AD-A190 172 RESULTS OF INAPPROPRIATE EEP (ELLIPTIC ERROR PROBABLE) 1/1
NORMALIZATION METHODS IN CORRELATION(U) JET PROPULSION
LAB PASADENA CA N W RENNIE 17 SEP 85 JPL-D-4299
UNCLASSIFIED NAS7-918 F/G 12/3 NL

MICROCOPY RESOLUTION TEST CHART

(2)

7057-75

U.S. ARMY INTELLIGENCE CENTER AND SCHOOL
SOFTWARE ANALYSIS AND MANAGEMENT SYSTEM

RESULTS OF INAPPROPRIATE EEP NORMALIZATION METHODS IN CORRELATION

TECHNICAL MEMORANDUM NO. 11

MARC

Mathematical Analysis Research Corporation

17 September 1985

National Aeronautics and
Space Administration

JPL

JET PROPULSION LABORATORY
California Institute of Technology
Pasadena, California

JPL D-4299
ALGO_PUB_0025

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER ALGO-PUB-0025	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Technical Memo 11, "Results of Inappropriate EEP Normalization Methods in Correlation".		5. TYPE OF REPORT & PERIOD COVERED FINAL 6. PERFORMING ORG. REPORT NUMBER D-4299
7. AUTHOR(s) Mathematical Analysis Research Corp.		8. CONTRACT OR GRANT NUMBER(s) NAS7-918
9. PERFORMING ORGANIZATION NAME AND ADDRESS Jet Propulsion Laboratory ATTN: 171-209 California Institute of Technology 4800 Oak Grove, Pasadena, CA 91109		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS RE 182 AMEND #187
11. CONTROLLING OFFICE NAME AND ADDRESS Commander, USAICS ATTN: ATSI-CD-SF Ft. Huachuca, AZ 85613-7000		12. REPORT DATE 17 Sep 85 13. NUMBER OF PAGES 5
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office) Jet Propulsion Laboratory, ATTN: 171-209 California Institute of Technology 4800 Oak Grove, Pasadena, CA 91109		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE NONE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Dissemination

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Prepared by Jet Propulsion Laboratory for the US Army Intelligence Center and School's Combat Developer's Support Facility.

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

EEP, Ellipse, F-Statistic, Chi Square Statistic, Variance, Covariance, Sample Size, Confidence Level

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

When computing elliptic error probable (EEP) ellipse axes for location error with a given confidence level, the chi square statistics tables are used if the standard deviation is known. If the standard deviation is estimated from the data, then the F-statistics tables are used. This memo examined the error caused by use of the chi square statistics tables in place of the F-statistics tables. This error is quantified in terms of the number of LOBs used to obtain the location estimate.

7057-75

U.S. ARMY INTELLIGENCE CENTER AND SCHOOL
Software Analysis and Management System

Results of Inappropriate EEP Normalization Methods in Correlation

Technical Memorandum No. 11

17 September 1985

Author:

MARC
Mathematical Analysis Research Corporation

Approval:

James W. Gillis, Subgroup Leader
Algorithm Analysis Subgroup

Edward J. Records, Supervisor
USAMS Task

Concur:

A. F. Ellman, Manager
Ground Data Systems Section

Fred Vote, Manager
Advanced Tactical Systems

JET PROPULSION LABORATORY
California Institute of Technology
Pasadena, California

JPL D-4299

PREFACE

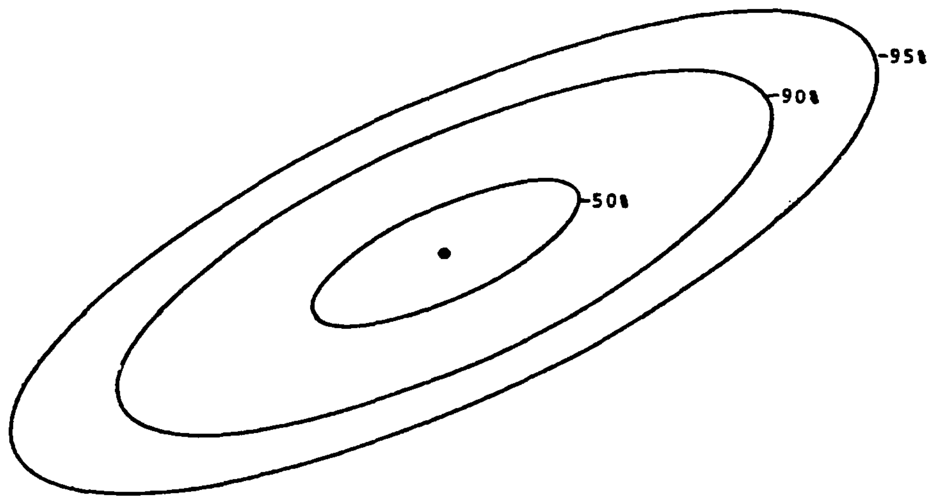
The work described in this publication was performed by the Mathematical Analysis Research Corporation (MARC) under contract to the Jet Propulsion Laboratory, an operating division of the California Institute of Technology. This activity is sponsored by the Jet Propulsion Laboratory under contract NAS7-918, RE182, A187 with the National Aeronautics and Space Administration, for the United States Army Intelligence Center and School.

Results of Inappropriate EEP Normalization Methods in Correlation

INTRODUCTION

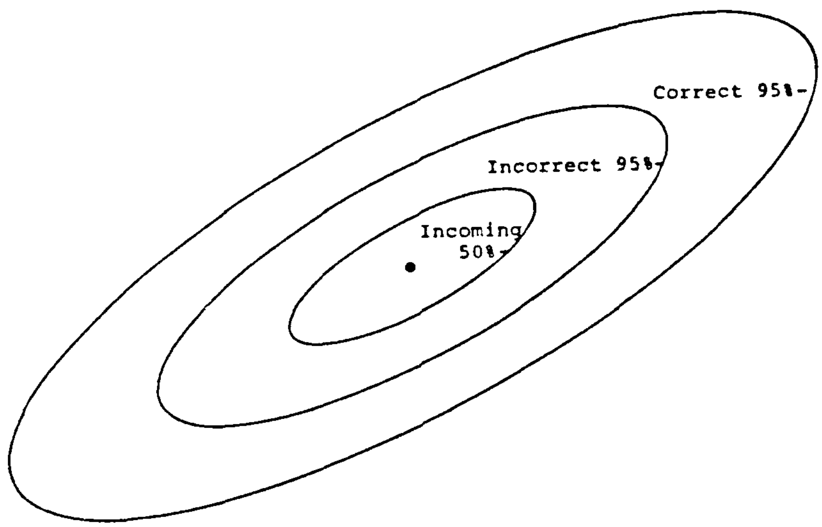
We assume that data arrives in terms of ellipses: the location of the center of the ellipse and the length and orientation of axes of the ellipse.

Ellipses are to be 95% ellipses, i.e. the probability that the specified ellipse contains a given emitter is .95 (elliptical error probable = EEP is 95%). If the incoming ellipse data are not for a 95% ellipse but are for a say 50% ellipse then the incoming data must be transformed. This transformation of incoming data only affects the length of the axes (not the center of the ellipse or orientation.) If the transformation is from a 50% to a 95% ellipse then the axes of the incoming ellipse are lengthened.



The inner most ellipse is a 50% ellipse.

The transformation from a non-95% ellipse depends on whether the incoming ellipse size was computed using a χ^2 value or using an F value*. The conversion algorithms presently used are χ^2 values regardless of how the incoming ellipse size was computed. When the incoming ellipse size is based on the F then the conversion is incorrect. That is, the converted ellipse is too small. The amount of error depends on the sample size used for the incoming ellipse. The smaller the sample size the greater the error.



This conversion error effects:

1) The accuracy of the test which determines whether or not to accept the incoming data as coming from an emitter already located in the data base.

2) The accuracy of the combination algorithm which combines the incoming ellipse with an ellipse already in the data base.

Footnote: In general a χ^2 value is used if the variance-covariance of the data is known and an F value is used if the variance-covariance of the data is estimated.

Under 1)

The error is too frequently stating that the incoming data do not come from a specific emitter when in fact they do.

Under 2)

There are two possible errors.

The resultant ellipse being too small.

The location of the center of the resultant ellipse being overly affected by data based on small sample sizes.

GENERAL

The F distribution is used in determination of EEPs (elliptical error probable) for some programs such as Guardrail. Ellipse confidence level conversion algorithms, combination algorithms and combination testing algorithms assume that EEPs are based on the Chi-square distribution. This memo is not concerned with all the problems that result from this discrepancy for these three types of algorithms. It is concerned with the direct impact on the confidence level conversion algorithm and with the implications of this impact for the combination and testing algorithms.

When the original distribution underlying EEPs is F, converting ellipse confidence coefficients to the 95% level is an approximation that is only valid for 'large' sample sizes. This memo is designed to illustrate the impact of using this approximation with small sample sizes. The ideas introduced include:

i) The interpretation of the conversion of confidence level as a rescaling of the size of the confidence ellipse.

ii) The factors affecting the scaling constant, specifically,
(a) the confidence level of the incoming ellipse
(b) whether the original distribution is Chi-square or F
(c) sample size (but only if the original distribution is F)

iii) The amount of error in ellipse size that can occur.

iv) What the difference between chi-square and F is supposed to represent.

v) The effect of scaling errors of the type being examined here on correlation testing. (And questions concerning use of that test when F is the basis for formation of the ellipse.)

vi) The effect of the scaling errors being studied here on the point estimate location determined by the combination algorithm. (And questions concerning use of this algorithm when F is the basis for formation of the ellipse.)

vii) The effect of the scaling errors being studied here on the EEP size of the resultant ellipse using the combination algorithm.

When incoming EEPs are based on sample sizes of 5 or smaller these considerations are very important.

The concepts outlined above will be discussed in the sections below. They are also illustrated in the graphs that follow, and may be pursued further using the tables attached in the appendices. (Other graphs concerning ellipse combination and testing for combination may easily be imagined in terms of the geometric characterization of the testing and combination algorithms.)

I. INTERPRETATION OF CONFIDENCE LEVEL CONVERSION AS RESCALING EEP'S

Changing confidence level may be thought of as scaling of the EEP (size of a confidence ellipse.) In fact given two confidence levels with specified confidence methodologies and sample size (for the F) it is possible to list the scaling factor. For example, if sample size is 5 (and assuming our applications have 2 spatial degrees of freedom) then a 90% EEP based on the F distribution is approximately 2.5 times as large as the corresponding 50% EEP. The 95% EEP approximately 3.3 times as large as the corresponding 50% EEP. This ellipse size ratio is illustrated in Graph 1 that follows.

The shape, location, and orientation of the base ellipse do not affect the scaling factor. The scaling factor does depend on which distribution and sample size is used for the base ellipse and converted ellipse, however. Because correlation algorithms assume 95% confidence levels the tables in the appendix assume that the converted ellipse has this confidence level. (The Ellipse Radius Ratios listed at the top of each table are Chi-square to Chi-square and the entry in the table are F to F). Because of this 95% convention the 2.5 scaling factor for sample size 5 between the 50% F and 90% F EEPs is not listed in the table. The 50% F to 95% scaling factor (for sample size 5) of 3.3 (3.292) is listed in this table, however.

II. FACTORS AFFECTING THE SCALING FACTOR

A. The Confidence Level of the Incoming Ellipse.

If the incoming ellipse is based on a 95% confidence level bound then no conversion is necessary.

If the incoming ellipse is based on a 90% confidence level then the scaling factor will be bigger than one and the resultant ellipse will be bigger. Examination of the 90% F radius ratios in tables in the Appendix confirms this. (As do the conversion factors for the Chi-square conversions which are listed above the tables.)

If the incoming ellipse is based on a 50% confidence level then the scaling factor for conversion will be even bigger than if the incoming ellipse had been a 90% confidence ellipse. This can be seen both in Graph 1 and by comparison of 50% radius ratios with 90% radius ratios in the appendix.

B. The Distribution (Chi-square or F) Underlying the EEPs.

Chi-square scaling factors are closer to 1 than F scaling factors. Note that the columns of F cutoff values are all bigger than the corresponding Chi-square cut-off value listed above the table.

C. The Sample Size.

Sample size affects the F scaling factors as can be seen in the tables in the appendix and in Graph 4. The tables and Graph 4 also illustrate that as sample size increases the F converges toward the Chi-square.

Sample size doesn't affect the Chi-square scaling factors. This is probably why sample size has not been sent to the correlation algorithm in the past.

D. The Degrees of Freedom

This factor is usually based on spatial dimension and hence is fixed in most applications and will not be discussed in detail here.

III. THE AMOUNT OF ERROR IN ELLIPSE SIZE THAT CAN OCCUR

This section is concerned about the difference between what a person sitting at a scope would observe and what he should be observing. The assumption is that a conversion has been done to make the ellipse a 95% ellipse but that that conversion was done using Chi-square scaling factors.

Of course, if the ellipses were based on the Chi-square distribution no error is made. But what if the distribution underlying the incoming ellipse was F?

If the incoming ellipse was a 95% F there is still no error as the conversion scale is 1. If, however, the incoming distribution wasn't at 95% then there is an error because a Chi-square radius ratio was used rather than an F radius ratio. The ratio between the Chi-square radius ratio and the F radius ratio is the natural measure of this error. This information is provided in the tables in the appendix under the unfortunate column title of CHI-F RATIO RATIO. The ratios depend on the confidence level so there are actually two columns. The entries in these columns are measures of the conversion error. Graph 2 might clarify these concepts. The inner ellipse in the graph is assumed to be a sample size 5, 2 degrees of freedom, 50% F incoming ellipse. This inner ellipse is probably not actually seen by the operator during correlation. The middle ellipse is what a chi-square conversion would construct and presumably report to the operator. The outer ellipse is the correctly converted ellipse.

Note that the difference is in Graph 2 is significant and examination of the tables in the appendix make it clear that even more significant differences exist. The larger RATIO RATIO entries reflect larger errors. This is illustrated in Graph 3 for which the 90-95% based RATIO RATIO is 1.1603 and the 50-95% based RATIO RATIO is 1.5860. This example and examination of the tables also show that the 50% to 95% conversion error is always larger than the 90-95% error, as one would expect. Graph 4 illustrates how this error decreases with sample size. In the limit, as sample size increases, the RATIO RATIO would approach one and there would be no significant error.

IV. WHAT DOES THE DIFFERENCE BETWEEN F AND CHI-SQUARE SCALING MEAN

The F based EEPs are bigger chi-square EEPs because chi-square assumes that the amount of error that one is subject to is known, whereas the F assumes the amount of error that one is subject to is unknown and only determined as data comes in and variation is found. The truth probably lies somewhere between these two assumptions. The fact that the F assumes an extra type of uncertainty (the amount of error that one is subject to) implies that it yields bigger EEPs. The higher the confidence level the more this uncertainty is reflected and hence the RATIO RATIOs get bigger as the difference between the base confidence level and the resultant increases.

The dependence on sample size is more problematical, however. The authors of this report suspect that unmodeled sources of error would prevent the uncertainty in ellipse size from going away to the extent it does as sample size increases using the F test.

V. THE IMPACT OF SCALING ERRORS ON TESTING FOR COMBINATION

The nature of scaling errors is that incoming ellipses appear to be smaller than they really are. This means that the incoming ellipse is more likely to overlap the base ellipse by enough to accept (i.e. there is less acceptance the bigger the RATIO RATIO Value.) This is illustrated in Graph 5 where the incorrectly converted ellipse rejects and the correctly converted ellipse accepts. (Examples may be constructed geometrically bearing in mind that non-intersecting ellipses reject and that if the center of one ellipse is in the other ellipse then the acceptance test will reject.)

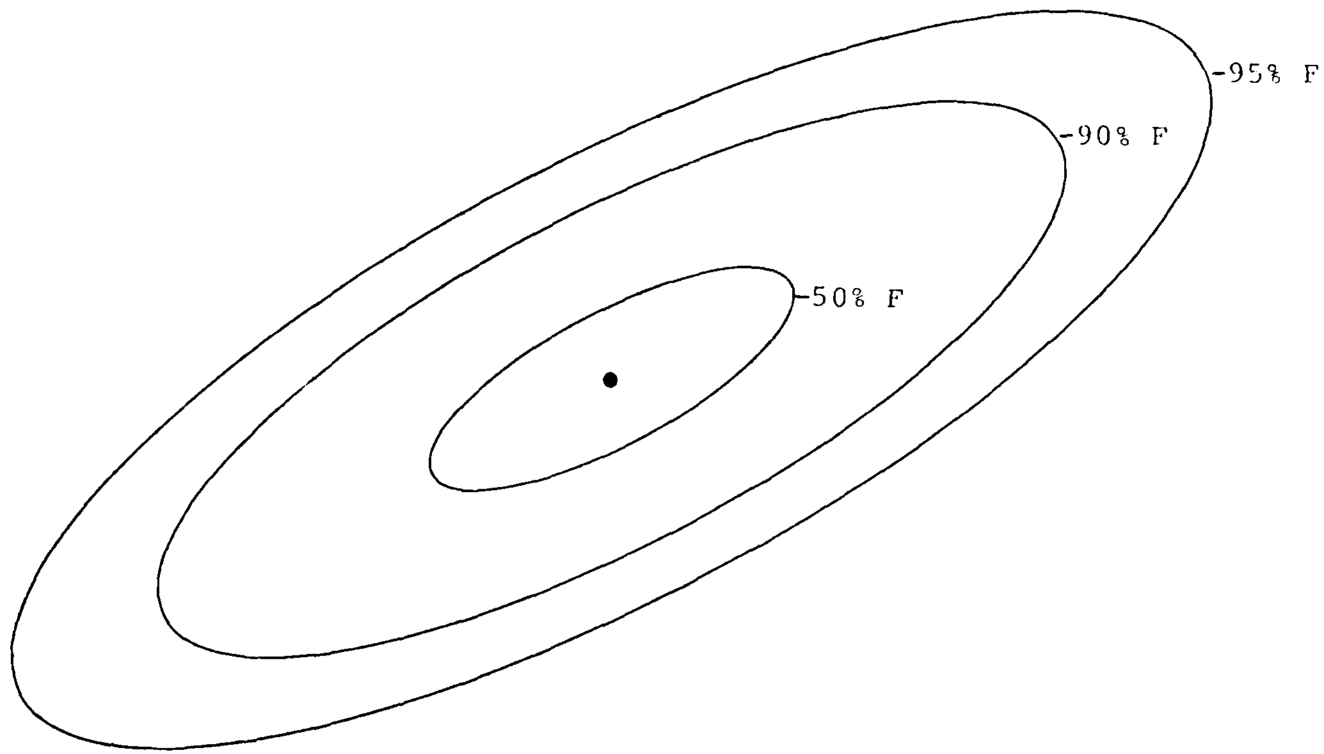
VI. THE IMPACT OF SCALING ERRORS ON DETERMINING A POINT ESTIMATE

The resultant point estimate is a type of weighted average between the two original point estimates (although the result is not necessarily actually on the line segment between the points.) The weights are based on ellipse "smallness." Changing ellipse size changes the weighting. The ellipse that is corrected to a larger size is weighted less and the location estimate will be adjusted away from the point estimate corresponding to that ellipse. Examination of graphs 6, 7, and 8 in sequence illustrate this point.

VII. THE IMPACT OF SCALING ERRORS ON DETERMINING A POINT ESTIMATE

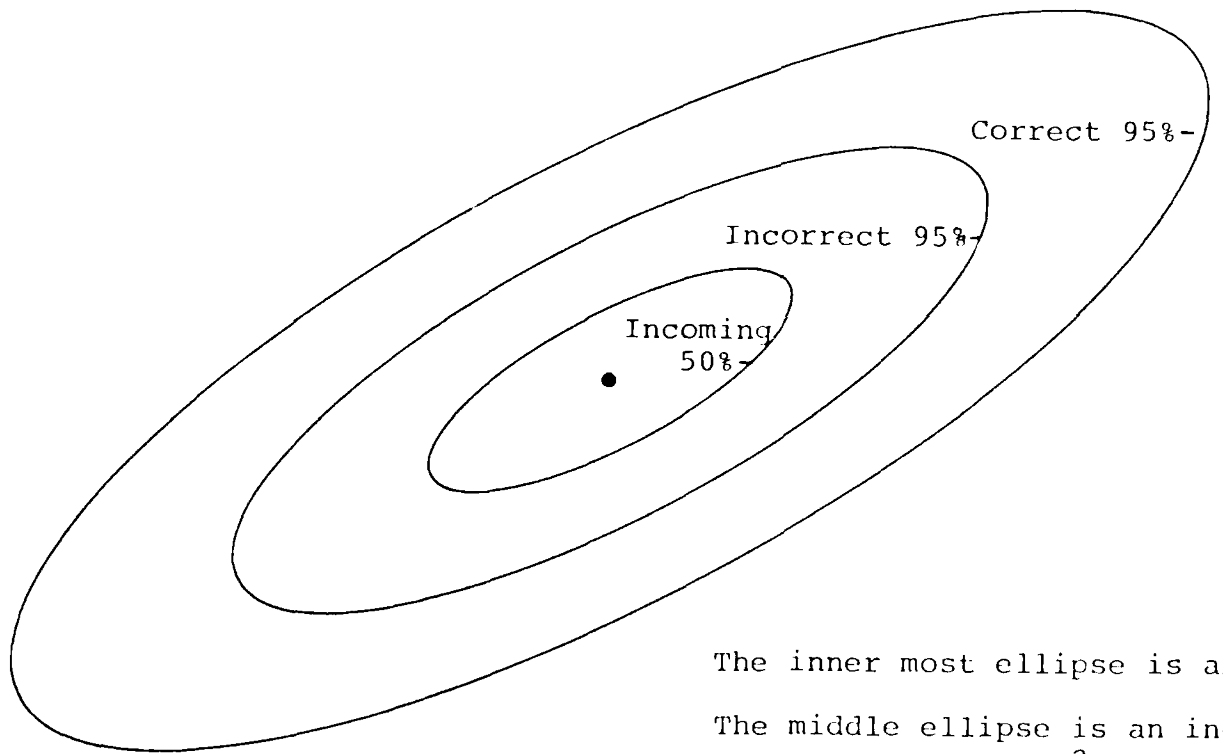
The resultant ellipse size is based on the ellipse sizes of the two input ellipses. Bigger input yields a bigger output. See Graphs 6, 7, and 8 for an example. See MARC's report, "Testing and Combination of Confidence Ellipses: A Geometric Analysis", for more details about the relationship of the geometry of input ellipses and output ellipses.

GRAPH 1



The inner most ellipse is a 50% ellipse obtained using an F distribution with sample size n=5. The outer two ellipses are the correctly converted ellipses using ratio's of F cut-offs.

GRAPH 2

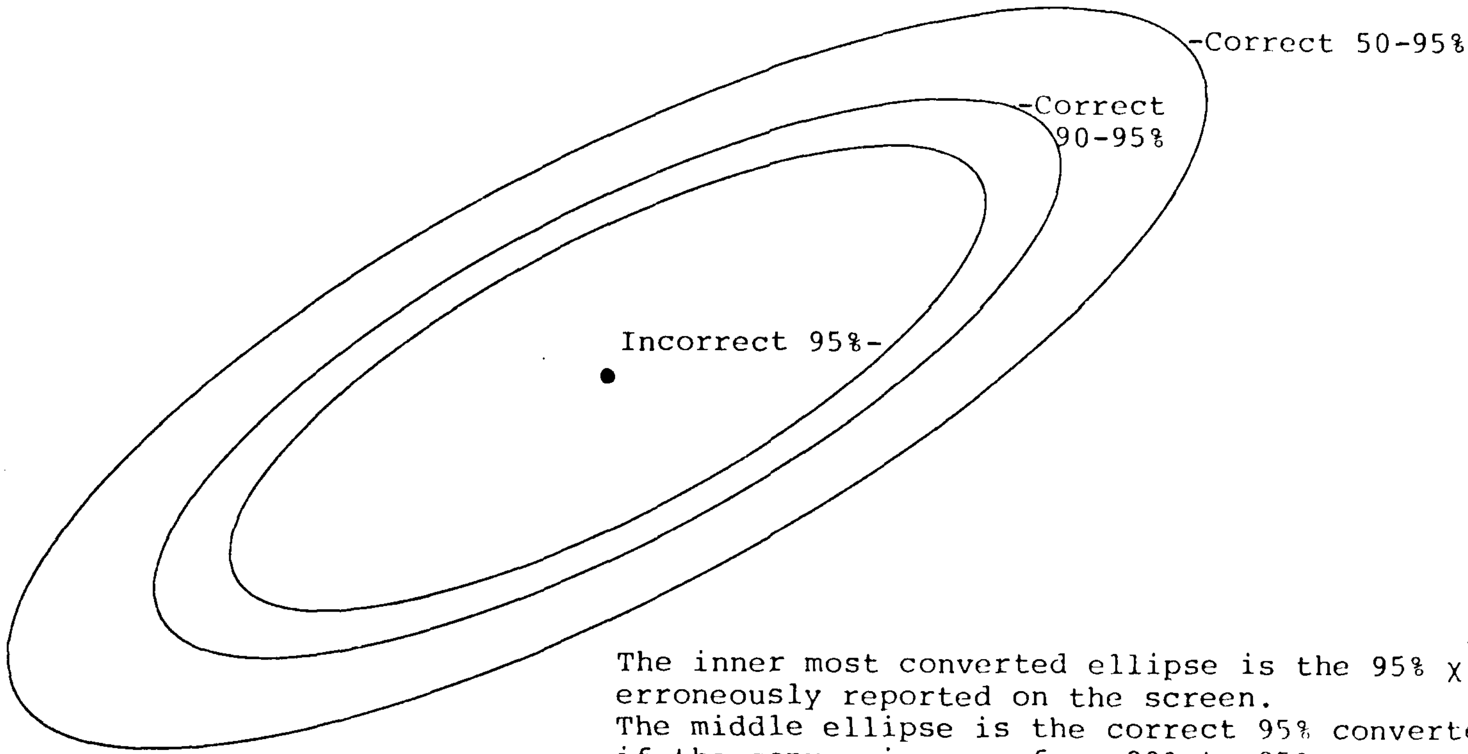


The inner most ellipse is an incoming 50% ellipse.

The middle ellipse is an incorrect conversion to a 95% ellipse using χ^2 cut-offs.

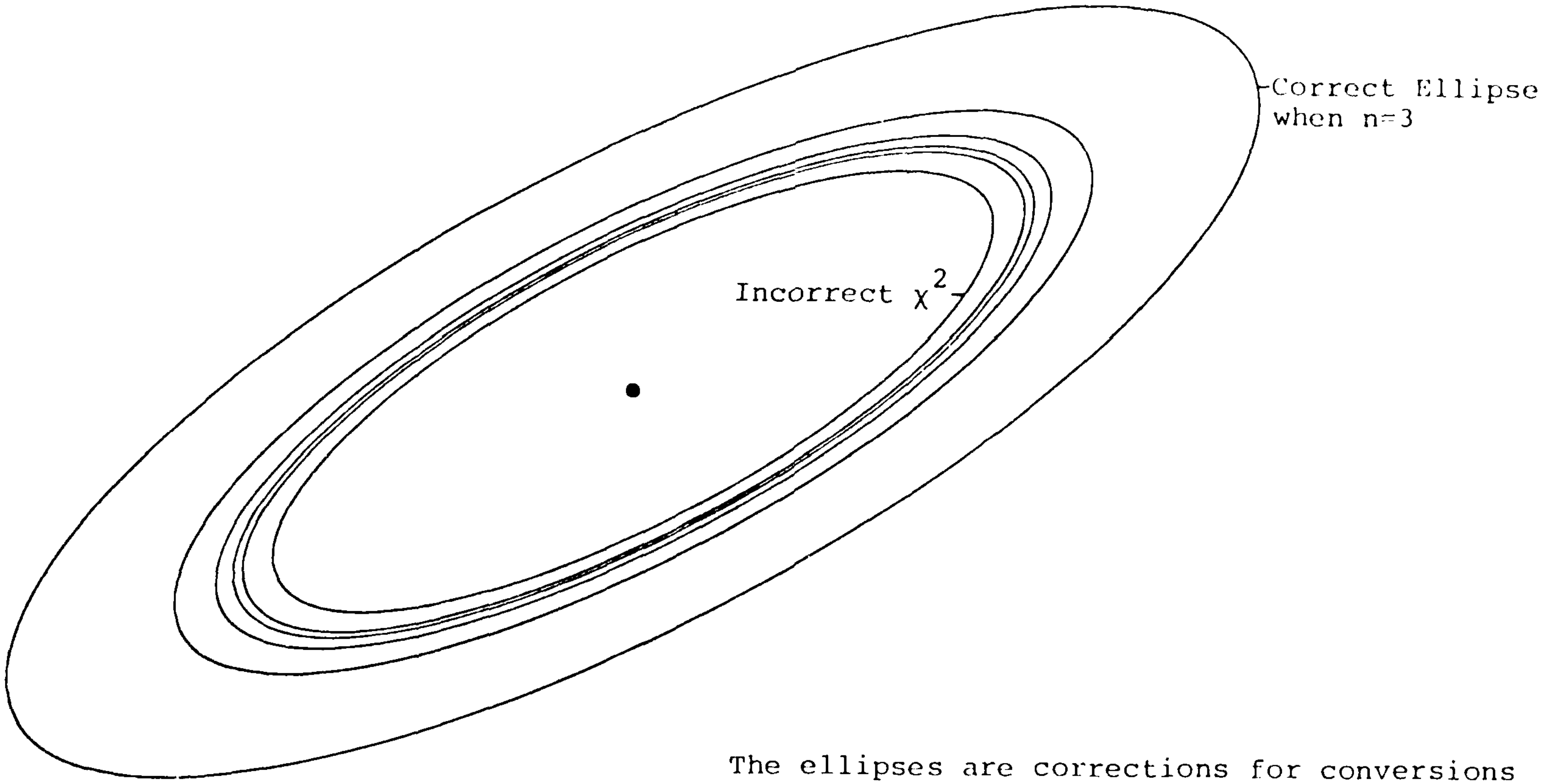
The outer most ellipse is a correct conversion to a 95% ellipse using Γ cut-offs. Sample size is 5.

GRAPH 3



The inner most converted ellipse is the 95% χ^2 ellipse erroneously reported on the screen.
The middle ellipse is the correct 95% converted ellipse if the conversion was from 90% to 95%.
The outer most ellipse is the correct 95% ellipse if the conversion was from 50% to 95%. Sample size is $n = 5$.

GRAPH 4

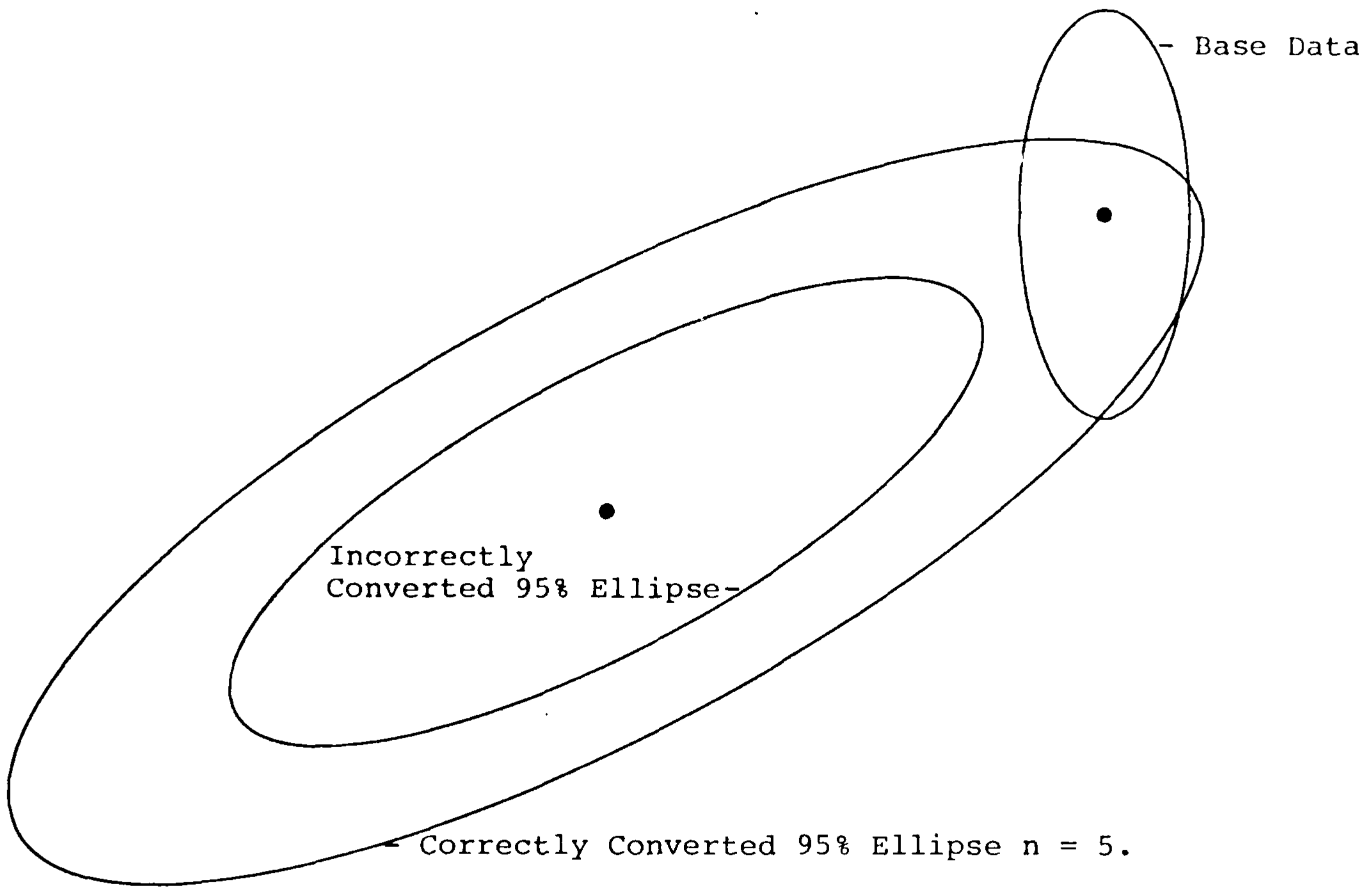


The ellipses are corrections for conversions from 90% to 95% ellipses.

The inner most ellipse is the χ^2 ellipse erroneously reported on the screen.

The outer 5 ellipses represent the correct conversion ellipse if the sample size was n = 3,4,5,6, or 7.

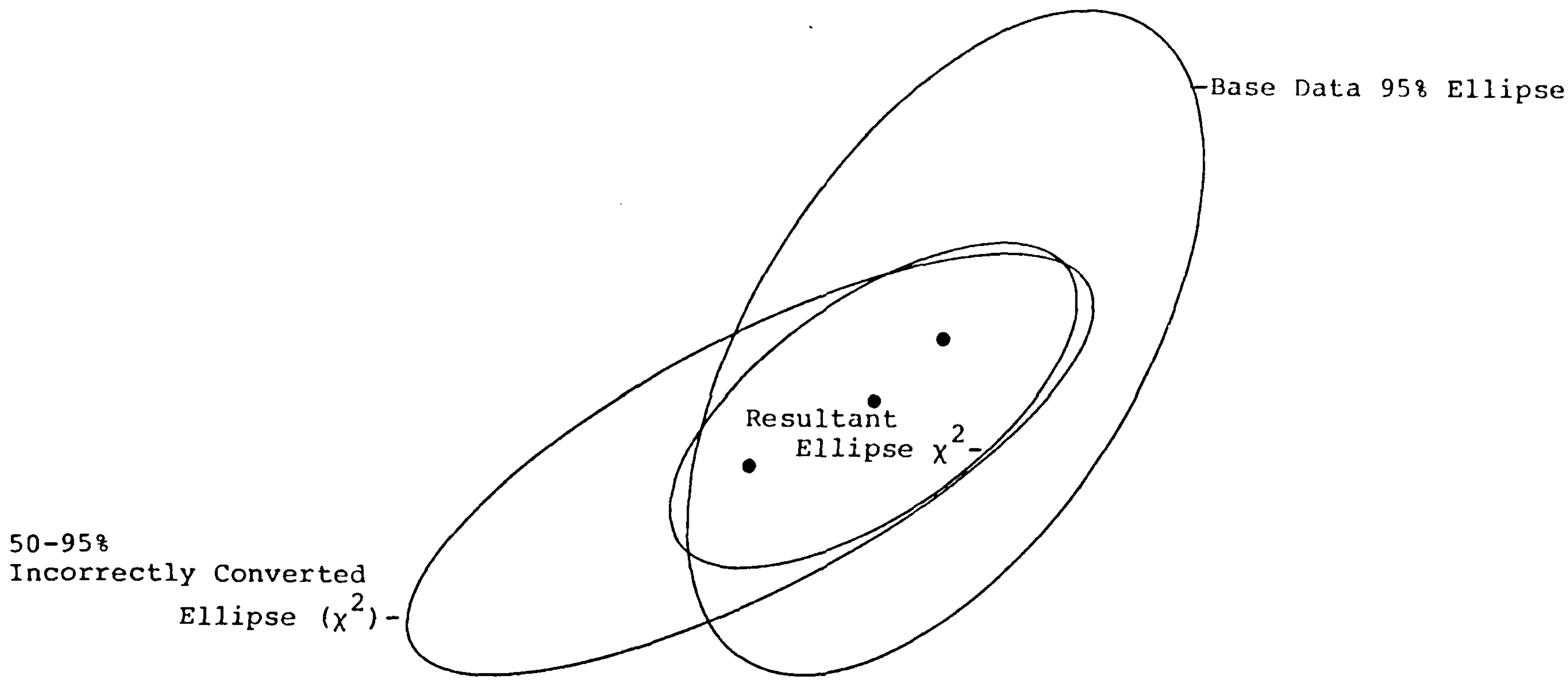
GRAPH 5



If we use the incorrect ellipse we reject that the new data are from the same emitter as the Base Data.

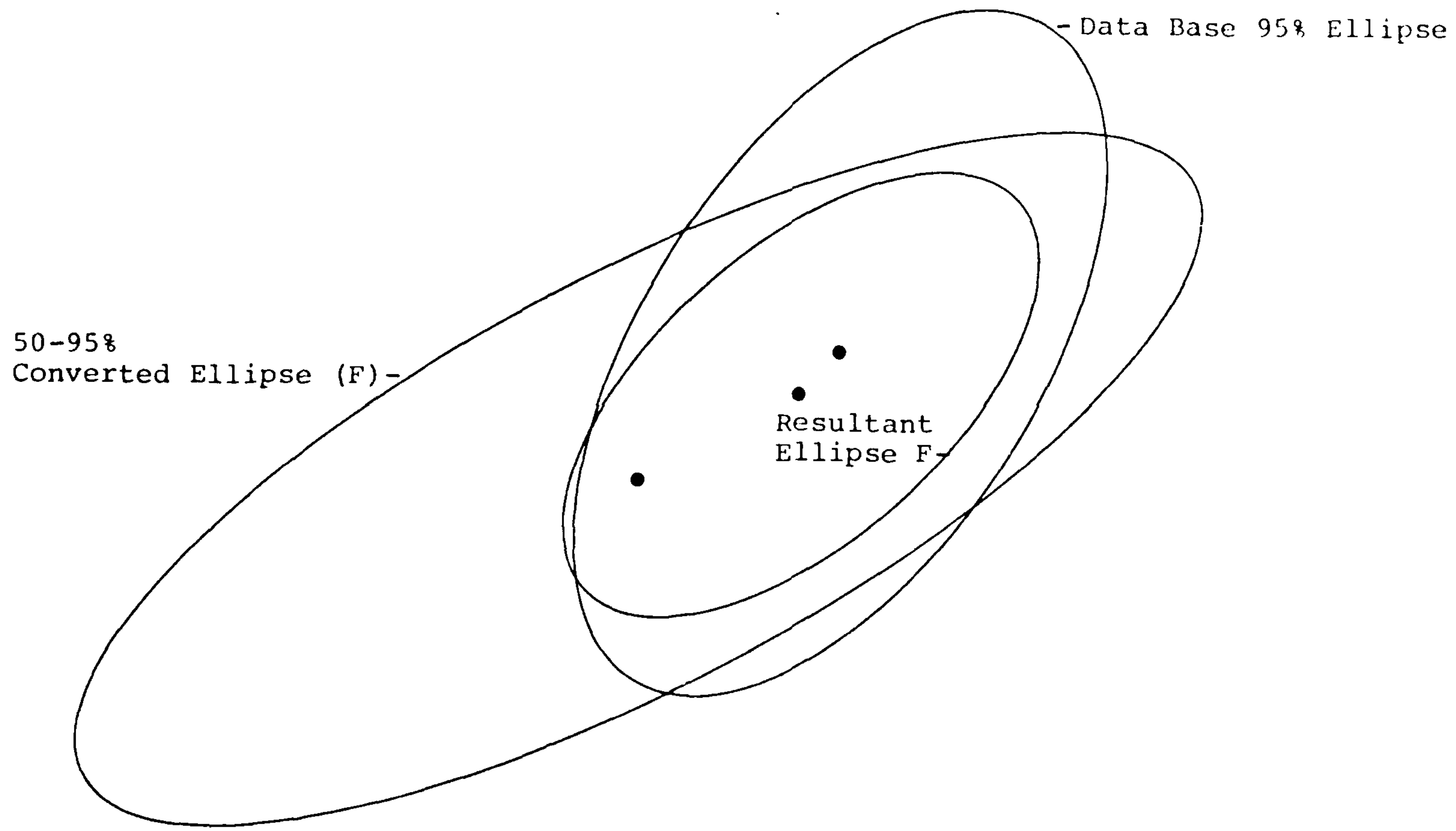
If we use the correct ellipse we do not reject that the new data are from the same emitter as the Base Data.

GRAPH 6

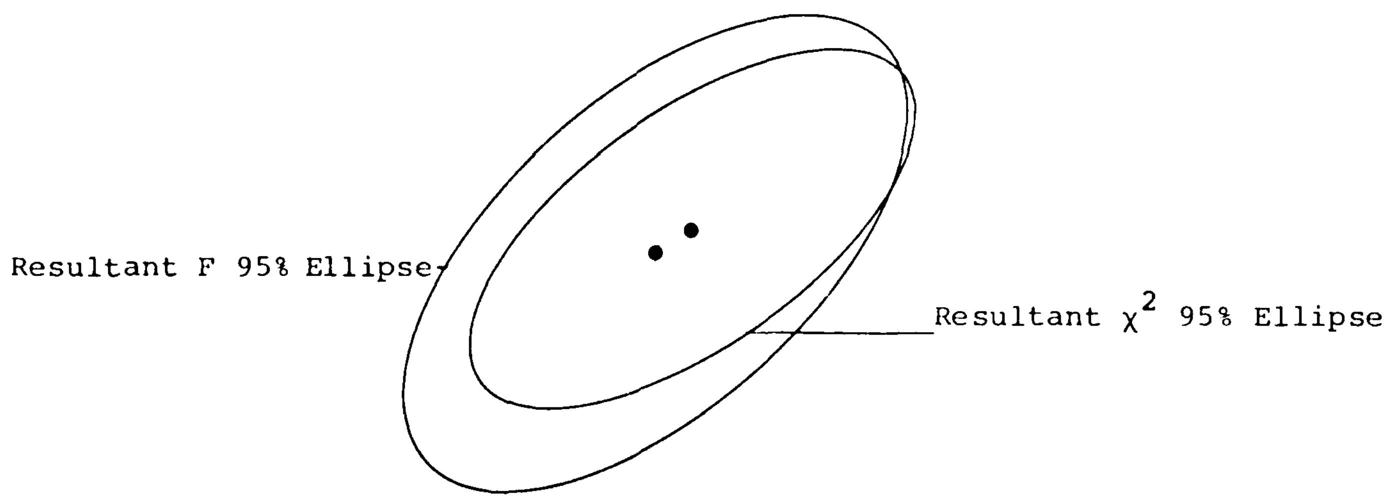


Graphs 6, 7, and 8 illustrate the results (resultant ellipse) when incorrectly and correctly converted ellipses are combined with Base Data.

GRAPH 7


Data Base 95% Ellipse
50-95%
Converted Ellipse (F)
Resultant
Ellipse F

GRAPH 8



Chi-square vs. F test CONFIDENCE LEVEL Conversion
2 DEGREES OF FREEDOM

CHI-SQUARE cut-off values does not depend on the number of data points.

50% cut-off value= 1.39
90% cut-off value= 4.61
95% cut-off value= 5.99
50% TO 95% cut-off ratio= 4.309   ELLIPSE RADIUS RATIO= 2.075
90% TO 95% cut-off ratio= 1.299   ELLIPSE RADIUS RATIO= 1.139

SAMPLE SIZE	50% CUTOFF VALUE	90% CUTOFF VALUE	95% CUTOFF VALUE	50% CUTOFF RATIO	90% CUTOFF RATIO	50% RADIUS RATIO	90% RADIUS RATIO	—	50 CHI-F RATIO RATIO	90 CHI-F RATIO RATIO
3	3	99	399	133	4.030	11.53	2.007	—	5.5556	1.7614
4	2	18	38	19	2.111	4.358	1.452	—	2.0998	1.2748
5	1.762	10.92	19.1	10.83	1.749	3.292	1.322	—	1.5860	1.1603
6	1.656	8.64	13.88	8.381	1.606	2.895	1.267	—	1.3946	1.1120
7	1.598	7.56	11.58	7.246	1.531	2.691	1.237	—	1.2968	1.0853
8	1.56	6.92	10.28	6.589	1.485	2.567	1.218	—	1.2366	1.0693
9	1.534	6.52	9.48	6.179	1.453	2.485	1.205	—	1.1975	1.0579
10	1.504	6.22	8.92	5.930	1.434	2.435	1.197	—	1.1731	1.0507
11	1.493	6.02	8.52	5.687	1.415	2.384	1.189	—	1.1488	1.0437
12	1.486	5.84	8.2	5.518	1.404	2.349	1.184	—	1.1316	1.0396
13	1.473	5.72	7.96	5.385	1.391	2.320	1.179	—	1.1179	1.0350
14	1.47	5.62	7.78	5.292	1.384	2.300	1.176	—	1.1082	1.0323
17	1.452	5.4	7.36	5.068	1.362	2.251	1.167	—	1.0845	1.0243
22	1.436	5.18	6.98	4.860	1.347	2.204	1.160	—	1.0620	1.0184

APPENDIX

Chi-square vs. F test CONFIDENCE LEVEL Conversion
3 DEGREES OF FREEDOM

CHI-SQUARE cut-off values does not depend on the number of data points.

50% cut-off value= 2.37
90% cut-off value= 6.25
95% cut-off value= 7.81
50% TO 95% cut-off ratio= 3.295 ELLIPSE RADIUS RATIO= 1.815
90% TO 95% cut-off ratio= 1.249 ELLIPSE RADIUS RATIO= 1.117

SAMPLE SIZE	50% CUTOFF VALUE	90% CUTOFF VALUE	95% CUTOFF VALUE	50% CUTOFF RATIO	90% CUTOFF RATIO	50% RADIUS RATIO	90% RADIUS RATIO	—	50 CHI-F RATIO RATIO	90 CHI-F RATIO RATIO
3	5.13	160.8	648	126.3	4.029	11.23	2.007	—	6.1915	1.7962
4	3.39	27.48	57.6	16.99	2.096	4.122	1.447	—	2.2708	1.2954
5	3	16.17	27.84	9.28	1.721	3.046	1.312	—	1.6782	1.1740
6	2.823	12.57	19.77	7.003	1.572	2.646	1.254	—	1.4578	1.1221
7	2.721	10.86	16.23	5.964	1.494	2.442	1.222	—	1.3454	1.0938
8	2.653	9.87	14.28	5.372	1.446	2.317	1.202	—	1.2769	1.0762
9	2.613	9.21	13.05	4.994	1.416	2.234	1.190	—	1.2311	1.0651
10	2.58	8.76	12.21	4.732	1.393	2.175	1.180	—	1.1984	1.0563
11	2.556	8.43	11.58	4.530	1.373	2.128	1.172	—	1.1725	1.0487
12	2.535	8.19	11.13	4.390	1.358	2.095	1.165	—	1.1543	1.0430
13	2.52	7.98	10.77	4.273	1.349	2.067	1.161	—	1.1388	1.0395
14	2.505	7.83	10.47	4.179	1.337	2.044	1.156	—	1.1262	1.0346
17	2.478	7.47	9.87	3.983	1.321	1.995	1.149	—	1.0994	1.0285
22	2.448	7.14	9.3	3.799	1.302	1.949	1.141	—	1.0737	1.0212

END
DATE
FILMED
5-88
DTIC