical Assumptions of an Exponential Distribution, Romeu, J.L., RAC START, Volume 8, Number 2, http://rac.alionscience.com/pdf/E_ASSUME.pdf.
8. Empirical Assessment of Normal and Lognormal Distribution Assumptions, Romeu, J.L., RAC START, Volume 9, Number 6, http://rac.alionscience.com/pdf/NLDIST.pdf.
9. Empirical Assessment of Weibull Distribution Assumptions, Romeu, J.L., RAC START, Volume 10, Number 3.

About the Author

Dr. Jorge Luis Romeu has over thirty years of statistical and operations research experience in consulting, research, and teaching. He was a consultant for the petrochemical, construction, and agricultural industries. Dr. Romeu has also worked in statistical and simulation modeling and in data analysis of software and hardware reliability, software engineering and ecological problems.

Dr. Romeu has taught undergraduate and graduate statistics, operations research, and computer science in several American and foreign universities. He teaches short, intensive professional training courses. He is currently an Adjunct Professor of Statistics and Operations Research for Syracuse University and a Practicing Faculty of that school's Institute for Manufacturing Enterprises.

For his work in education and research and for his publications and presentations, Dr. Romeu has been elected Chartered Statistician Fellow of the Royal Statistical Society, Full Member of the Operations Research Society of America, and Fellow of the Institute of Statisticians.

Romeu has received several international grants and awards, including a Fulbright Senior Lectureship and a Speaker Specialist Grant from the Department of State, in Mexico. He has extensive experience in international assignments in Spain and Latin America and is fluent in Spanish, English, and French.

Romeu is a senior technical advisor for reliability and advanced information technology research with Alion Science and Technology Corporation. Since joining Alion in 1998, Romeu has provided consulting for several statistical and operations research projects. He has written a State of the Art Report on Statistical Analysis of Materials Data, designed and taught a three-day intensive statistics course for practicing engineers, and written a series of articles on statistics and data analysis for the AMPTIAC Newsletter and RAC Journal.

Other START Sheets Available

Many Selected Topics in Assurance Related Technologies (START) sheets have been published on subjects of interest in reliability, maintainability, quality, and supportability. START sheets are available on-line in their entirety at <http://rac.alionscience.com/DATA/START>.

For further information on RAC START Sheets contact the:

Reliability Analysis Center
201 Mill Street
Rome, NY 13440-6916
Toll Free: (888) RAC-USER
Fax: (315) 337-9932

or visit our web site at:

<http://rac.alionscience.com>

About the Reliability Analysis Center

The Reliability Analysis Center is a Department of Defense Information Analysis Center (IAC). RAC serves as a government and industry focal point for efforts to improve the reliability, maintainability, supportability and quality of manufactured components and systems. To this end, RAC collects, analyzes, archives in computerized databases, and publishes data concerning the quality and reliability of equipments and systems, as well as the microcircuit, discrete semiconductor, and electromechanical and mechanical components that comprise them. RAC also evaluates and publishes information on engineering techniques and methods. Information is distributed through data compilations, application guides, data products and programs on computer media, public and private training courses, and consulting services. Located in Rome, NY, the Reliability Analysis Center is sponsored by the Defense Technical Information Center (DTIC). Alion, and its predecessor company IIT Research Institute, have operated the RAC continuously since its creation in 1968. Technical management of the RAC is provided by the U.S. Air Force's Research Laboratory Information Directorate (formerly Rome Laboratory).