DEPARTMENT OF NATIONAL DEFENCE
CANADA

OPERATIONAL RESEARCH DIVISION

DIRECTORATE OF OPERATIONAL RESEARCH (JOINT)/
DIRECTORATE OF OPERATIONAL RESEARCH (CORPORATE)

ORD PROJECT REPORT PR 2003/15

A GUIDEBOOK FOR THE ANALYSIS OF CAMOUFLAGE DATA

By

Dr. Lise Arseneau
Edward J. Emond

SEPTEMBER 2003

OTTAWA, CANADA

National Defence Défense nationale

OPERATIONAL RESEARCH DIVISION

CATEGORIES OF PUBLICATION

ORD Reports are the most authoritative and most carefully considered publications of the DGOR scientific community. They normally embody the results of major research activities or are significant works of lasting value or provide a comprehensive view on major defence research initiatives. ORD Reports are approved personally by DGOR, and are subject to peer review.

ORD Project Reports record the analysis and results of studies conducted for specific sponsors. This Category is the main vehicle to report completed research to the sponsors and may also describe a significant milestone in ongoing work. They are approved by DGOR and are subject to peer review. They are released initially to sponsors and may, with sponsor approval, be released to other agencies having an interest in the material.

Directorate Research Notes are issued by directorates. They are intended to outline, develop or document proposals, ideas, analysis or models which do not warrant more formal publication. They may record development work done in support of sponsored projects which could be applied elsewhere in the future. As such they help serve as the corporate scientific memory of the directorates.

ORD Journal Reprints provide readily available copies of articles published with DGOR approval, by OR researchers in learned journals, open technical publications, proceedings, etc.

ORD Contractor Reports document research done under contract of DGOR agencies by industrial concerns, universities, consultants, other government departments or agencies, etc. The scientific content is the responsibility of the originator but has been reviewed by the scientific authority for the contract and approved for release by DGOR.

DEPARTMENT OF NATIONAL DEFENCE

CANADA

OPERATIONAL RESEARCH DIVISION

DIRECTORATE OF OPERATIONAL RESEARCH (JOINT)/
DIRECTORATE OF OPERATIONAL RESEARCH (CORPORATE)

ORD PROJECT REPORT PR 2003/15

A GUIDEBOOK FOR THE ANALYSIS OF CAMOUFLAGE DATA

by

Dr. Lise Arseneau
Edward J. Emond

Recommended by: ____________________________	Approved by: ________________________________
R.G. Dickinson DOR(Joint)	A.P. Bradfield DGOR

OTTAWA, ONTARIO SEPTEMBER 2003

ABSTRACT

This report was produced for NATO Task Group SCI-095 "Enhancement of Camouflage Assessment Techniques". It describes the analysis of data collected from different types of camouflage trials, including photosimulation experiments, live camouflage trials and video simulation. Since visual camouflage data are often statistically censored, the report offers procedures on how to adjust the raw observations, depending on the type of analysis to be carried out. Recommended methods for calculating descriptive statistics, including the median and first percentile of a dataset, are provided. Statistical tests for different experimental designs are presented with examples. The two main characteristics that determine which statistical test is most appropriate are the number of camouflage treatments being evaluated and whether the observations are independent or paired. For each case, the report recommends a statistical tool and gives an example of its use. Finally, the General Linear Model as an analysis tool for camouflage data is discussed and two examples are given.

RÉSUMÉ

Le présent rapport a été produit pour le groupe d'étude SCI-095 « Enhancement of Camouflage Assessment Techniques » de l'OTAN, et décrit l'analyse des données recueillies d'après différents essais de camouflage, dont des expériences de photosimulation, des essais en direct et des simulations vidéo. Les données statistiques sur le camouflage visuel étant souvent censurées, le rapport offre des procédures permettant d'ajuster les observations brutes selon le type d'analyse à effectuer. En outre, le document recommande des méthodes pour calculer des statistiques descriptives, y compris la médiane et le premier centile d'un ensemble de données, et présente des tests statistiques accompagnés d'exemples pour différents plans expérimentaux. Les deux principaux facteurs permettant de déterminer quel test est le plus approprié sont le nombre de traitements de camouflage évalués et le fait de savoir si les observations sont indépendantes ou jumelées. Pour chaque cas, le rapport suggère un outil statistique et fournit un exemple d'utilisation. Il traite finalement du modèle linéaire général en tant qu'outil d'analyse des données en matière de camouflage et fournit deux exemples à ce sujet.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

In February 2001, the NATO Task Group SCI-095 "Enhancement of Camouflage Assessment Techniques" was created. As indicated by its title, the purpose of SCI-095 is to provide advice and recommendations on ways to enhance the measurement and analysis of camouflage effectiveness. The members of SCI-095 elected to produce a guidebook explaining the methodologies developed over the years and currently used within the group for the analysis of camouflage data in the visible portion of the spectrum.

This report describes the types of camouflage trials, including photosimulation experiments, live camouflage trials and video simulation. The data collected in each of these trials pertains to detection events consisting of initial detection, recognition and identification. Much of the data collected in camouflage trials is censored in various ways. (See Section III, paragraph 14.) The report includes a discussion on the kinds of censoring encountered in camouflage trials, specifically interval censoring, left censoring and right censoring. Two methods for adjusting censored camouflage data are described. The uniform spacing method is applicable when the analysis is only concerned with the distributional properties of the data; for example, in the calculation of descriptive statistics. The single value spacing method is applicable when the analysis involves distributional properties of the data as well as covariate information, for example, in the analysis of the effect of variables such as observer age and experience on initial detection performance.

The analysis of camouflage data is discussed in detail. Visualization of the data using a cumulative frequency distribution curve is described. Recommended methods for calculating descriptive statistics, including the median and first percentile, are given.

The analysis of camouflage effectiveness is most often concerned with the evaluation of observer performance for differing camouflage treatments. Several statistical tests are used depending on the experimental design of the camouflage trial. The two main characteristics that determine which statistical test is appropriate are the number of camouflage treatments being evaluated and whether the observations are paired or independent. If each observer views more than one of the camouflage treatments being evaluated, then the data are paired. Alternatively, if each observer views only one of the camouflage treatments, then the data are independent. Table ES-I presents the appropriate statistical procedure for the four situations which commonly occur. The report describes each test in detail and gives an appropriate example.

TABLE ES-I

STATISTICAL TESTS FOR DIFFERENT EXPERIMENTAL DESIGNS

	Independent Observations	**Paired Observations**
2 Camouflage Treatments	Wilcoxon Two-Sample Test	Wilcoxon Two-Sample Paired Test
3 or more Treatments	Kruskal-Wallis Test	Friedman Test

Finally, the report discusses the use of the General Linear Model in the analysis of camouflage data. The General Linear Model investigates the relationship between a response variable such as initial detection range and independent variables such as observer age and previous experience. Two examples are given, one with independent observations and the other with paired observations.

A GUIDEBOOK FOR THE ANALYSIS OF CAMOUFLAGE DATA

I. INTRODUCTION

1. In February 2001, the NATO Task Group SCI-095 "Enhancement of Camouflage Assessment Techniques" was created. As indicated by its title, the purpose of SCI-095 is to provide advice and recommendations on ways to enhance the measurement and analysis of camouflage effectiveness. The Directorate of Soldier Systems Program Management and the Operational Research Division represent Canada on this Task Group. The members of SCI-095 elected to produce a guidebook explaining the methodologies developed over the years and currently used within the group for the analysis of camouflage data in the visible portion of the spectrum.

2. This report will describe several types of camouflage trials and experiments and the different kinds of data that these produce. The methods for converting raw observations from different trials into usable data for analysis will also be discussed. The calculation of descriptive statistics (median and first percentile) will be presented with examples. Statistical tests and procedures for the evaluation of camouflage effectiveness will also be described. It is intended that this document be included as part of the final report of NATO Task Group SCI-095.

AIM

3. The aim of this report is to act as a guidebook for analysts involved in the evaluation of camouflage effectiveness. Although the discussion focuses on the visual spectrum, some of the procedures may be applicable to the analysis of data from other wavelength regions.

II. CAMOUFLAGE TRIALS AND EXPERIMENTS

PHOTOSIMULATION TRIALS

4. A standard technique used by the NATO community to assess visual camouflage effectiveness is photosimulation. This procedure is comprised of two stages, the first of which consists of taking photographs of camouflaged vehicles in a field setting. Typically the targets are military vehicles such as tanks, armoured vehicles or trucks. Targets may be given different camouflage treatments, such as being covered with nets or painted with different patterns. Often an uncamouflaged vehicle is included to facilitate comparisons. The accepted standard for a photosimulation experiment is a set of ten to twenty 35 mm colour slides. The photographs are typically taken from a hovering helicopter starting at long range and decreasing in range and altitude at regular intervals.

5. The second stage of a photosimulation experiment is conducted under controlled conditions in a room set up for this purpose. Observers sit on a chair placed at a fixed location in front of a standard projection screen. Each observer participating in the trial views each 35 mm slide for a fixed time starting from the longest range slide. The range of the slide at which the observer first detects each target is noted. In like manner, recognition and identification data can also be collected. Further details on photosimulation procedures are included in References 1 and 2.

6. A recent innovation is to use a laptop computer and digital imagery instead of a projection screen and photosimulation slides. Again each observer views the imagery from long range to short range for a fixed time and the image at which each detection event first occurs is recorded.

LIVE CAMOUFLAGE TRIALS

7. Live camouflage trials are conducted from time to time by NATO nations. In a live camouflage trial, each observer is brought to the field and given the opportunity to detect targets. There are several possible variants of live camouflage trials. Observers may for example be driven to predetermined locations and allowed to search for targets. Another option exercised during the NATO SCI-095 Camouflage Assessment Trial involved observers searching for targets while descending in a cable car. See Reference 3 for more details of this trial. In live camouflage trials, the distance from the observer to each target when detection, recognition and identification occur may be recorded.

VIDEO SIMULATION

8. In a video simulation, continuous imagery or imagery clips of several seconds each is recorded in the field with one or more targets starting at long range and decreasing to short range. Observers are then shown the video in controlled conditions and the ranges at which detection events occur are recorded.

III. CAMOUFLAGE DATA

DETECTION EVENTS

9. In camouflage assessment there are generally three categories of range information that can be collected: initial detection ranges, recognition ranges and identification ranges. The three corresponding detection event types are defined below.

10. Initial Detection: Initial detection occurs when the observer first correctly indicates the presence of a target of potential military interest. In camouflage trials, this indication by the observer must be specific and observable. In photosimulation trials a laser pointer may be used to point at the target. In live trials it may be necessary to improvise in order to ensure that there is a discernable event to define initial detection. For example, in the SCI-095 Camouflage Assessment Trial held in October 2002, observers used a picture of the background scene (without any targets) to point to the location of targets of interest.

11. Recognition: Recognition occurs when the observer differentiates between various categories of targets or objects. For example, the observer may state that the target of potential military interest is an armored fighting vehicle as opposed to a cargo vehicle.

12. Identification: Identification occurs when the identity of a target within a given category is indicated. For example, the observer may state that the target is a 5-ton cargo truck or is an Armored Personnel Carrier (APC).

13. Detection Event: In order to simplify the discussion in this report, the generic term “detection event” will sometimes be used rather than initial detection, recognition and identification. The reason for this is that the methodology for analyzing the range observations is the same for the three types of data.

DATA CENSORING

14. Much of the data collected in camouflage trials is censored in various ways. A definition of data censoring is given below.

> “A data sample is said to be censored when, either by accident or design, the value of the random variable under investigation is unobserved for some of the items in the sample. A *censored* observation is distinct from a *missing* observation in that the order of the censored observation relative to some of the uncensored observations is known and conveys information regarding the distribution being sampled.” [1]

15. In photosimulation experiments where observers are shown a series of slides taken at decreasing ranges from the target of interest, detection events are <u>interval censored</u>. For example, suppose a detection event takes place at the 1000 m slide and that the previous slide was taken at 1200 m. Although the detection event took place at 1000 m, there was a chance for it to happen between 1200 m and 1000 m and it may well have happened for example at 1100 m if there had been a slide at this range. Thus the detection event is interval censored, meaning that our best information is that it would have occurred within the interval 1000 m to 1200 m.

16. Camouflage detection event data may be <u>right censored</u>. Right censoring occurs when a detection event takes place at the longest possible range. For example, in a photosimulation experiment, if a detection event happens on the first (longest range) slide, the range for this observation is right censored, meaning that we know only that the detection event would have occurred at a range greater than or equal to the first slide range. This can also occur in live trials and video simulation if a detection event takes place at the longest possible range.

17. Camouflage detection event data may also be <u>left censored</u>. Left censoring occurs when there is a lower bound to a detection event and the detection event does not happen. For example if the final slide in a photosimulation experiment is at 400 m and a detection event does not occur at or before this slide, our best information is that the event would have taken place somewhere between 0 and 400 m. We say that the detection event is left censored at 400 m. This can also occur in live trials and video simulation if a detection event does not take place by the end of the observation period.

18. For analysis purposes, whenever the observations from a camouflage trial are censored in some way, it is necessary to adjust the raw data. The use of censored data without adjustment can lead to statistical bias and incorrect conclusions in the

---

[1] Kotz and Johnson, Encyclopedia of Statistical Sciences

analysis of the results. The adjustment of censored camouflage data depends on both the type of censoring and the type of analysis.

METHODS FOR ADJUSTING CENSORED CAMOUFLAGE DATA

19. Two methods for adjusting censored camouflage data will be discussed. The first method is uniform spacing and the second method is single value spacing. Note that for mathematical purposes, all ranges are treated as negative values in the calculations.

Uniform Spacing

20. This method applies to interval and left censored camouflage data. Suppose there are n detection event observations that are interval censored between ranges R_a and R_b. For photosimulation data, R_a is the range of the slide where the n observations occurred and R_b is the range of the immediately preceding (longer range) slide. For left censored data, R_a is 0 and R_b is the minimum range at which a detection event is possible. In uniform spacing the n observations are given range values that are evenly distributed between R_a and R_b with $(R_a - R_b) / (n+1)$ as the spacing. This method assumes that the camouflage data derive from a uniform distribution. In most cases,Lacking any distributional information, this assumption is the most reasonable. However, if the analyst has information that challenges this assumption, it may be more accurate to adjust the values from the raw data observations accordingly.

21. Adjustment of data using uniform spacing is applicable when the analysis is only concerned with the distributional properties of the data, for example, in the calculation of descriptive statistics. This method may not be appropriate when the analysis also involves covariate information, such as the age or experience of the observers.

22. <u>Example of Data Adjustment Using Uniform Spacing</u>: The data in Table I were obtained in a photosimulation trial conducted at CFB Petawawa, Canada, in December 2001. The target in the slide series was an uncamouflaged light armoured vehicle. The first column in Table I gives the slide range in metres and the second column gives the number of observers that first detected the target at a given slide range. The number of observers was 27.

TABLE I

PHOTOSIMULATION INITIAL DETECTION RANGES

23. Table II shows the adjusted data values from Table I using uniform spacing. The first column of Table II gives the slide range of each observation and the second column gives the adjusted value using uniform spacing.

TABLE II
ADJUSTED INITIAL DETECTION RANGES BY UNIFORM SPACING

Initial Detection Slide Ranges (m)	Adjusted Initial Detection Ranges (m)
1600	-1700
1400	-1550
1400	-1500
1400	-1450
1000	-1075
1000	-1050
1000	-1025
900	-990
900	-980
900	-970
900	-960
900	-950
900	-940
900	-930
900	-920
900	-910
800	-880
800	-860
800	-840
800	-820
700	-767
700	-733
600	-650
500	-567
500	-533
400	-450
300	-350

Single Value Spacing

24. This method applies to interval and left censored camouflage data. Suppose there are n detection event observations that are interval censored between ranges R_a and R_b. For photosimulation data, R_a is the range of the slide where the n observations occurred and R_b is the range of the immediately preceding (longer range) slide. For left censored data, R_a is 0 and R_b is the minimum range at which a detection event is possible. In single value spacing the n observations are given a single range value that is halfway between R_a and R_b with $(R_a – R_b) / 2$ as the spacing. This single range value is chosen since the underlying distribution is not known and it is preferable with

camouflage data to use a value that is conservative (i.e., a value that is greater than R_a).

25. Adjustment of data using single value spacing is applicable when the analysis involves distributional properties of the data as well as covariate information. For example, in the analysis of the effect of variables such as observer age and experience on initial detection performance.

26. <u>Example of Data Adjustment Using Single Value Spacing</u>: Table III shows the adjusted data values from Table I using single value spacing. The first column of Table III gives the slide range of each observation and the second column gives the adjusted value using single value spacing.

TABLE III
ADJUSTED INITIAL DETECTION RANGES BY SINGLE VALUE SPACING

Initial Detection Slide Ranges (m)	Adjusted Initial Detection Ranges
1600	-1700
1400	-1500
1400	-1500
1400	-1500
1000	-1050
1000	-1050
1000	-1050
900	-950
900	-950
900	-950
900	-950
900	-950
900	-950
900	-950
900	-950
900	-950
800	-850
800	-850
800	-850
800	-850
700	-750
700	-750
600	-650
500	-550
500	-550
400	-450

300	-350

Right Censored Data

27. The uniform spacing and single value spacing methods apply to both interval and left censored camouflage data. However, when a detection event is right censored and occurs at the longest possible slide range, it is not possible to adjust the raw observations that occur at this range. The reason is that there is no slide immediately preceding the longest range slide to give a bound on the range where no detections took place. It may be possible to use the value of the longest slide range for these raw observations in the analysis, depending on the situation and the statistical tool being applied. For example, the median can still be calculated as long as more than half the raw observations are not right censored. In this case, the longest slide range will be assigned as the value corresponding to the observations that are right censored (i.e., the ‘adjusted value’ is the same as the raw observation). Note that in the design of a photosimulation experiment, the longest range slide should be at a distance where no detection events will occur to prevent camouflage data being right censored.

IV. ANALYSIS OF CAMOUFLAGE DATA

28. An important principle in the analysis of camouflage data is that the type of analysis to be carried out determines the way the data must be handled. This principle will be explained and illustrated with practical examples. There are two different types of analysis that will be discussed in this report. In the first type of analysis, the data values are treated as sample points from a distribution and the analysis is concerned with statistical estimates of distribution parameters. In this case the order of the data values is not important and there is no other information associated with any given data point. In the second type of analysis, the distinguishing feature is that each data point has one or more covariates associated with it. For example, if the analysis is concerned with the possible effect on detection event performance of observer age, then age is a covariate and each detection event observation has an observer age associated with it. The time of each observation and even the order of observations are examples of possible covariates of interest.

29. Although detection, recognition and identification ranges are always reported as positive values, it is mathematically convenient in the analysis to treat ranges as negative values. This convention applies for both adjusted and non-adjusted data values.

CUMULATIVE FREQUENCY DISTRIBUTION

30. It is often useful and informative to create a cumulative frequency distribution curve for camouflage data because it illustrates the distribution of detection event observations for the entire sample at once. This curve shows the percentage of observations that occurred at or before each range. To create a cumulative frequency distribution curve, it is appropriate to use the uniform spacing method to adjust the raw observations. The data from Table II were used to create the curve shown in Figure 1.

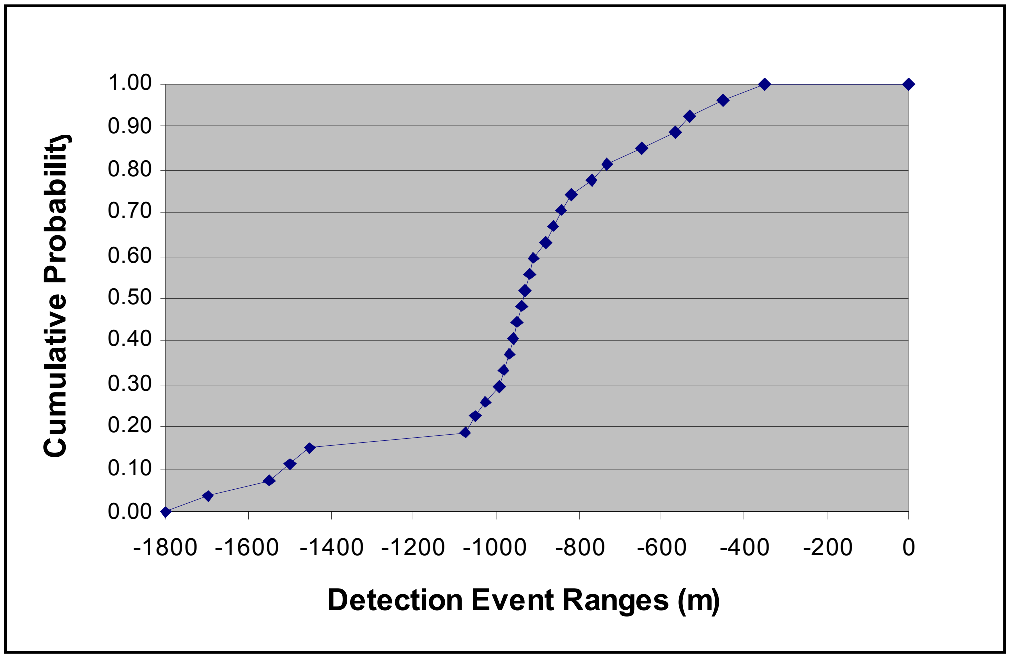

Figure 1: Cumulative Initial Detection Probability Curve for the Data in Table II

DESCRIPTIVE STATISTICS

Median

31. A standard way to summarize a set of data using numerical methods is to calculate a central value for the data such as the mean or median. The median is preferred over the mean for camouflage data because extreme data points, such as unusually high or low initial detection range, could skew the mean and make the central value appear higher or lower. The median is unaffected by outliers.

32. To calculate the median, first adjust the raw data observations, if necessary, using the uniform spacing method. The median is the middle value of an ordered set of observations (i.e., the ordered set of the adjusted values). When the number of observations is odd, the median is the middle value. When the number of observations is even, the median is the average of the two middle values.

33. For example, the median of the 27 ordered data values in Table II is the fourteenth value, 930 m.

First Percentile

34. In some cases one may be less interested in understanding average performance and more concerned about exceptional performance. There are several possible statistics that can be used to describe the longest range at which initial detection might take place.

35. It is recommended that the estimated first percentile of the distribution of observations be used as a standard descriptive statistic. The first percentile is the distance at which an observer has a one percent probability of initially detecting the target. Thus it is reasonable to use the first percentile as the threshold for initial detection of a target.

36. In order to calculate the first percentile, first adjust the N raw data observations, if necessary, using the uniform spacing method.[2] Then sort the data in decreasing order of magnitude and denote the data value with the largest magnitude as $X_{(1)}$. Next calculate the median as explained above and denote it by $X_{.50}$. The estimated first percentile, denoted by $X_{.01}$ is given by the following formula.

$$X_{.01} \quad = \quad X_{.50} - 2.326\left(\frac{X_{(1)} - X_{.50}}{\alpha(N)}\right)$$

37. The value α(N) may be obtained from Annex A. A detailed derivation of the first percentile formula is given in Annex B. To illustrate the calculation of the first percentile consider the data in Table II. In this case, $N = 27$, $X_{.50} = -930$, $X_{(1)} = -1700$ and $\alpha(27) = -1.998$. Therefore the estimated first percentile is 1826 m.

$$X_{.01} \quad = \quad -930 - 2.326\left(\frac{-1700-(-930)}{-1.998}\right) \quad = \quad -1826$$

TESTING FOR SIGNIFICANT DIFFERENCES

38. The analysis of camouflage effectiveness is often concerned with the evaluation of observer performance for differing camouflage treatments. Several statistical tests are used depending on the experimental design of the camouflage trial. The two main

---

[2] Note that if the data are right censored, it is not possible to estimate the first percentile.

characteristics that determine which statistical test is appropriate are the number of camouflage treatments being evaluated and whether the observations are paired or independent. If each observer views more than one of the camouflage treatments being evaluated, then the data are paired. Conversely, if each observer views only one of the camouflage treatments, then the data are independent. Table IV gives the appropriate statistical procedure for the four situations which commonly occur. Each of the tests will be discussed below.

TABLE IV
STATISTICAL TESTS FOR DIFFERENT EXPERIMENTAL DESIGNS

	Independent Observations	**Paired Observations**
2 Camouflage Treatments	Wilcoxon Two-Sample Test	Wilcoxon Two-Sample Paired Test
3 or more Treatments	Kruskal-Wallis Test	Friedman Test

Wilcoxon Two-Sample Test

39. The Wilcoxon two-sample test provides a methodology to determine if there is a statistically significant difference between the medians of two sets of data, irrespective of their distribution. The Wilcoxon test is most often used to assess differences between detection performance against two differently camouflaged targets for two groups of observers.[3] Note that the Wilcoxon two-sample test requires that the two samples of observers be independent. This means that each observer views only one of the targets under investigation. If the same observer views both targets, the observations are said to be paired and the Wilcoxon test is not applicable because the samples are not independent. The analysis of paired observations will be discussed below.

40. The Wilcoxon test is a nonparametric test. Nonparametric implies that there is no assumption of a specific distribution for the population. As a result, nonparametric tests are statistically robust. However, the Wilcoxon test may be less sensitive to small differences between medians of two data sets than statistical tests that make a specific assumption about the underlying distribution of the population.

41. The output from a Wilcoxon test is a "p-value" indicating the strength of evidence that there is a significant difference between the median detection event ranges for two camouflage treatments. For camouflage effectiveness analysis, a p-value less than or equal to 0.05 is said to indicate a statistically significant difference. A p-value between 0.05 and 0.10 is said to indicate some evidence of a difference.

42. In order to apply the Wilcoxon two-sample test, first adjust the values of the raw observations, if necessary, using the uniform spacing method. Then, a statistical software package such as the Statistical Analysis System (SAS) can be used to apply the test on the adjusted data values. See Reference 1 for more details.

---

[3] This test can also be used to evaluate the detection performance of two independent groups of observers for the same target on two different occasions.

43. Example of the Application of the Wilcoxon Two-Sample Test: An example of the use of the Wilcoxon two-sample test is the comparison of observer performance in a live trial with a photosimulation experiment. The data in Table V were collected during the Camouflage Assessment Trial in October 2002 at Oberjettenberg, Germany. Observers participating in this live trial descended towards the target area in a cable car. See Reference 3 for more details. The first column of Table V presents the raw observations for initial detection range and the second column gives the adjusted values by the uniform spacing method. Since the observations were collected during a live trial, there was no interval censoring and only one value required adjustment because it was left censored. One observer did not detect the target before the line of sight to the target area was lost due to the intervening tree line. Therefore this observation was left censored at 428 m. Accordingly, an adjusted value of 214 m was assigned to the raw observation of 0 m.

TABLE V
DATA FROM CAMOUFLAGE ASSESSMENT TRIAL

Initial Detection Range (m)	Adjusted Values by Uniform Spacing (m)
0	-214
464	-464
941	-941
1012	-1012
1115	-1115
1161	-1161
1285	-1285
1299	-1299
1809	-1809
1937	-1937
2003	-2003
2089	-2089

44. The data in Table VI were collected during a photosimulation experiment in January 2003 at CFB Gagetown, Canada. The imagery for the photosimulation was collected during the Camouflage Assessment Trial in October 2002 mentioned above. The purpose of the photosimulation experiment was to compare the results obtained with those from the live trial. The first column in Table VI gives the range associated with the slide where initial detection occurred. The second column gives the adjusted values using the uniform spacing method.

TABLE VI
DATA FROM GAGETOWN PHOTOSIMULATION TRIAL

Initial Detection Ranges (m)	
Raw Data	**Adjusted Values Using Uniform Spacing**
0	-93
0	-187
0	-280
0	-374
467	-496
524	-538
524	-551
524	-565
578	-593
578	-607
578	-622
696	-728
947	-979
1071	-1092
1071	-1113

45. The median initial detection range for the live trial was 1223 m and for the photosimulation trial was 565 m. The Wilcoxon two-sample test was applied to test the hypothesis that there was no statistically significant difference between the medians for the datasets given in Tables V and VI. This test is appropriate because the observations in both datasets are independent. The Wilcoxon test was applied using SAS. The p-value was 0.0032 indicating that there is a statistically significant difference between the medians of the initial detection ranges for the two datasets.

Wilcoxon Two-Sample Paired Test

46. The Wilcoxon two-sample paired test is used to assess differences between detection performance against two differently camouflaged targets when the observations are paired.[4] This means that each observer views both of the targets under investigation. Since the same observer views both targets, the observations are said to be paired and the Wilcoxon two-sample test is not applicable because the samples are not independent.

---

[4] This test can also be used to evaluate the detection performance of paired observations for the same target on two different occasions.

47. In order to apply the Wilcoxon two-sample paired test, first adjust the values of the raw observations, if necessary, using the single value spacing method. Then, a statistical software package such as SAS can be used to apply the test on the adjusted data values. See Reference 4 for more details.

48. Example of the Wilcoxon Two-Sample Paired Test: The data in Table VII were collected during a photosimulation experiment in January 2003 at CFB Gagetown, Canada. Sixteen observers were shown a series of 20 slides starting at range 2000 m and steadily decreasing by 100 m. The targets of interest in this example were two different camouflage nets deployed over military vehicles. The first column gives the observer sequence number.[5] Columns 2 and 4 of Table VII give the slide range at which initial detections occurred for each observer. Columns 3 and 5 give the adjusted data values using the single value spacing method. This is appropriate since the observations were paired interval censored data from a photosimulation trial.

TABLE VII
DATA FOR WILCOXON TWO-SAMPLE PAIRED TEST EXAMPLE

Sequence Number	Initial Detection Ranges (m)			
	Camouflage Net 1		Camouflage Net 2	
	Raw Data	Adjusted Values by Single Value Spacing	Raw Data	Adjusted Values by Single Value Spacing
3	300	-350	300	-350
11	400	-450	400	-450
13	200	-250	200	-250
24	600	-650	600	-650
27	100	-150	200	-250
31	0	-50	200	-250
37	200	-250	200	-250
43	300	-350	300	-350
49	400	-450	400	-450
58	400	-450	500	-550
65	400	-450	400	-450
72	300	-350	700	-750
74	300	-350	500	-550
84	300	-350	300	-350
94	400	-450	400	-450
98	500	-550	500	-550

---

[5] The sequence numbers are not consecutive because there were 99 observers who participated in this photosimulation experiment but only those that viewed the slide set being discussed are noted in Table VII.

49. The median initial detection range for Camouflage Net 1 is 375 m and for Camouflage Net 2 is 430 m. The Wilcoxon two-sample paired test was applied to test the hypothesis that there was no statistically significant difference between the medians for the datasets given in Table VII. This test is appropriate because the observations in both datasets are paired. The Wilcoxon two-sample paired test was applied using SAS. The p-value was 0.0625 indicating that there is some evidence of a difference between the medians of the initial detection ranges for the two camouflage nets. Since the p-value is greater than 0.05, the evidence is not sufficient to declare a statistically significant difference.

Kruskal-Wallis Test

50. The Kruskal-Wallis Test is a nonparametric test that is used to determine if there is a statistically significant difference between three or more independent groups of sampled data, regardless of their distribution. When there are only two groups in the comparison, the test becomes a Wilcoxon two-sample test. To apply the Kruskal-Wallis Test, first adjust the values of the raw observations, if necessary, using the uniform spacing method. Then, a statistical package such as SAS can be used. If the test returns a p-value that is less than or equal to 0.05, then two or more groups of the sampled data have medians which are significantly different. In this case, further analysis is required in order to determine how to partition the datasets into groups with equal medians.

51. Example of the Kruskal-Wallis Test: The data in Table VIII were collected during a photosimulation experiment in January 2003 at CFB Gagetown, Canada. The imagery for the photosimulation was collected during the Camouflage Assessment Trial in October 2002. The purpose of the photosimulation experiment was to compare the results obtained with those from the live trial. Three groups of observers were shown three separate series of slides taken over the three days of the Camouflage Assessment Trial. The target of interest in this example was a small wheeled vehicle, partially camouflaged. The first, third and fifth columns of Table VIII give the slide range at which initial detections occurred for the three different slide series, labeled NATO Series A, B and C, respectively. Note that the number of observers for NATO Series A and B was sixteen and for NATO Series C was fifteen. Column two, four and six give the adjusted values using the uniform spacing method. This is appropriate since the observations were independent interval censored data from a photosimulation trial.

TABLE VIII
DATA FOR KRUSKAL-WALLIS TEST EXAMPLE

Initial Detection Range (m)					
NATO Series A		**NATO Series B**		**NATO Series C**	
Raw Data	**Adjusted Values Using Uniform Spacing**	**Raw Data**	**Adjusted Values Using Uniform Spacing**	**Raw Data**	**Adjusted Values Using Uniform Spacing**
0	-30	0	-40	0	-93
0	-60	0	-81	0	-187
0	-90	0	-121	0	-280
0	-121	0	-162	0	-374
0	-151	0	-202	467	-496
0	-181	0	-243	524	-538
0	-211	0	-283	524	-551
0	-241	0	-324	524	-565
0	-271	0	-364	578	-593
0	-301	0	-405	578	-607
0	-331	445	-459	578	-622
0	-362	445	-473	696	-728
0	-392	445	-486	947	-979
0	-422	500	-518	1071	-1092
452	-480	500	-535	1071	-1113
508	-535	670	-701		

52. The median initial detection ranges for NATO Series A, B and C are 256 m, 344 m and 565 m, respectively. The Kruskal-Wallis test was applied to test the hypothesis that there was no statistically significant difference between the medians for the datasets given in Table VIII. This test is appropriate because the observations in the datasets are independent. The Kruskal-Wallis test was applied using SAS. The p-value was 0.0013 indicating that at least one of the median initial detection ranges for the three slide series is significantly different from the others. Further analysis[6] indicated that the median initial detection range for NATO Series C was significantly greater than the medians for NATO Series A and B. In addition, the analysis indicated that the medians for NATO Series A and B were not significantly different.

Friedman Test

53. The Friedman Test is a nonparametric test that is used to determine if there is a statistically significant difference between three or more groups of paired data,

---

[6] Pairwise comparisons using the Wilcoxon two-sample test with an adjusted p-value were used in this case.

regardless of their distribution. When there are only two groups in the comparison, the Wilcoxon two-sample paired test should be used. To apply the Friedman Test, first adjust the values of the raw observations, if necessary, using the single value spacing method. Then, a statistical package such as SAS can be used. If the test returns a p-value that is less than or equal to 0.05, then two or more groups of the sampled data have medians which are significantly different. In this case, further analysis is required in order to determine how to partition the datasets into groups with equal medians.

54. Example of the Friedman Test: The data in Table IX were collected during a photosimulation experiment in January 2003 at CFB Gagetown, Canada. Sixteen observers were shown a series of 20 slides starting at range 2000 m and steadily decreasing by 100 m. The targets of interest in this example were three camouflage nets deployed over military vehicles. The first column of Table IX gives the observer sequence number. Columns 2, 4 and 6 give the slide range at which initial detections occurred for each observer. Columns 3, 5 and 7 give the adjusted data values using the single value spacing method. This is appropriate since the observations were paired interval censored data from a photosimulation trial.

TABLE IX
DATA FOR FRIEDMAN TEST EXAMPLE

Sequence Number	Initial Detection (m)					
	Camouflage Net 1		Camouflage Net 2		Camouflage Net 3	
	Raw Data	Adjusted Values by Single Value Spacing	Raw Data	Adjusted Values by Single Value Spacing	Raw Data	Adjusted Values by Single Value Spacing
3	300	-350	300	-350	300	-350
11	400	-450	400	-450	400	-450
13	200	-250	200	-250	200	-250
24	600	-650	600	-650	600	-650
27	100	-150	100	-150	200	-250
31	0	-50	0	-50	200	-250
37	200	-250	200	-250	200	-250
43	300	-350	300	-350	300	-350
49	400	-450	400	-450	400	-450
58	600	-650	400	-450	500	-550
65	400	-450	400	-450	400	-450
72	300	-350	300	-350	700	-750
74	100	-150	300	-350	500	-550
84	300	-350	300	-350	300	-350
94	400	-450	400	-450	400	-450
98	1800	-1850	500	-550	500	-550

55. The median initial detection range for Camouflage Net 1 is 350 m, for Camouflage Net 2 is 375 m and for Camouflage Net 3 is 430 m. The Friedman test was applied to test the hypothesis that there was no statistically significant difference between the medians for the datasets given in Table IX. This test is appropriate because the observations in the datasets are paired. The Friedman test was applied using SAS. The p-value was 0.2982 indicating that there is no statistically significant difference between the medians of the initial detection ranges for the three camouflage nets.

ANALYSIS USING THE GENERAL LINEAR MODEL

56. Both multiple linear regression analysis and the analysis of variance with several factors fall under the heading of the General Linear Model. The General Linear Model is used to investigate the relationship between a response variable such as

initial detection range and independent variables such as observer age and previous experience. The General Linear Model requires more assumptions than the Wilcoxon test but permits the analysis of multiple variables simultaneously. Since the General Linear Model involves associating variables with data values, the first step is to adjust the values of the raw observations, if necessary, using the single value spacing method. Then a statistical software package such as SAS can be used to apply the General Linear Model using the adjusted values as input. Two examples will be given, one with independent observations and the other with paired observations.

57. Example of the General Linear Model with Independent Observations: The data in Tables X, XI and XII were collected during a photosimulation experiment in January 2003 at CFB Gagetown, Canada. The imagery for the photosimulation was collected during the Camouflage Assessment Trial in October 2002. Three groups of observers were shown three separate series of slides taken over the three days of the Camouflage Assessment Trial. The target of interest in this example was a small wheeled vehicle, partially camouflaged. For each table, the first column denotes the observer sequence number, the second column gives the age of the observer and the third column indicates whether or not the observer reported having previous experience in detecting and identifying military vehicles in the field. A value of 0 indicates no previous experience. Otherwise, a value of 1 is assigned to the variable 'Experience'. Column four gives the slide range at which initial detection occurred. Column five gives the adjusted values using the single value spacing method. This is appropriate since the observations were interval censored data, including covariates, from a photosimulation trial.

TABLE X
DATA FOR GENERAL LINEAR MODEL (NATO A)

Sequence Number	Age (years)	Experience	Initial Detection Ranges (m)	
			Raw Data	Adjusted Values Using Single Value Spacing
5	28	0	0	-226
15	37	1	0	-226
22	25	0	0	-226
25	32	1	452	-480
34	26	0	0	-226
42	45	0	0	-226
47	31	1	0	-226
50	44	0	0	-226
57	25	0	0	-226

63	24	1	508	-535
70	46	0	0	-226
73	20	0	0	-226
83	54	0	0	-226
85	35	0	0	-226
93	39	0	0	-226
99	23	0	0	-226

TABLE XI
DATA FOR GENERAL LINEAR MODEL (NATO B)

Sequence Number	**Age (years)**	**Experience**	**Initial Detection Ranges (m)**	
			Raw Data	**Adjusted Values Using Single Value Spacing**
4	36	0	500	-527
7	37	0	670	-701
16	41	1	500	-527
26	36	1	0	-223
28	25	1	0	-223
32	22	0	0	-223
38	20	0	0	-223
44	33	1	0	-223
52	34	1	0	-223
56	39	1	445	-473
64	23	1	0	-223
71	20	0	0	-223
75	23	0	0	-223
81	36	1	0	-223
89	28	0	445	-473
91	38	0	445	-473

TABLE XII
DATA FOR GENERAL LINEAR MODEL (NATO C)

Sequence Number	**Age (years)**	**Experience**	**Initial Detection Ranges (m)**	
			Raw Data	**Adjusted Values Using Single Value Spacing**
2	40	1	524	-551
10	38	1	0	-234
18	37	0	0	-234
19	26	1	578	-607
33	27	0	0	-234
39	36	0	696	-728
45	37	1	947	-979
53	48	0	524	-551

60	40	1	467	-496
61	25	1	1071	-1103
68	35	0	578	-607
76	20	0	1071	-1103
80	30	1	578	-607
90	41	1	524	-551
96	36	1	0	-234

58. The purpose of applying the General Linear Model to the data in Tables X, XI, and XII was to determine if there was a statistically significant difference between the mean initial detection performance for slide series NATO A, NATO B and NATO C. In addition, the test would determine if the variables age and experience had a statistically significant effect on observer performance. The General Linear Model procedure was applied using SAS. The results indicate that there was a statistically significant difference between the mean initial detection ranges of the three slide series and that there was no evidence that age and experience had a statistically significant effect on initial detection performance. The SAS code and output for this example are given in Annex C.

59. <u>Example of the General Linear Model with Paired Observations</u>: The data in Table XIII were collected during a photosimulation experiment in January 2003 at CFB Gagetown, Canada. Sixteen observers were shown a series of 20 slides starting at range 2000 m and steadily decreasing by 100 m. The targets of interest in this example were three camouflage nets deployed over military vehicles. The first column of Table XIII gives the observer sequence number and column 2 gives the observer age in years. Column 3 indicates whether or not the observer reported having previous experience in detecting and identifying military vehicles in the field. A value of 0 indicates no previous experience. Otherwise, a value of 1 is assigned to the variable 'Experience'. Columns 4, 6 and 8 give the slide range at which initial detections occurred. Columns 5, 7 and 9 give the adjusted data values using the single value spacing method. This is appropriate since the observations were interval censored data from a photosimulation trial and the analysis includes additional information about the observers.

TABLE XIII
DATA FOR GENERAL LINEAR MODEL (PAIRED OBSERVATIONS)

Sequence Number	Age (years)	Experience	Initial Detection Ranges (m)					
			Camouflage Net 1		Camouflage Net 2		Camouflage Net 3	
			Raw Data	Adjusted Values by Single Value Spacing	Raw Data	Adjusted Values by Single Value Spacing	Raw Data	Adjusted Values by Single Value Spacing
3	32	1	300	-350	300	-350	300	-350
11	33	1	400	-450	400	-450	400	-450
13	25	0	200	-250	200	-250	200	-250
24	31	1	600	-650	600	-650	600	-650
27	27	1	100	-150	100	-150	200	-250
31	33	0	0	-50	0	-50	200	-250
37	26	0	200	-250	200	-250	200	-250
43	39	0	300	-350	300	-350	300	-350
49	41	1	400	-450	400	-450	400	-450
58	48	1	600	-650	400	-450	500	-550
65	26	1	400	-450	400	-450	400	-450
72	26	0	300	-350	300	-350	700	-750
74	25	0	100	-150	300	-350	500	-550
84	36	1	300	-350	300	-350	300	-350
94	37	0	400	-450	400	-450	400	-450
98	21	0	1800	-1850	500	-550	500	-550

60. The purpose of applying the General Linear Model to the data in Table XIII was to determine if there was a statistically significant difference between the mean initial detection performance for the three camouflage nets. In addition, the test would determine if the variables age and experience had a statistically significant effect on observer performance. The General Linear Model procedure was applied using SAS. The results indicate that there was no statistically significant difference between the mean initial detection ranges of the three camouflage nets and that there was no evidence that the variables age and experience had a statistically significant effect on initial detection performance. The SAS code and output for this example are given in Annex C.

V. SUMMARY

61. This report discusses the analysis of data collected from different types of camouflage trials, including photosimulation experiments, live camouflage trials and video simulation. Since visual camouflage data are often statistically censored, the report gives procedures on how to adjust the raw observations that depend on the type of analysis to be carried out. Recommended methods for calculating descriptive statistics, including the median and first percentile of a dataset, are given. Statistical tests for different experimental designs are presented with examples. The two main characteristics that determine which statistical test is appropriate are the number of camouflage treatments being evaluated and whether or not the observations are independent or paired. For each case, the report recommends a statistical tool and gives an example of its use. Finally, the General Linear Model as an analysis tool for camouflage data is discussed and two examples are given.

VI. REFERENCES

1. Ashforth M., and Collins, J.H. (1991), "Determination of Detection Range by Analysis of Recorded Imagery", U.K. Ministry of Defence Technical Memorandum SCRDE/91/6, Stores and Clothing Research and Development Establishment, Colchester, Essex, June 1991.

2. Emond, E.J., Arseneau, L. and Karpoff, C., "Photosimulation Procedures and Analysis to Enhance Camouflage Assessment", Department of National Defence, Canada, ORD Project Report PR2002/16, Operational Research Division, Ottawa, Canada, November 2002.

3. Arseneau, L. and Emond, E.J., "The NATO SCI-095 Camouflage Assessment Trial", Department of National Defence, Canada, ORD Project Report PR2003/04, Operational Research Division, Ottawa, Canada, March 2003.

4. Hays, W.L. and Winkler, R.L., Statistics: Probability, Inference and Decision, Holt, Rinehart and Winston, Inc., 1971.

ANNEX A
ORD PROJECT REPORT PR 2003/15
SEPTEMBER 2003

VALUES OF $\alpha(N)$ FOR FIRST PERCENTILE CALCULATION

1. This annex gives the expected value of the first order statistic of a standardized Normal distribution, denoted by $\alpha(N)$, for use in the calculation of the first percentile of a set of camouflage data. The value depends only on N, the total number of observations in the dataset being evaluated.

N	**$\alpha(N)$**	**N**	**$\alpha(N)$**	**N**	**$\alpha(N)$**	**N**	**$\alpha(N)$**
1.	-	26.	-1.982	51.	-2.256	76.	-2.406
2.	-0.564	27.	-1.998	52.	-2.263	77.	-2.411
3.	-0.846	28.	-2.013	53.	-2.271	78.	-2.416
4.	-1.029	29.	-2.028	54.	-2.278	79.	-2.421
5.	-1.163	30.	-2.042	55.	-2.285	80.	-2.425
6.	-1.267	31.	-2.056	56.	-2.292	81.	-2.430
7.	-1.352	32.	-2.069	57.	-2.299	82.	-2.434
8.	-1.424	33.	-2.082	58.	-2.305	83.	-2.439
9.	-1.485	34.	-2.094	59.	-2.312	84.	-2.443
10.	-1.539	35.	-2.106	60.	-2.318	85.	-2.447
11.	-1.586	36.	-2.118	61.	-2.324	86.	-2.451
12.	-1.629	37.	-2.129	62.	-2.331	87.	-2.456
13.	-1.668	38.	-2.140	63.	-2.337	88.	-2.460
14.	-1.703	39.	-2.150	64.	-2.342	89.	-2.464
15.	-1.736	40.	-2.160	65.	-2.348	90.	-2.468
16.	-1.766	41.	-2.170	66.	-2.354	91.	-2.472
17.	-1.794	42.	-2.180	67.	-2.360	92.	-2.476
18.	-1.820	43.	-2.189	68.	-2.365	93.	-2.480
19.	-1.844	44.	-2.198	69.	-2.371	94.	-2.483
20.	-1.867	45.	-2.207	70.	-2.376	95.	-2.487
21.	-1.889	46.	-2.216	71.	-2.381	96.	-2.491
22.	-1.909	47.	-2.224	72.	-2.386	97.	-2.495
23.	-1.929	48.	-2.232	73.	-2.392	98.	-2.498
24.	-1.947	49.	-2.240	74.	-2.397	99.	-2.502
25.	-1.965	50.	-2.248	75.	-2.401	100.	-2.505

ANNEX B
ORD PROJECT REPORT PR 2003/15
SEPTEMBER 2003

DERIVATION OF THE FORMULA FOR CALCULATING THE FIRST PERCENTILE

1. This annex will present the derivation of a method for calculating the first percentile from a set of camouflage range data. The first percentile is defined as the point on the distribution of range values for a detection event at which an observer has probability 0.01 or one percent of detecting the target. The method outlined in this annex is simple to apply, gives intuitively reasonable results and requires minimal assumptions.

2. The range values or observations of interest must first be adjusted, if necessary, using the uniform spacing method to account for any data censoring that may exist. The adjusted range values are then sorted in decreasing order of magnitude. Let the sorted values be represented for mathematical purposes by: $X_{(1)}$, $X_{(2)}$, … , $X_{(N)}$. The observation $X_{(1)}$ is referred to as the first order statistic.

3. Next the median for this set of range values is calculated. Denote the median value by $X_{.50}$.

4. We make an assumption that the probability density function for the range of the detection event has a maximum value near its median and diminishes steadily at longer ranges, similar to a Normal distribution. This assumption is considered to be reasonable from both the theoretical and empirical points of view. Note that no assumption is required about the other half of the probability density function.

5. Given this assumption, we may approximate the lower or long range part of the probability density function for the range of a detection event using the Normal distribution. We will use two characteristics of the Normal distribution to derive an estimator of the first percentile for a set of camouflage range data which is consistent with the data, simple to apply and statistically sound.

6. For a sample of size N from a standardized Normal distribution, the expected value of the first order statistic, denoted here by $\alpha(N)$, can be calculated using numerical methods. The results of this calculation for a standardized Normal distribution and for values of N from 2 to 100 have been tabulated in Annex A. In order to derive an estimator for the first percentile of a dataset, we equate a sample statistic to its

expected value, as is done in the method of moments. The unknowns in the equation are the parameters μ and σ, representing the mean and standard deviation of the Normal distribution of the upper half of the range values.

$$\frac{X_{(1)} - \mu}{\sigma} \approx \alpha(N)$$

7. Another characteristic of the Standard Normal distribution is that the theoretical first percentile value is 2.326 standard deviations below the mean.

$$X_{.01} = \mu - 2.326\sigma$$

8. Next we combine the two equations to eliminate the parameter σ. In addition, we replace the parameter μ with the estimated sample median, $X_{.50}$. Note that the mean and median are the same for the Normal distribution.

$$X_{.01} \approx X_{.50} - 2.326\left(\frac{X_{(1)} - X_{.50}}{\alpha(N)}\right)$$

9. Given a table of values for α(N), this equation can be used to estimate the first percentile from a sample of N detection event observations.

10. Note that the estimated first percentile depends on the difference between the median and the largest observed value. In rare cases it is possible for the formula to produce inconsistent values when this difference is large for one set of data and relatively small for another. In a real example the estimated first percentile for the recognition range for a set of data was 1018 m while the associated identification range first percentile was 1034 m. This inconsistent result was caused by distributional differences between the two sets of data. The inconsistency was resolved by reducing the value of the first percentile identification range from 1034 m to 1018 m.

SAS CODE AND OUTPUT FOR GENERAL LINEAR MODEL EXAMPLE

1. The SAS code for the General Linear Model example with independent observations is given below.

```
DATA NATO_ABC ;
INPUT series $ age experience $ detect_range ;
CARDS ;

     NATO_A  28  0   -226
     NATO_A  37  1   -226
     NATO_A  25  0   -226
     NATO_A  32  1   -480
     NATO_A  26  0   -226
     NATO_A  45  0   -226
     NATO_A  31  1   -226
     NATO_A  44  0   -226
     NATO_A  25  0   -226
     NATO_A  24  1   -535
     NATO_A  46  0   -226
     NATO_A  20  0   -226
     NATO_A  54  0   -226
     NATO_A  35  0   -226
     NATO_A  39  0   -226
     NATO_A  23  0   -226
     NATO_B  36  0   -527
     NATO_B  37  0   -701
     NATO_B  41  1   -527
     NATO_B  36  1   -223
     NATO_B  25  1   -223
     NATO_B  22  0   -223
     NATO_B  20  0   -223
     NATO_B  33  1   -223
     NATO_B  34  1   -223
     NATO_B  39  1   -473
     NATO_B  23  1   -223
     NATO_B  20  0   -223
     NATO_B  23  0   -223
     NATO_B  36  1   -223
     NATO_B  28  0   -473
     NATO_B  38  0   -473
     NATO_C  40  1   -551
     NATO_C  38  1   -234
     NATO_C  37  0   -234
     NATO_C  26  1   -607
     NATO_C  27  0   -234
     NATO_C  36  0   -728
     NATO_C  37  1   -979
     NATO_C  48  0   -551
```

```
     NATO_C  40  1   -496
     NATO_C  25  1 -1103
     NATO_C  35  0   -607
     NATO_C  20  0 -1103
     NATO_C  30  1   -607
     NATO_C  41  1   -551
     NATO_C  36  1   -234
;

PROC GLM;
  CLASS series experience ;
  MODEL detect_range = series age experience ;
RUN;
```

2. The SAS output for the code given above is as follows.

```
                                The SAS System
                               The GLM Procedure

                           Class Level Information

                        Class      Levels      Values

                        series       3    NATO_A NATO_B NATO_C

                        experience   2           0 1


                         Number of observations    47


    Dependent Variable: detect_range

                                        Sum of
    Source              DF          Squares     Mean Square   F Value   Pr > F

    Model               4        907437.517      226859.379     5.48    0.0012

    Error               42      1737450.185      41367.862

    Corrected Total     46       2644887.702


    R-Square     Coeff Var      Root MSE    detect_range Mean

    0.343091     -51.94747      203.3909           -391.5319


    Source              DF     Type I SS     Mean Square    F Value    Pr > F

    series               2    896717.3313     448358.6656    10.84     0.0002
    age                  1      8455.8123       8455.8123     0.20     0.6535
    experience           1      2264.3733       2264.3733     0.05     0.8162


    Source              DF     Type III SS    Mean Square    F Value    Pr > F

    series               2     830259.1580     415129.5790    10.04     0.0003
    age                  1       9153.9287       9153.9287     0.22     0.6405
    experience           1       2264.3733       2264.3733     0.05     0.8162
```

3. The SAS code for the General Linear Model example with paired observations is given below.

```
DATA cam_nets;
INPUT age experience $ detect_range1 detect_range2 detect_range3 ;
CARDS ;

    32  1   -350   -350   -350
    33  1   -450   -450   -450
    25  0   -250   -250   -250
    31  1   -650   -650   -650
    27  1   -150   -150   -250
    33  0     -50     -50   -250
    26  0   -250   -250   -250
    39  0   -350   -350   -350
    41  1   -450   -450   -450
    48  1   -650   -450   -550
    26  1   -450   -450   -450
    26  0   -350   -350   -750
    25  0   -150   -350   -550
    36  1   -350   -350   -350
    37  0   -450   -450   -450
    21  0  -1850    -550  -550
;

PROC GLM  ;
   CLASS experience ;
   MODEL detect_range1 detect_range2 detect_range3 = age experience
        / NOUNI ;
   REPEATED target 3 ;
RUN;
```

4. The SAS output for the code given above is as follows.

```
                    The SAS System
                  The GLM Procedure

               Class  Level Information

            Class        Levels      Values

            experience     2           0 1

             Number of observations    16

         Repeated Measures Analysis of Variance

          Repeated Measures Level Information
```

```
                                            detect_   detect_   detect_
                    Dependent Variable      range1    range2    range3

                     Level of target           1        2        3


       Manova Test Criteria and Exact F Statistics for the Hypothesis of no
target Effect
                      H = Type III SSCP Matrix for target
                             E = Error SSCP Matrix

                                       S=1    M=0    N=5

Statistic                      Value      F Value    Num DF    Den DF   Pr > F

Wilks' Lambda               0.88445286     0.78         2        12     0.4787
Pillai's Trace              0.11554714     0.78         2        12     0.4787
Hotelling-Lawley Trace      0.13064251     0.78         2        12     0.4787
Roy's Greatest Root         0.13064251     0.78         2        12     0.4787


     Manova Test Criteria and Exact F Statistics for the Hypothesis of no
target*age Effect
                    H = Type III SSCP Matrix for target*age
                             E = Error SSCP Matrix

                                       S=1    M=0    N=5

Statistic                     Value       F Value    Num DF    Den DF   Pr > F

Wilks' Lambda               0.94506215     0.35         2        12     0.7125
Pillai's Trace              0.05493785     0.35         2        12     0.7125
Hotelling-Lawley Trace      0.05813147     0.35         2        12     0.7125
Roy's Greatest Root         0.05813147     0.35         2        12     0.7125


 Manova Test Criteria and Exact F Statistics for the Hypothesis of no
target*experience Effect
                 H = Type III SSCP Matrix for target*experience
                             E = Error SSCP Matrix

                                       S=1    M=0    N=5

Statistic                        Value    F Value    Num DF    Den DF    Pr > F

Wilks' Lambda               0.88458279       0.78         2        12    0.4791
Pillai's Trace              0.11541721       0.78         2        12    0.4791
Hotelling-Lawley Trace      0.13047644       0.78         2        12    0.4791
Roy's Greatest Root         0.13047644       0.78         2        12    0.4791





                     Repeated Measures Analysis of Variance
                Tests of Hypotheses for Between Subjects Effects

Source                  DF     Type III SS     Mean Square    F Value   Pr > F

age                     1       21523.836       21523.836       0.15    0.7022
experience              1       18336.650       18336.650       0.13    0.7240
Error                  13     1830976.164      140844.320
```

```
                    Repeated Measures Analysis of Variance
           Univariate Tests of Hypotheses for Within Subject Effects


Adj Pr > F
Source           DF    Type III SS    Mean Square   F Value   Pr > F    G - G
H - F

target            2      58208.670      29104.335      0.61   0.5506   0.4652
0.4915
target*age        2      43273.587      21636.794      0.45   0.6400   0.5330
0.5659
target*experience  2      9838.458       4919.229      0.10   0.9023   0.7813
0.8232
Error(target)    26    1239226.413      47662.554


                     Greenhouse-Geisser Epsilon    0.5608
                     Huynh-Feldt Epsilon           0.6712
```

DOCUMENT CONTROL DATA
(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared e.g. Establishment Sponsoring a contractor's report, or tasking agency, are entered in Section 8). Operational Research Division Department of National Defence Ottawa, Ontario K1A 0K2	2. SECURITY CLASSIFICATION (overall security classification of the document, including special warning terms if applicable) UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title)

A GUIDEBOOK FOR THE ANALYSIS OF CAMOUFLAGE DATA

4. AUTHORS (last name, first name, middle initial)

Arseneau, Dr. Lise; Emond, Edward J.

5. DATE OF PUBLICATION (month Year of Publication of document) SEPTEMBER 2003	6a. NO OF PAGES (total containing information. Include Annexes, Appendices, etc.) 43	6b. NO OF REFS (total cited in document) 4

7. DESCRIPTIVE NOTES (the category of document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Project Report

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address).

DSSPM and NATO SCI-095

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) 3552-1	9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written.)
10a. ORIGINATOR's document number (the official document number by which the document is identified by the originating activity. This number must be unique to this document.) ORD Project Report PR 2003/15	10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification.)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors: further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the test is bilingual).

This report was produced for NATO Task Group SCI-095 “Enhancement of Camouflage Assessment Techniques”. It describes the analysis of data collected from different types of camouflage trials, including photosimulation experiments, live camouflage trials and video simulation. Since visual camouflage data are often statistically censored, the report offers procedures on how to adjust the raw observations, depending on the type of analysis to be carried out. Recommended methods for calculating descriptive statistics, including the median and first percentile of a dataset, are provided. Statistical tests for different experimental designs are presented with examples. The two main characteristics that determine which statistical test is most appropriate are the number of camouflage treatments being evaluated and whether the observations are independent or paired. For each case, the report recommends a statistical tool and gives an example of its use. Finally, the General Linear Model as an analysis tool for camouflage data is discussed and two examples are given.

Canad

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified . If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Photosimulation
Camouflage Analysis
Censored Data
Significance Testing
First Percentile