



RDECOM



AMSAA TECHNICAL REPORT NO. TR-2010-35

CONFIDENCE INTERVAL
METHODOLOGY FOR RATIO
MEANS (CIM4RM)

AUGUST 2010

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED

U.S. ARMY MATERIEL SYSTEMS ANALYSIS ACTIVITY
ABERDEEN PROVING GROUND, MARYLAND 21005-5071

20101014466

DESTRUCTION NOTICE

Destroy by any method that will prevent disclosure of contents or reconstruction of the document.

DISCLAIMER

The findings in this report are not to be construed as an official Department of the Army position unless so specified by other official documentation.

WARNING

Information and data contained in this document are based on the input available at the time of preparation.

TRADE NAMES

The use of trade names in this report does not constitute an official endorsement or approval of the use of such commercial hardware or software. The report may not be cited for purposes of advertisement.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503				
1. AGENCY USE ONLY (LEAVE BLANK)	2. REPORT DATE August 2010	3. REPORT TYPE AND DATES COVERED Technical Report		
4. TITLE AND SUBTITLE Confidence Interval Methodology for Ratio Means (CIM4RM)			5. FUNDING NUMBERS	
6. AUTHOR(S) John Nierwinski, Jr.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Director, U.S. Army Materiel Systems Analysis Activity 392 Hopkins Road Aberdeen Proving Ground, MD 21005-5071			8. PERFORMING ORGANIZATION REPORT NUMBER TR-2010-35	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED			12b. DISTRIBUTION CODE B	
13. ABSTRACT (Maximum 200 words) The U.S. Army and many other government and private organizations need to evaluate ratio means to help them make informed life cycle management decisions. A ratio mean is the ratio of the mean of two random variables, X and Y, whose corresponding terms are paired. An example of a ratio mean that the Army evaluates is the maintenance ratio (MR). The Army collects MR data from a sample of vehicles within a selected system and applies CIM4RM to the sample data to construct approximate confidence intervals (CI's) for the MR. CIM4RM is the combined effort of bootstrapping and creativity in constructing CI's around ratio means. CIM4RM was tested in many ratio mean applications and shown to be consistently valid. Prior to the development of CIM4RM, no documented tool existed that could produce consistently valid ratio mean CI's for various scenarios. The Army is currently using this methodology to evaluate ratio means for its fielded ground and aviation systems. The Office of Inspector General for Health and Human Services is utilizing this methodology for reporting ratio mean CI's to the U.S. Congress. Other existing applications include: performance evaluations for Army test systems, evaluations of an Improvised Explosive Device detection demonstration, and hypothesis test development for many applications that compare two ratio means. Although current applications are government centric, there are countless other areas in private industry (e.g. banking, automotive) where CIM4RM can be used for improving decision analysis.				
14. SUBJECT TERMS			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAME AS REPORT	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)

Prescribed by ANSI Std. Z39-18
298-102

THIS PAGE INTENTIONALLY LEFT BLANK.

CONTENTS

	Page
LIST OF TABLES	iv
ACKNOWLEDGEMENTS	v
LIST OF ACRONYMS	vi
LIST OF SYMBOLS	vii
1. INTRODUCTION	1
2. DEVELOPMENT OF METHODOLOGY (CIM4RM)	3
3. CONCLUSIONS	6
REFERENCES	41
DISTRIBUTION LIST	43

LIST OF TABLES

Table No.	Title	Page
1	Definition of Ratio Mean	1
2	Redefined Maintenance Ratio.....	3

ACKNOWLEDGEMENTS

The US Army Materiel Systems Analysis Activity (AMSAA) recognizes the following individuals for their contributions to this report:

The author is:

John Nierwinski, Logistics Analysis Division.

LIST OF ACRONYMS

CIM4RM	- Confidence Interval Methodology for Ratio Means
MH	- Man-Hours
MR	- Maintenance Ratio
PCM	- Parts Cost per Mile
CI	- Confidence Interval
BSS	- Bootstrap Sample
SE	- Standard Error
IID	- Independent Identical Distribution

LIST OF SYMBOLS

$miles_j$	Collected miles for vehicle or repair period
$man - hours_j$	Man-hours per vehicle or repair
$\hat{MR}_i$	Estimated maintenance ratio for i^{th} BSS
$\hat{adj} MR_j$	Weighted estimated maintenance ratio for a vehicle or repair
$\overline{AdjMR}$	Arithmetic mean of the population for adjMR
$\hat{Z}$	A standardized distribution
$\overline{man - hours}$	Arithmetic mean of the population for man-hours.
N	Population size
n	Sample size
f	Sampling fraction (n / N)
s_{miles}^2	Sample variance for miles
$s_{man-hours}^2$	Sample variance for man-hours
$covariance(man - hours, miles)$	Covariance for man-hours and miles
$\overline{MR}$	Arithmetic mean of the population for MR
$\hat{\overline{MR}}$	Estimate of $\overline{MR}$
$\hat{SE}$	Standard error estimate
α	Type I Risk
$(1 - 2\alpha)\%$	Confidence level
$\hat{\overline{MR}}$	Population MR estimate
$\hat{l}_\alpha$	α^{th} percentile for any Z distribution
$\hat{l}_{1-\alpha}$	$(1 - \alpha)^{th}$ percentile for any Z distribution
B	Number of BSS's required for CIM4RM

THIS PAGE INTENTIONALLY LEFT BLANK

CONFIDENCE INTERVAL METHODOLOGY FOR RATIO MEANS (CIM4RM)

1. INTRODUCTION

The U.S. Army and many other government and private organizations need to evaluate ratio means to help them make informed life cycle management decisions. A ratio mean is the ratio of the mean of two random variables, X and Y, whose corresponding terms are paired. The following chart depicts this:

Table 1. Definition of Ratio Mean.

X	Y	
x_1	y_1	
.	.	
.	.	
x_n	y_n	
$\sum_{j=1}^n x_j$	$\sum_{j=1}^n y_j$	$\frac{\sum_{j=1}^n x_j}{n} \bigg/ \frac{\sum_{j=1}^n y_j}{n} = \frac{\sum_{j=1}^n x_j}{\sum_{j=1}^n y_j}$

This is the sample ratio mean.

The pairs (x_j, y_j) are assumed to be independent. The correlation between X and Y may be positive, negative or zero. X & Y are random variables are considered independent and identically distributed (i.i.d.), with some distribution but unknown.

For example, the Army tracks and evaluates the performance of many weapon systems using ratio mean metrics, such as the maintenance ratio (MR). A MR estimate

$$\hat{MR} = \frac{\sum_{j=1}^n \text{man-hours}_j}{\sum_{j=1}^n \text{miles}_j}$$

is based on a random sample (without replacement) of n vehicles from a

finite population, where the pair (man-hours and miles) are associated with each vehicle. This example assumes there is no variation from visit to visit within each vehicle. A visit is defined to be any timeline event that requires repair.

The Army cannot always afford to track every vehicle in its inventory. Therefore, ratio mean performance metrics are tracked using a sample of vehicles over a given time period. The Army has multiple goals or objectives. One of them is to decide whether maintenance augmentation is necessary for a fleet of vehicles before a mission. Another goal might be to determine if a fleet of vehicles has achieved acquisition thresholds and targets. These goals and objectives are based on inferences from the sample using approximate confidence intervals (CI) for ratio means. This paper discusses and develops the methodology that produces these

confidence intervals (Confidence Interval Methodology for Ratio Means – CIM4RM). The MR will be used to develop the methodology.

CIM4RM is the combined effort of an existing tool (bootstrap-t approach [1] with no parametric assumptions on the distributions) and creativity to compute approximate confidence intervals for a ratio mean metric. The bootstrap-t approach is very applicable to a location statistic such as the ratio mean [2].

It is known that the bootstrap-t requires the estimation of three parameters based on a random sample from a population. They are: mean, standard error (SE), and bootstrap standardized Z distribution.

It is known that the SE estimate for a ratio mean is [3]:

$$\hat{SE} \approx \frac{\sqrt{\frac{(1-f)}{n * \overline{man-hours}} [s_{miles}^2 + \hat{MR}^2 s_{man-hours}^2 - 2 * \hat{MR} * \text{covariance}(man-hours, miles)]}}{\sqrt{n}} \quad (1)$$

Where n is the sample size and f is the sampling fraction (n/N).

s_{miles}^2 & $s_{man-hours}^2$ are the sample variances of miles and man-hours, respectively.

This SE estimate in equation (1) is only dependable for sample sizes greater than 30 where the coefficient of variation for both variables are less than 10% [3]. Therefore, another approach is needed to estimate these three parameters for smaller samples and larger samples with high variation. The approach that accomplishes this is CIM4RM.

2. DEVELOPMENT OF METHODOLOGY (CIM4RM)

Mean and SE estimates for ratio mean

Let's take a random sample (without replacement) of n vehicles (man-hours and miles for each vehicle) from the population of N vehicles. Let's redefine the sample ratio mean to be a sample arithmetic mean. The following chart depicts this redefinition.

Table 2. Redefined Maintenance Ratio.

<u>Vehicle #</u>	<u>Man-hours</u>	<u>Miles</u>	<u>Adj MR</u>
1	Man-hours ₁	Miles ₁	Adj MR ₁
.	.	.	
k	Man-hours _k	Miles _k	$adjMR_k = n * Man - hours_k / \sum_{j=1}^n Miles_j$
.	.	.	
.	.	.	
n	Man-hours _n	Miles _n	Adj MR _n
Sample Ratio Mean = $\frac{\sum_{j=1}^n Man - hours_j}{\sum_{j=1}^n Miles_j}$ = Sample Arithmetic Mean			

It can be shown that the sample ratio mean and the sample arithmetic mean are the same quantity. The redefined variable AdjMR only accounts for variation in man-hours and does not account for variation in miles. Correlation is also not accounted for because the pairing for man-hours and miles for each vehicle has been eliminated.

The following is an estimate of the arithmetic mean of the population ($\overline{AdjMR}$) based on the sample of n AdjMR's [1]:

$$\hat{MR} = \frac{\sum_{j=1}^n Man - hours_j / n}{\sum_{j=1}^n Miles_j / n} \quad (2)$$

The estimate of the SE of $\hat{MR}$ using this same sample of n AdjMR's is the following [1]:

$$\hat{SE} = \sqrt{\frac{\sum_{j=1}^n (AdjMR_j - \hat{MR})^2}{n^2}} \quad (3)$$

$\hat{MR}$ is also the estimate for $\overline{MR}$ [3], where $\overline{MR}$ is the mean of man-hours divided by the mean of miles based on the population. Let $\hat{SE}$ serve as the estimate for the SE for $\overline{MR}$. We know that this is not the best SE estimate for $\overline{MR}$ because it doesn't fully account for variation in miles. Nevertheless, using these estimates in the procedures described in the next section leads to excellent coverage, non-coverage and efficiency (distributions and sizes of CI lengths) results for many ratio mean problems.

Standardized Z estimate for ratio mean & the CI

It follows from the bootstrap-t that bootstrapping (based on a random sample) can be used to obtain an estimate for the standardized Z distribution[1]. In statistics, bootstrapping is a modern, computer-intensive general purpose approach to statistical inference, falling within a broader class of re-sampling methods. A bootstrap sample (BSS) of n vehicles is a random sample one vehicle at a time with replacement from the original sample of n vehicles. The probability of selecting a given vehicle for each of the n random selections is 1/n.

Bootstrapping has been in existence for over 25 years and has facilitated solutions of various kinds of problems (confidence intervals, variance reduction, hypothesis testing, etc.). This research uses bootstrapping to generate an approximate CI around a ratio mean. This is done by generating a large number of BSS's through a Monte Carlo [4] simulation procedure.

Now, let's estimate a bootstrap standardized $\hat{Z}$ distribution from the sample of n vehicles. Generate B BSS's of size n from the sample of n vehicles. Each of the n elements of BSS i (i=1 to B) will contain all vehicle information (i.e. man-hours, miles). For each BSS compute the mean and SE estimate the same way that was done for the original raw sample. Compute a standardized $\hat{Z}$ value for each BSS [1].

$$\hat{Z}_{(i)} = \frac{\hat{MR}_{(i)} - \hat{MR}}{\hat{SE}_{(i)}} \text{ where } \hat{MR}_{(i)} \text{ is an estimate of the mean for the } i^{th} \text{ BSS;} \quad (4)$$

$$\hat{SE}_{(i)} = \sqrt{\frac{\sum_{j=1}^n [AdjMR_{j(i)} - \hat{MR}_{(i)}]^2}{n^2}} \text{ is an estimate of the SE for the } i^{th} \text{ BSS}$$

$\hat{MR}$ is the original sample mean

Now we use ($\hat{SE}$, $\hat{MR}$, and $\hat{Z}$) estimates to compute an approximate 2-sided $100(1-2\alpha)\%$ confidence interval (α is the desired area in each tail) for the ratio mean metric. First compute the α^{th} percentile and $(1-\alpha)^{th}$ percentile of the $\hat{Z}$ distribution and call them $\hat{t}_\alpha$ and $\hat{t}_{1-\alpha}$, respectively. It follows that the $100(1-2\alpha)\%$ CI is [1]:

$$\begin{aligned} \text{Lower CI bound} &= \hat{MR} - \hat{SE} * \hat{t}_{1-\alpha} \quad \text{and} \\ \text{Upper CI bound} &= \hat{MR} + \hat{SE} * \hat{t}_\alpha \end{aligned} \quad (5)$$

The $\hat{Z}$ distribution accounts for the correlation of man-hours & miles and the variation in man-hours and miles. The following mathematical argument creates the framework for proceeding to the next level of validation.

Recall, $\hat{SE}$ was used to serve as the estimate for the SE for $\overline{MR}$. Also recall that we know that this is not the best SE estimate for $\overline{MR}$ because it doesn't fully account for variation in miles. The argument and hypothesis are that the product, $\hat{SE} * \hat{t}_{(\alpha \text{ or } 1-\alpha)}$ is not compromised because the lack of variation of miles in $\hat{SE}$ is accounted for in the $\hat{Z}$ distribution and ultimately $\hat{t}_{(\alpha \text{ or } 1-\alpha)}$. This hypothesis is tested and validated by simulating the confidence interval properties (coverage, non-coverage and efficiency) for many ratio mean problems. It was shown that the validation results for various ratio mean problems were excellent for many kinds of scenarios.

3. CONCLUSIONS

Many situations arise that require the need to develop reliable approximate confidence intervals for ratio means. CIM4RM is a tool that was developed to satisfy this need. Prior to CIM4RM, no documented tool existed that could produce stable validation results.

Coverage and its related properties for CIM4RM were tested with many ratio mean problems and were shown to perform very well for various measures of sample size, correlation, location and distribution mix. Therefore, the CIM4RM methodology is a reliable and stable tool for building CI's around ratio means.

The only scenarios where coverage starts to deviate away from the required level is for extremely high correlation with highly skewed data, existence of outliers in the data that cause the location parameter to shift or cases where the sample sizes are extremely small. When the two variables for the ratio mean are highly correlated, the confidence interval tends to be extremely short.

The Army is currently using this methodology to quantitatively analyze maintenance ratios and other ratio mean performance metrics for fielded Army ground and aviation systems. The Office of Inspector General for the Department of Health and Human Services is utilizing this methodology for reporting ratio mean confidence intervals to the U.S. Congress. Some other existing applications include: performance evaluations for Army test systems, evaluations of an Improvised Explosive Device Detection demonstration, hypothesis testing development for many applications that compare two ratio means, Aging Effects hypothesis testing, and paired reliability hypothesis testing.

Although current applications are government centric, there are countless other areas in private industry (e.g. banking, automotive) where CIM4RM can be used for improving decision analysis.

THIS PAGE INTENTIONALLY LEFT BLANK.

REFERENCES

1. Efron, B. and Tibshirani, R.J., "An Introduction to the Bootstrap", London U.K.: Chapman & Hall/CRC, 1998
2. DiCiccio, T. and Efron, B., "Bootstrap Confidence Intervals", Statistical Science, Vol. 11, No. 3, 1996
3. Cochran, W.G., "Sampling Techniques (Third edition)", John Wiley & Sons, New York, 1977
4. Ross, S.M., "Introduction to Probability Models", New York: Academic Press, 2003

APPENDIX A-DISTRIBUTION LIST

THIS PAGE INTENTIONALLY LEFT BLANK

DISTRIBUTION LIST

No. of Copies	Organization
8	Director US Army Materiel Systems Analysis Activity AMXAA-L AMXAA-B (1 copy) 392 Hopkins Road Aberdeen Proving Ground, MD 21009-5071
1	Deputy Commanding General for Systems of Systems Integration USArmy Research, Development and Engineering Command AMSRD-SS 6000 6th Street, Suite 100 Fort Belvoir, VA 22060-5608
2	US Army TRADOC Analysis Center ATRC-PR (Susan Matus) CD 255 Sedgwick Avenue Fort Leavenworth, KS 66027-2345
1	US Army TRAC-White Sands Missile Range ATRC-W (Dr. Lambert) Bldg 1400 Martin Luther King Drive White Sands Missile Range, NM 88002-5502
1	Defense Technical Information Center DTIC-O 8725 John J. Kingman Road, Suite 0944 Fort Belvoir, VA 22060-6218

THIS PAGE INTENTIONALLY LEFT BLANK