

AD A 024307

9

SACLANT ASW  
RESEARCH CENTRE

MEMORANDUM *repts.*

12

14  
SACLANTCEN Memorandum  
- SM - 79

FC

6  
APPROXIMATE CONFIDENCE LIMITS FOR A PRODUCT  
OF INDEPENDENT POSITIVE RANDOM VARIABLES  
AS APPLIED TO ASW EXERCISE DATA .

10  
by  
NIELS/BACHE

11  
1 MARCH 1976

12 37 p.

AD DDC  
MAY 18 1976  
RECEIVED  
C

NORTH  
ATLANTIC  
TREATY  
ORGANIZATION

VIALE SAN BARTOLOMEO 400  
I - 19025 - LA SPEZIA, ITALY

This document is unclassified. The information it contains is published subject to the conditions of the legend printed on the inside cover. Short quotations from it may be made in other publications if credit is given to the author(s). Except for working copies for research purposes or for use in official NATO publications, reproduction requires the authorization of the Director of SACLANTCEN.

DISTRIBUTION STATEMENT A  
Approved for public release;  
Distribution Unlimited

319950 ✓

This document is released to a NATO Government at the direction of the SACLANTCEN subject to the following conditions:

1. The recipient NATO Government agrees to use its best endeavours to ensure that the information herein disclosed, whether or not it bears a security classification, is not dealt with in any manner (a) contrary to the intent of the provisions of the Charter of the Centre, or (b) prejudicial to the rights of the owner thereof to obtain patent, copyright, or other like statutory protection therefor.

2. If the technical information was originally released to the Centre by a NATO Government subject to restrictions clearly marked on this document the recipient NATO Government agrees to use its best endeavours to abide by the terms of the restrictions so imposed by the releasing Government.

ACCESS TO	
NTIS	Write Section <input checked="" type="checkbox"/>
DTC	Buy Section <input type="checkbox"/>
UNCLASSIFIED	<input type="checkbox"/>
JUSTIFICATION	
BY	
DISTRIBUTION AVAILABILITY CODES	
Dist.	Avail. W. & C. Code
A	

Published by



SACLANTCEN MEMORANDUM SM-79

NORTH ATLANTIC TREATY ORGANIZATION  
SACLANT ASW Research Centre ✓  
Viale San Bartolomeo 400  
I 19026 - La Spezia, Italy

APPROXIMATE CONFIDENCE LIMITS FOR A PRODUCT  
OF INDEPENDENT POSITIVE RANDOM VARIABLES  
AS APPLIED TO ASW EXERCISE DATA

by

Niels Bache

1 March 1976

This memorandum has been prepared within the SACLANTCEN Operational Research Group and does not necessarily represent the considered opinion of the SACLANT ASW Research Centre, of SACLANT, or of NATO.



R. Nagelhout  
Group Leader

**Manuscript Completed:**  
**22 August 1975**

## TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT	1
PREFACE	1
INTRODUCTION	2
1. GENERAL METHOD	4
1.1 Definition of the Problem	4
1.2 Outline of the Method	4
1.3 Transformation	5
1.4 Cumulants	6
1.5 Inversion	7
1.6 A Limitation of the Method	8
2. APPLICATIONS	8
2.1 Confidence Limits for Products of Probabilities	8
CONCLUSION	13
REFERENCES	14
APPENDIX A — ASYMPTOTIC EXPANSION FOR THE INVERSE FUNCTION OF AN ARBITRARY DISTRIBUTION FUNCTION	15
APPENDIX B — THE FIRST SIX CUMULANTS	17
APPENDIX C — EVALUATING A DEFINITE INTEGRAL	20
APPENDIX D — LOCATION OF THE MAXIMUM ERROR	21
APPENDIX E — TWO NUMERICAL EXAMPLES	23
APPENDIX F — CONFIDENCE LIMITS OF A FUTURE OBSERVED VALUE OF PRODUCTS OF PROBABILITIES, BASED ON A SAMPLE ALREADY TAKEN	27

### List of Figures

1. Definition of confidence limit.	4
2. The lognormal distribution.	5
E.1 A sample of detections and opportunities.	24
E.2 Interval.	25
E.3 Confidence limits.	26
F.1 Limits for an observed value.	29

LIST OF SYMBOLS IN MAIN TEXT

$c_i$	detection probability in range bracket $i$
$C$	a positive random variable
$Y$	the random variable $Y = \ln C$
$C_i$	a positive random variable, which is a factor number $i$ in a product $C$ , $C_i$ and $C_k$ are statistically independent, $i \neq k$ , $i = 1, \dots, N$
$Y_i$	the random variable $Y_i = \ln C_i$
$f_i(c_i)$	density function of $C_i$ with argument $c_i$
$c_p$	confidence limit for $C$ corresponding to the confidence level $p$
$y_p$	confidence limit for $Y$ corresponding to the confidence level $p$
$p$	confidence level or coefficient
$g(c)$	density function of $C$ with argument $c$
$h(y)$	density function of $Y$ with argument $y$
$X$	a random variable
$f(x)$	density function of $X$ with argument $x$
$i$	the imaginary unit $\sqrt{-1}$ , also used as a suffix
$j$	a suffix
$\varphi(t)$	the characteristic function with argument $t$ of a random variable
$t$	the argument in characteristic functions, also used as a parameter

List of Symbols (Cont'd)

- $\varphi_g(t)$  the characteristic function of a function  $g$  of a random variable
- $K_r$  the cumulant of  $r^{\text{th}}$  order of a distribution
- $\mu_r$  the moment of  $r^{\text{th}}$  order of a distribution
- $I_i(s, t)$  the integral  $\int (\ln c_i)^s c_i^{it} f_i(c_i) dc_i$ ,  $s \geq 0$ , integer
- $B(x|n, p)$  the binomial distribution with argument  $x$  and parameters  $n$  and  $p$
- $\gamma(t) \cdot dt$  probability of success in a small interval  $dt$  around the parameter  $t$
- $g(\gamma)$  density function of  $\gamma$ ,  $\gamma \geq 0$
- $K$  negative constant
- $Q(x, n)$  the sum  $\sum_{j=0}^{n-x} (-1)^j \binom{n-x}{j} \ln(j+n)$ ,  $1 \leq x \leq n-1$
- $\sigma$  standard deviation of a normal distribution

APPROXIMATE CONFIDENCE LIMITS FOR A PRODUCT  
OF INDEPENDENT POSITIVE RANDOM VARIABLES  
AS APPLIED TO ASW EXERCISE DATA

by

Niels Bache

ABSTRACT

A general, but approximate, method is suggested to determine confidence limits for a product of independent positive random variables. The method is developed and the accuracy of the approximation is discussed in one of its applications: the Bayesian confidence limits for the product of  $N$  binomial parameters. Four applications (including the above), in which the factors in the product are probabilities are given together with numerical examples. These are from the antisubmarine warfare (ASW) field and constitute to the evaluation of operational data in that field and statistical data in many other fields.

PREFACE

This work is done as part of the exercise research programme at SACLANTCEN and supports the evaluation of measures of operational ASW performance collected in the "SACLANTCEN Compendium of NATO ASW Exercises",



## INTRODUCTION

In SACLANTCEN's analysis of the data obtained from NATO's ASW exercises, as in similar statistical analyses for other purposes, probabilities are frequently used as measures of performance. For example, the performance of shipborne sonars in detecting submarines can be measured statistically by recording the frequencies of detection at different ranges and using those data to estimate the probabilities of detection at selected ranges of interest. However, like all statistically-derived values, these estimates have little value to the user, and may even mislead him, unless they can be provided with some indication of the statistical uncertainties resulting from the inadequacies of the basic data. The purpose of this paper and of a preceding paper [Ref. 1] by the same author is to present a method of giving approximate confidence limits to such estimates of probability.

The original problem [Ref. 1] was to find confidence limits for the probability that a sonar-fitted surface ship would detect a submarine by its sonar outside a given range from the ship to the submarine. To that end a sample of relative tracks between the ship and submarine was collected, including notes about detections and opportunities for detection as functions of range between ship and submarine. The range was divided into brackets; calling  $c_i$  the probability of detection in range bracket  $i$ , the probability of detection outside a given range could be written as

$$1 - \prod_i (1 - c_i),$$

where the product was taken over all brackets outside the given range.

We would like to find confidence limits for that expression, given some distribution for  $c_i$ .

It was noted that this problem was a special case of a more general problem: to find confidence limits for a product of independent positive random variables. If one could solve this problem one would not only have solved the present problem but many other related problems involving a product of random variables.

Considerable work has been done in this area. The only exact method known to the author is that by Springer and Thompson [Ref. 2], which assumes a particular distribution for each factor. Other methods [Refs. 1, 3, 4] are both approximate and assume particular distributions for the factors. The purpose of this paper is to present a method that does not assume particular distributions for the factors. The method however is only approximate.

All data collectors, who are trying to estimate products of random (stochastic) variables on the basis of sampling experiments on the individual factors, have an interest in confidence limits. As examples, one can mention:

- a) Estimation of sonar, or radar, detection probabilities as a function of range or time.
- b) The use of a product of probabilities of detection, correct classification, localization, and successful attack, as a measure of effectiveness for a tactical process.
- c) Estimation of failure rates (from small samples) of series, parallel and series-parallel systems (electronic apparatus, missiles, aircraft, etc.) in reliability and quality control problems.

The author believes that it is more important to present confidence limits for a measured quantity than to give an estimate of it. At once it must be said that it is generally more difficult to determine the confidence limits than to determine an estimate, because confidence limits require some knowledge of the distribution of the quantity whereas an estimate often does not. Confidence limits contain more information than an estimate; having confidence limits one can always find an estimate but not vice versa. Confidence limits are important when comparing two measures, such as results from two experiments (in the form of figures or curves), as they can give a first indication of whether the two quantities could stem from the same population or not (however a proper test has to be performed in the end). Estimates for the two quantities give no information in that respect.

This paper generalizes the determination of confidence limits for a product of probabilities, so as to cover problems involving a product of independent positive random variables in general.

Because of this generalization, the problem posed by the present author in Ref. 1 is solved with fewer approximations.

The opportunity is also taken here to correct an error in Sect. 1.4 of Ref. 1, which should read:

Equation 1 to be used as an exact expression in both the discrete and the continuous case.

Equation 3 to be used only as an approximation for the easy calculation of the confidence limits in Ch. 2.

However this latter approximation is no longer needed due to the improved method presented in the present memorandum, which in fact, should now be regarded as a replacement for Ref. 1.

## 1. GENERAL METHOD

### 1.1 Definition of the Problem

We want confidence limits for the random variable  $C$ , which is given by

$$C = \prod_{i=1}^N C_i, \quad [\text{Eq. 1}]$$

where  $C_i$  is a set of independent positive random variables with density functions  $f_i(c_i)$ .

We denote a random variable by an upper-case letter (e.g.  $C$ ) and use the corresponding lower-case letter (e.g.  $c$ ) for a particular value that the random variable assumes. The problem is to find a confidence limit for  $C$ ;  $c_p$ , such that

$$\text{Prob}\{C < c_p\} = p, \quad [\text{Eq. 2}]$$

$p$  is usually called the confidence level (or coefficient) and has to be decided in advance.

$g(c)$  is the density function of  $C$  and  $c_p$  means that  $c$  is dependent on the chosen  $p$ .

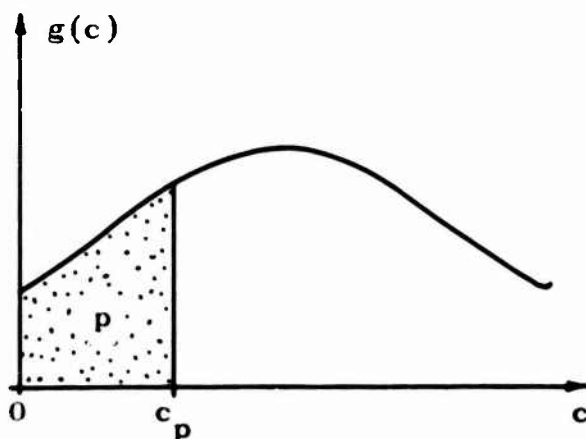


FIG. 1 DEFINITION OF CONFIDENCE LIMIT

### 1.2 Outline of the Method

The method presented in this paper can be divided into four stages:

- a. Transform the problem of finding confidence limits for a product of independent positive random variables to the problem of finding confidence limits for a sum of new independent (not necessarily positive) random variables by means of a logarithmic transformation.

- b. Find the characteristic function (related to the moment generating function) as a sum of independent random variables. When this is done, the distribution of the sum is in principle determined.
- c. Determine the cumulants (related to the moments) of this distribution by means of the characteristic function.
- d. Apply a method of finding the inverse of a distribution function by means of its cumulants.

### 1.3 Transformation

Already at this stage something can be said about  $g(c)$ : under fairly general conditions [Ref. 5],  $g(c)$  will approach the lognormal distribution as  $N \rightarrow \infty$ , independent of the form of the density functions for the factors  $f_i(c_i)$ . This is a direct consequence of the central limit theorem.

Figure 2 shows the shape of the lognormal distribution.

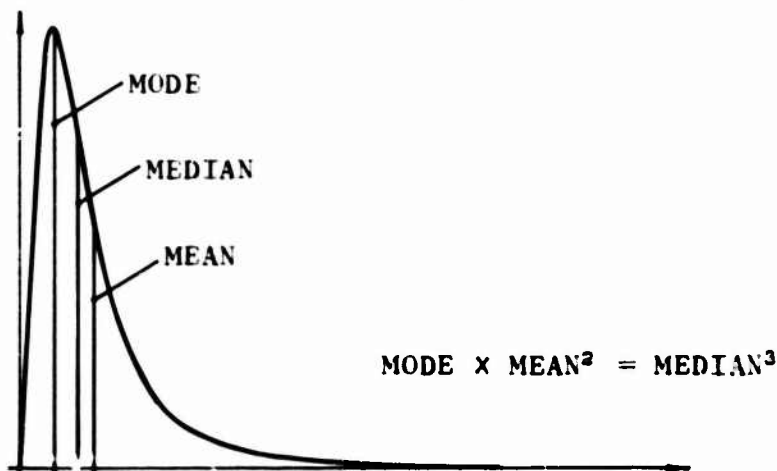


FIG 2 THE LOGNORMAL DISTRIBUTION

In most practical cases it is impossible or very difficult to find  $g(c)$  by analytical means (for example, transformation by the characteristic function or by the Jacobian determinant) for most of the distributions (binomial, poisson, normal, etc.) assumed for  $f_i(c_i)$  (the rectangular and beta distributions are exceptions in some cases, see Ref. 3).

The approach in this paper uses a general method to determine an arbitrary distribution developed by Cornish and Fisher in 1937 [quoted in Refs. 6, 7 and 9], which takes as its starting point the normal distribution and finds a transformation between this well-known distribution and the actual distribution. This transformation is a function of the higher moments of the actual distribution. The Cornish-Fisher approximation [copied from Ref. 9] is given in Appendix A.

The important point at this stage is that the method is based on the normal distribution. Therefore we will transform the stochastic variable  $C$  into a new stochastic variable  $Y$  so that instead of  $g(c)$  we will have a new density function  $h(y)$ , which will approach the normal distribution as  $N \rightarrow \infty$ . This transformation is simply  $Y = \ln C$ :

Adding the additional assumption  $C_i > 0$  we are able to define a new set of independent random variables  $Y_i$ , as

$$Y_i = \ln C_i \quad [\text{Eq. 3}]$$

and

$$Y = \sum_{i=1}^N Y_i . \quad [\text{Eq. 4}]$$

Then Eq. 2 is equivalent to (the function  $e^x$  is single valued and monotone)

$$\text{Prob}\{Y < y_p\} = p ,$$

where

$$y_p = \ln c_p . \quad [\text{Eq. 5}]$$

The problem is now to find  $y_p$ , the confidence limit for a sum  $Y$  of independent random variables  $Y_i$ , where the density function  $h(y)$  of  $Y$  will approach the normal distribution as  $N \rightarrow \infty$ , because of the central limit theorem.

We have now assured that the Cornish-Fisher method will be accurate for large  $N$ . For small and moderate  $N$  nothing can be said in general, except that the more  $h(y)$  differs from normality, the more inaccurate the method will be. The cases of small  $N$  is discussed later in the special case where  $f_i(c_i)$  is beta distributed.

#### 1.4 Cumulants

In this section we will make use of the characteristic function of a random variable,  $X$ , defined as

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx , \quad [\text{Eq. 6}]$$

where  $t$  is a real variable,  $i = \sqrt{-1}$  ( $i$  is also used as a suffix, this is believed not to create confusion) and  $f(x)$  is the density function of  $X$ .

We will also use the characteristic function of a function of the random variable  $X$ ,  $g(x)$ , defined as

$$\varphi_g(t) = \int_{-\infty}^{\infty} e^{itg(x)} f(x) dx. \quad [\text{Eq. 7}]$$

The cumulants  $K_1, K_2, \dots, K_r, \dots$  are defined formally by the identity in  $t$

$$\begin{aligned} & \exp\left\{K_1 t + K_2 \frac{t^2}{2!} + \dots + K_r \frac{t^r}{r!} + \dots\right\} = \\ & = 1 + \mu_1' t + \mu_2' \frac{t^2}{2!} + \dots + \mu_r' \frac{t^r}{r!} + \dots, \end{aligned}$$

where  $\mu_r'$  is the  $r^{\text{th}}$  moment about the origin:

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x) dx,$$

which is assumed to exist.

Provided an expansion in power series exists for  $\ln \varphi(t)$  it can be shown that

$$K_1 it + K_2 \frac{(it)^2}{2!} + \dots + K_r \frac{(it)^r}{r!} + \dots = \ln \varphi(t). \quad [\text{Eq. 8}]$$

Therefore

$$K_r = \left[ \frac{d^r}{d(it)^r} \ln \varphi(t) \right]_{t=0};$$

$\ln \varphi(t)$  is called the cumulant generating function or just the cumulative function.

As mentioned before, the Cornish-Fisher method is a transformation of a variable. The transformation used is a function of the first six cumulants of the distribution  $h(y)$  (see Appendix A).

Appendix B derives a means by which these first six cumulants of the distribution of  $Y$  may be determined from a knowledge of the distribution functions of the  $C_i$ .

## 1.5 Inversion

Once the first six cumulants of the distribution of  $Y$  are available to us we may make use of the Cornish-Fisher expansion (Appendix A) to find an approximation for the required confidence

limits, i.e. that value  $y_p$  from which  $\text{prob}(Y > y_p) = p$ . This can be transformed back into a value of  $c_p$  by means of Eq. 5.

It is of interest to note that Cornish-Fisher asymptotic expansions can also be used to derive an estimate of the distribution function of a stochastic variable, which is the sum of random variables drawn from different distributions, so in principle we are not limited to finding only the confidence limits, although this is possibly the most important application.

## 1.6 A Limitation of the Method

Applications of the method are dependent on whether the integral, which appears in Appendix B,

$$I_i(s, 0) \equiv \int (\ln c_i)^s \cdot f_i(c_i) \cdot dc_i, \quad s \geq 1 \text{ integer}$$

can be evaluated (exactly or approximately).

Since  $C_i \geq 0$  by definition, the only point where  $\ln c_i$  is not defined is  $c_i = 0$ . Therefore a necessary condition, which is not always fulfilled, is that

$$f_i(c_i) \xrightarrow{c_i \rightarrow 0^+} 0$$

in such a way that the integral  $I_i(s, 0)$  is finite. A way of avoiding this problem is to redefine  $C_i$  such that  $C_i > 0$ . Examples of both these cases will be given later. If  $C_i$  has no upper limit then we must also have as a necessary condition, that  $I_i(s, 0)$  be finite. If  $C_i$  is a discrete random variable we must define  $C_i > 0$  or have  $f_i(0) = 0$ .

## 2. APPLICATIONS

### 2.1 Confidence Limits for Products of Probabilities

#### 2.1.1 Product of probabilities

In this case the factors  $C_i$  in the product  $C$  are probabilities, so  $f_i(c_i) = 0$  for  $c_i$  outside the interval  $(0, 1)$ . The family of integrals that we must evaluate in order to apply the method can thus be written as

$$I_i(s, 0) = \int_0^1 (\ln c_i)^s \cdot f_i(c_i) \cdot dc_i. \quad [\text{Eq. 9}]$$

Let us assume that for each factor  $C_i$  there have been run a number of trials to estimate  $C_i$ . Let the number of trials be  $n_i$  and the resultant number of successes be  $x_i$ , where  $x_i \leq n_i$ . The number  $x_i$  will then be binomially distributed.

If  $c_i$  is the true, but unknown, probability on which we would like to place confidence limits, then the probability of  $x_i$  successes from  $n_i$  trials is

$$B(x_i | n_i, c_i) = \binom{n_i}{x_i} \cdot c_i^{x_i} \cdot (1 - c_i)^{n_i - x_i}, \quad [\text{Eq. 10}]$$

We adopt a Bayesian approach to find the posterior density distribution  $f_i(c_i | x_i, n_i)$ , in which the number of successes  $x_i$  of the  $n_i$  experiments are given. The posterior distribution will be used in Eq. 9 instead of  $f_i(c_i)$ . The density function  $f_i(c_i | x_i, n_i)$  expresses our limited knowledge about the probability  $c_i$ .

From Bayes' theorem we have

$$f_i(c_i | x_i, n_i) = \frac{B(x_i | n_i, c_i) f_i(c_i)}{\int_0^1 B(x_i | n_i, c_i) f_i(c_i) dc_i}. \quad [\text{Eq. 11}]$$

Without any a priori knowledge of  $C_i$  we have, according to Bayes' postulate:

$$f_i(c_i) = 1, \quad 0 \leq c_i \leq 1. \quad [\text{Eq. 12}]$$

This is the most important assumption in this application and means that, before we have carried out any trials to determine the value of  $C_i$ , any value of  $C_i$  between 0 and 1 must be accepted as equally likely.

Substituting Eq. 10 and Eq. 12 into Eq. 11 gives

$$f_i(c_i | x_i, n_i) = (n_i + 1) \binom{n_i}{x_i} c_i^{x_i} (1 - c_i)^{n_i - x_i}, \quad [\text{Eq. 13}]$$

which is the beta distribution.

When we use this density function in the definition of the family of integrals, Eq. 9, we obtain

$$I_i(s, 0) = (n_i + 1) \binom{n_i}{x_i} \int_0^1 (\ln c_i)^s c_i^{x_i} (1 - c_i)^{n_i - x_i} dc_i \quad [\text{Eq. 14}]$$



In Appendix C this is shown to reduce to the computationally-easier form of

$$I_i(s, 0) = (-1)^s s! (n_i+1) \binom{n_i}{x_i} \sum_{j=0}^{n_i-x_i} (-1)^j \frac{\binom{n_i-x_i}{j}}{(x_i+j+1)^{s+1}}. \quad [\text{Eq. 15}]$$

We are now in a position to use the general method given previously.

A computer program has been written and to check and determine the accuracy of the method for this application a comparison has been made between values obtained by the method and corresponding exact values for the important cases of small sample sizes,  $n_i$ , and few factors,  $N$ , in the product. The exact values have been obtained by deriving an analytical expression in each case.

TABLE 1  
MAXIMUM ABSOLUTE ERROR FOR UPPER AND LOWER CONFIDENCE LIMITS  
OF 10% EACH, MULTIPLIED BY 10<sup>3</sup>

$n_i \backslash N$	1	2	3	4
0	11.7	1.4	0.25	0.045
1	6.2	0.90	0.22	0.054
2	4.2	0.65	0.18	-
3	3.2	0.51	-	-

Each cell in the matrix above contains the maximum absolute error for all combinations of  $x_i$  and  $n_i$ , where at least one  $n_i$  has the value shown in the left-hand column. Table 1 is valid only for upper and lower confidence limits of 10% each and only for this application.

The main conclusion from Table 1 is that the accuracy is sufficient for all practical applications, because usually  $N$  will be greater than 2 or 3. One could have feared that the error in Table 1 would begin to increase outside the frame of the table. This is not so, as shown in Appendix D. The highest value of the error will be in the cell  $(n_i, N) = (0, 1)$ . However, Appendix D has not taken into account the computational inaccuracy.

Simpler and more accurate methods exist for the case  $N=1$  and finding confidence limits for cases where the values of  $n_i$  are large is usually of little interest.

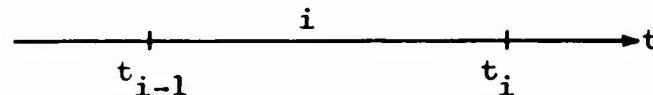
The material of this section makes use of Bayes's theorem (or formula) and Bayes's postulate. Bayes's postulate is far from being universally accepted and therefore the application in this section should be seen in this light.

An example is given in Appendix E.

### 2.1.2 Cumulative probability curve

The application in this section is also based on Bayes's theorem but instead of using Bayes's postulate we use an extension of this postulate proposed by H. Jeffrey (1948) [Ref. 8] and therefore the method in this section is on an even looser ground than the application in Sect. 2.1.1.

Let a process (for example a detection process) be dependent on a continuous parameter  $t$  (for example a distance or time) and let the index  $i$  in Sect. 2.1.1 represent an interval of the parameter  $t$  from  $t_{i-1}$  to  $t_i$ :



and call  $x_i$  the number of failures of  $n_i$  trials run when  $t$  was known to be in the interval  $i$ . Then, following the notation in Sect. 2.1.1, where  $c_i$  now is the (unknown) probability of failure for the interval  $i$ ,

$$1 - \prod_{i=1}^N C_i \quad [\text{Eq. 16}]$$

represents the probability of at least one success in all intervals  $i = 1, \dots, N$ . We have assumed that an outcome (success, failure) in one interval is independent of outcomes in other intervals.

Section 2.1.1 can be used again to find confidence limits for Eq. 16, except that  $C_i$  is no longer a random variable of which we have no a priori knowledge. It is possible in this case to say something about  $c_i$  a priori because of the independence assumption stated above.

Define

$\gamma(t) \cdot dt$  = probability of success in a small interval  $dt$  around  $t$ .

$c(t_{i-1}, t)$  = probability of failure in the interval  $t_{i-1}$  to  $t$ .

Then we can set up a differential equation using the independence assumption

$$c(t_{i-1}, t) = c(t_{i-1}, t - dt)[1 - \gamma(t)dt]$$

leading to

$$c_i = \exp \left\{ - \int_{t_{i-1}}^{t_i} \gamma(t) dt \right\} \quad [\text{Eq. 17}]$$

$\gamma$  is now the basic random variable of which we assume to have no a priori knowledge.

It is noted that Eq. 17 is not a one-to-one transformation between  $c_i$  and  $\gamma(t)$ . For a given  $c_i$  there is an infinite number of functions  $\gamma(t)$  that satisfy Eq. 17, but not vice-versa. Therefore Eq. 17 is regarded to contain some information about  $c_i$ , and  $\gamma$  is regarded to be a more basic variable than  $c_i$ .

Since  $0 \leq \gamma < \infty$  we must use the Jeffrey extension [Ref.8] of Bayes' postulate, which states that if  $\gamma$  ranges from 0 to  $\infty$  the prior distribution is taken as proportional to  $d\gamma/\gamma$  (instead of just to  $d\gamma$  as in the original postulate of Bayes, which deals only with a finite definition range):

$$g(\gamma)d\gamma = \frac{d\gamma}{\gamma}, \quad 0 \leq \gamma.$$

To find the density function  $f_i(c_i)$  of  $c_i$  the (Jacobian) transformation

$$f_i(c_i) |dc_i| = g(\gamma) |d\gamma|$$

gives, instead of Eq. 12,

$$f_i(c_i) dc_i = \frac{K}{c_i \ln c_i} dc_i, \quad K \text{ negative constant.}$$

Proceeding as in Sect. 2.1.1:

$$f_i(c_i | x_i, n_i) = \frac{c_i^{x_i-1} (1-c_i)^{n_i-x_i}}{Q(x_i, n_i) \ln c_i}, \quad [\text{Eq. 18}]$$

where, provided  $1 \leq x_i \leq n_i - 1$

$$Q(x_i, n_i) = \int_0^1 c_i^{x_i-1} \cdot (1-c_i)^{n_i-x_i} \frac{dc_i}{\ln c_i} = \sum_{j=0}^{n_i-x_i} (-1)^j \binom{n_i-x_i}{j} \cdot \ln(j+x_i) \quad [\text{Eq. 19}]$$

$$\bar{I}_i(s, 0) = \frac{1}{Q(x_i, n_i)} \int_0^1 c_i^{x_i-1} \cdot (1-c_i)^{n_i-x_i} \cdot (\ln c_i)^{s-1} \cdot dc_i$$

$$I_i(s, 0) = \frac{(-1)^{s-1} \cdot (s-1)!}{Q(x_i, n_i)} \sum_{j=0}^{n_i-x_i} (-1)^j \cdot \frac{\binom{n_i-x_i}{j}}{(j+x_i)^s}, \quad s \geq 1.$$

Again we are in a position to use the general method given previously.

During the determination of  $Q(x_i, n_i)$  the following restrictions on  $x_i$  were necessary

$$1 \leq x_i \leq n_i - 1.$$

It can be shown that

$$I_i(s, 0) \xrightarrow{x_i \rightarrow n_i} 0, \quad \text{all } s \geq 1.$$

Although from Eq. 18 and Eq. 19 we have

$$I_i(s, 0) = \lim_{x_i \rightarrow 0} \left\{ \frac{(-1)^{s-1} \cdot (s-1)!}{\ln(x_i) \cdot x_i^s} \right\},$$

this application is not defined for  $x_i = 0$  (no failures), since  $\sigma \rightarrow -\infty$  (See Appendix A).

It means that this application can be used only if there is at least one failure in each interval ( $x_i > 0$ ). The minimum values of  $x_i$  and  $n_i$  is  $(x_i, n_i) = (1, 1)$ . Intervals with no successes ( $x_i = n_i$ ) need not to be included in the calculation of the cumulants, since they do not contribute to Eq. B.2 in Appendix B.

The need to have a well-defined binomial process (constant  $n_i$ ) in each interval determines the intervals.

The major assumption in this section is the use of Jeffrey's extension of Bayes' postulate. Other distributions than  $dY/Y$  can be used if better arguments than Jeffrey's can be given.

An example is given in Appendix E.

### CONCLUSION

The purpose of this paper was to present a method for obtaining confidence intervals for a product of independent positive random variables, that does not assume a special type of distribution for the factors in the product. This is done, although the method is only approximative.

The problem of finding confidence limits has been reduced to the task of finding a family of definite integrals for each factor

in the product separately. The important thing is that the area of work has been transformed from the product to each factor.

The applications show that in some important cases the integrals can be found exactly.

In the application where the factors are binomial distributed the accuracy is acceptable even for very small sample sizes. All the applications given can be used down to sample sizes from zero to two depending on the application.

In general the accuracy of the method is no better than that of the asymptotic expansion used and is sometimes worse due to additional computation before applying the expansion.

The method cannot be upheld as an elegant mathematical technique. It was believed that elegance should be sacrificed in favour of the development of a usable general method.

#### REFERENCES

1. BACHE, N. On estimating the cumulative probability curve and its upper and lower confidence limits, SACLANTCEN SM-29. La Spezia, Italy, SACLANT ASW Research Centre, 1973. [AD 770246]
2. SPRINGER, H.D. & THOMPSON, W.E. Bayesian confidence limits for the product of N binomial parameters. *Biometrics* 53, 1966: 611-613.
3. PERNESKI, B.W., WAART, P.H. van der, & McDONALD, L.L. Selected statistical techniques applicable to ASW exercise design and analysis, Rpt No P-66-1-72. North Stonington, Conn., Analysis and Technology Inc., 1973.
4. MADANSKY, A. Approximate confidence intervals for the reliability of series and parallel systems. *Technometrics* 7, 1965: 495-503.
5. AITCHISON, J. & BROWN, J.A.C. *The Lognormal Distribution*. Cambridge, The University Press, 1969.
6. FISHER, R.A. *Contributions to Mathematical Statistics*. New York, Wiley, 1950: p.30.a.
7. KENDALL, M.C. & STUART, A. *The Advanced Theory of Statistics*, Vol. 1: *Distribution Theory*. London, Griffin & Co Ltd, 1958: p.163.
8. KENDALL, M.C. & STUART, A. *The Advanced Theory of Statistics*, Vol. 2: *Inference & Relationships*, 2nd ed. London, Griffin & Co Ltd, 1967: p.150.
9. ABRAMOWITZ, M. & STEGUN, I.A. (eds.). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. Washington, D.C., National Bureau of Standards, 1964: p.935.

A P P E N D I C E S

APPENDIX A

ASYMPTOTIC EXPANSION FOR THE INVERSE FUNCTION  
OF AN ARBITRARY DISTRIBUTION FUNCTION

For a detailed discussion, see Refs. A.1, A.2. A briefer description is given in Ref. A.3, as follows:

Let the distribution function of a stochastic variable  $Y$  be denoted by  $F(y)$  and its cumulants by  $K_r$ . Then the (Cornish-Fisher) asymptotic expansion for the value of  $y_p$  such that  $F(y_p) = 1 - p$  is ( $\sim$  means asymptotically equal)

$$y_p \sim m + \sigma \cdot w$$

where  $m = K_1$ ,  $\sigma = \sqrt{K_2}$  and

$$\begin{aligned} w = & x + \gamma_1 h_1(x) \\ & + \gamma_2 h_2(x) + \gamma_1^2 h_{11}(x) \\ & + \gamma_3 h_3(x) + \gamma_1 \gamma_2 h_{12}(x) + \gamma_1^3 h_{111}(x) \\ & + \gamma_4 h_4(x) + \gamma_2^2 h_{22}(x) + \gamma_1 \gamma_3 h_{13}(x) + \gamma_1^2 \gamma_2 h_{112}(x) + \gamma_1^4 h_{1111}(x) \\ & + \dots \end{aligned} \quad [\text{Eq. A.1}]$$

where  $\gamma$  is defined as  $\gamma_{r-2} = \frac{K_r}{\sigma^r}$ ,  $r \geq 3$

and  $x$  determined by

$$\frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left\{-\frac{t^2}{2}\right\} dt = p$$

$x$  is the fractile of the normalized normal distribution. The terms on each line in Eq. A.1 are the same order of magnitude.

$$\begin{aligned}
h_1(x) &= \frac{1}{6} \text{He}_2(x) \\
h_2(x) &= \frac{1}{24} \text{He}_3(x) \\
h_{11}(x) &= -\frac{1}{36} [2\text{He}_3(x) + \text{He}_1(x)] \\
h_3(x) &= \frac{1}{120} \text{He}_4(x) \\
h_{12}(x) &= -\frac{1}{24} [\text{He}_4(x) + \text{He}_2(x)] \\
h_{111}(x) &= \frac{1}{324} [12\text{He}_4(x) + 19\text{He}_2(x)] \\
h_4(x) &= \frac{1}{720} \text{He}_5(x) \\
h_{22}(x) &= -\frac{1}{384} [3\text{He}_5(x) + 6\text{He}_3(x) + 2\text{He}_1(x)] \\
h_{13}(x) &= -\frac{1}{180} [2\text{He}_5(x) + 3\text{He}_3(x)] \\
h_{112}(x) &= \frac{1}{288} [14\text{He}_5(x) + 37\text{He}_3(x) + 8\text{He}_1(x)] \\
h_{1111}(x) &= -\frac{1}{7776} [252\text{He}_5(x) + 832\text{He}_3(x) + 227\text{He}_1(x)].
\end{aligned}$$

The  $\text{He}_n(x)$  are the Hermite polynomials, which can be calculated recursively from

$$\text{He}_0(x) = 1, \quad \text{He}_1(x) = x$$

and

$$\text{He}_{n+1}(x) = x \cdot \text{He}_n(x) - n \cdot \text{He}_{n-1}(x), \quad n \geq 1.$$

#### REFERENCES

- A.1 FISHER, R.A. Contributions to Mathematical Statistics. New York, Wiley, 1950: p.30.a.
- A.2 KENDALL, M.C. & STUART, A. The Advanced Theory of Statistics, Vol. 1: Distribution Theory. London, Griffin & Co Ltd, 1958: p.163.
- A.3 ABRAMOWITZ, M. & STEGUN, I.A..(eds.). Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables. Washington, D.C., National Bureau of Standards, 1964: p.935.



## APPENDIX B

### THE FIRST SIX CUMULANTS

We will find expressions for the first six cumulants by means of the characteristic function  $\varphi_Y(t)$  of  $Y$ .

Using Eqs. 4 and 6 of the main text and the assumption that the  $Y_i$  are independent leads to

$$\varphi_Y(t) = \prod_{i=1}^N \varphi_{Y_i}(t),$$

where  $\varphi_{Y_i}(t)$  is the characteristic function for the  $Y_i$ .

Using Eq. 7 of the main text for the function in Eq. 3 of the main text gives

$$\varphi_{Y_i}(t) = \int \exp\{it \ln c_i\} f_i(c_i) dc_i,$$

with limits of integration from zero to infinity; thus

$$\varphi_{Y_i}(t) = \int c_i^{it} f_i(c_i) dc_i. \quad [\text{Eq. B.1}]$$

From Eq. 8 of the main text we have

$$\ln \varphi_Y(t) = \sum_{r=1}^{\infty} K_r \frac{(it)^r}{r!}.$$

Denoting the cumulants of the distribution of  $Y_i$  by  $K_{i,r}$  we have again from Eq. 8 of the main text

$$\ln \varphi_{Y_i}(t) = \sum_{r=1}^{\infty} K_{i,r} \frac{(it)^r}{r!},$$

leading to

$$K_r = \sum_{i=1}^N K_{i,r}. \quad [\text{Eq. B.2}]$$

$K_{i,r}$  can be determined by

$$K_{i,r} = \left[ \frac{d^r}{d(it)^r} \varrho_n \varphi_{Y_i}(t) \right]_{t=0} . \quad [\text{Eq. B.3}]$$

Substituting Eq. B.1 in Eq. B.3 gives

$$K_{i,r} = \left[ \frac{d^r}{d(it)^r} \left\{ \varrho_n \int c_i^{it} f_i(c_i) dc_i \right\} \right]_{t=0} . \quad [\text{Eq. B.4}]$$

Define a function  $I_i(s, t)$  as

$$I_i(s, t) = \int (\varrho_n c_i)^s c_i^{it} f_i(c_i) dc_i ,$$

then we have the property of  $I_i(s, t)$ :

$$\frac{d}{d(it)} I_i(s, t) = I_i(s+1, t) .$$

Equation B.4 can then be written as

$$K_{i,r} = \left[ \frac{d^r}{d(it)^r} \varrho_n I_i(0, t) \right]_{t=0} . \quad [\text{Eq. B.5}]$$

From Eq. B.5 we find

$$K_{i,1} = I_i(1, 0) .$$

$$K_{i,2} = I_i(2, 0) - K_{i,1}^2 .$$

$$K_{i,3} = I_i(3, 0) - 3K_{i,1} \cdot K_{i,2} - K_{i,1}^3 .$$

$$K_{i,4} = I_i(4, 0) - 4K_{i,1} \cdot K_{i,3} - 6K_{i,1}^2 \cdot K_{i,2} - 3K_{i,2}^2 - K_{i,1}^4 .$$

$$K_{i,5} = I_i(5, 0) - 5K_{i,1} \cdot K_{i,4} - 10K_{i,2} \cdot K_{i,3} - 10K_{i,1}^2 \cdot K_{i,3} - 15K_{i,1} \cdot K_{i,2}^2 - 10K_{i,1}^3 \cdot K_{i,2} - K_{i,1}^5 .$$

$$K_{i,6} = I_i(6, 0) - 6K_{i,1} \cdot K_{i,5} - 15K_{i,2} \cdot K_{i,4} - 15K_{i,1}^2 \cdot K_{i,4} - 60K_{i,1} \cdot K_{i,2} \cdot K_{i,3} - 20K_{i,1}^3 \cdot K_{i,3} - 45K_{i,1}^2 \cdot K_{i,2}^2 - 15K_{i,1}^4 \cdot K_{i,2} - 10K_{i,2}^3 - 15K_{i,2}^2 \cdot K_{i,1} - K_{i,1}^6 .$$

$K_r (r = 1, \dots, 6)$  can now be determined, provided  $I_i(s, 0)$  can be evaluated:

$$I_i(s, 0) = \int (\ln c_i)^s \cdot f_i(c_i) \cdot dc_i \quad [\text{Eq. B.6}]$$

APPENDIX C

EVALUATING A DEFINITE INTEGRAL

Dropping the index  $i$  in Eq. 14 of the main text we have:

$$I(s, 0) = (n+1) \cdot \binom{n}{x} \cdot \int_0^1 (\ln c)^s \cdot c^x \cdot (1-c)^{n-x} \cdot dc \quad [\text{Eq. C.1}]$$

$(1-c)^{n-x}$  can be expanded in a binomial series:

$$(1-c)^{n-x} = \sum_{k=0}^{n-x} \binom{n-x}{k} (-c)^k. \quad [\text{Eq. C.2}]$$

Using Eq. C.2 in Eq. C.1:

$$I(s, 0) = (n+1) \cdot \binom{n}{x} \cdot \sum_{k=0}^{n-x} (-1)^k \cdot \binom{n-x}{k} \cdot \int_0^1 (\ln c)^s \cdot c^{x+k} \cdot dc. \quad [\text{Eq. C.3}]$$

From standard integral tables:

$$\int (\ln c)^s \cdot c^{x+k} \cdot dc = s! \cdot \frac{c^{x+k+1}}{x+k+1} \cdot \sum_{j=0}^s \frac{(-1)^j \cdot (\ln c)^{s-j}}{(s-j)! \cdot (x+k+1)^j}.$$

Therefore

$$\int_{\epsilon}^1 (\ln c)^s \cdot c^{x+k} \cdot dc \xrightarrow{\epsilon \rightarrow 0^+} \frac{s!}{x+k+1} \cdot \frac{(-1)^s}{(x+k+1)^s}. \quad [\text{Eq. C.4}]$$

Substituting Eq. C.4 into Eq. C.3 gives Eq. 15 of the main text.  
As a partial check we must have  $I(0, 0) = 1$ .

APPENDIX D

LOCATION OF THE MAXIMUM ERROR

From Eq. B.2 in Appendix B we see that  $K_r$  is of order  $N$ ; we may write

$$K_r \sim O(N) .$$

Then

$$\gamma_{r-2} = \frac{K_r}{K_2^{r/2}} \sim O(N^{1-r/2}), \quad r \geq 3 \quad [\text{Eq. D.1}]$$

(see App.A)

The  $\gamma$ 's are a measure of the deviation of  $h(y)$  (see main text) from the normal distribution. Because of Eq. D.1 the absolute error will then decrease monotonically as  $N$  increases.

Therefore a maximum, apart from  $(n_i, N) = (0, 1)$ , must lie in the column  $N=1$  of Table 1 of the main text, if a maximum exists. For  $N=1$  the density function  $h(y)$  for  $Y = \ln C$  is (dropping the index  $i$ ):

$$h(y) = (n+1) \cdot \binom{n}{x} \cdot e^{(x+1) \cdot y} \cdot (1 - e^y)^{n-x} .$$

This function has a maximum, except for  $x = n$ , which is the case that differs most from the corresponding normal density function with the same mean and variance. This does not change with increasing  $n$ . Therefore the maximum error will always come from this family of curves ( $x = n$ ).

For these curves we have

$$I(s, 0) = \frac{(-1)^s \cdot s!}{(n+1)^s}$$

giving

$$K_r = \frac{A}{(n+1)^r} \quad A \text{ is a constant}$$

$$\gamma_{r-2} = \frac{K_r}{K_2^{r/2}} = A^{1-r/2}, \quad r \geq 3 \quad (\text{see App.A}) \quad [\text{Eq. D.2}]$$

Hence the  $\gamma$ 's are constants, independent of  $n$ . The argument  $y_p$  in Appendix A can then be written

$$y_p \sim \frac{-A}{n+1} \quad (A \text{ is a new constant } > 0).$$

So the argument for  $C$  will then be

$$\underline{c_p} \sim e^{-\frac{A}{n+1}} \quad (\text{see Eq. 5 of main text}).$$

A direct evaluation of  $c_p$  is

$$g(c) = (n+1)c^n$$

$$\int_0^{c_p} (n+1) c^n dc = 1-p \Rightarrow c_p^{n+1} = 1-p$$

$$\underline{c_p} = e^{\frac{\ln(1-p)}{n+1}}.$$

The difference

$$\left| e^{-\frac{A}{n+1}} - e^{\frac{\ln(1-p)}{n+1}} \right|,$$

which is the error given in column  $N=1$  in Table 1 of the main text, has only one maximum at  $n=0$  and tends to zero as  $n \rightarrow \infty$ .

Looking at Table 1 in the main text, the error is seen to fall off more slowly vertically than horizontally. This is because the  $\gamma$ 's (which are related to the error) remains constant (D.2) for increasing  $n$  ( $N$  fixed) whereas the  $\gamma$ 's decreases (D.1) towards zero for increasing  $N$  ( $n$  fixed).

APPENDIX E

TWO NUMERICAL EXAMPLES

E.1 Product of Probabilities [Referring to Sect. 2.1.1 of main text]

Maritime patrol aircrafts (MPA) detected submarines ten times in certain ASW exercises. On these ten occasions, the submarine was localized eight times. Another sample shows that out of nine occasions when submarines were localized, seven resulted in an attack on the submarine by the aircraft. A third sample shows that out of four occasions when the submarine was attacked three were evaluated as successful. The above data are hypothetical and, for sake of comparison, have been made equal to the data used in the example in Ref. E.1.

The question is: What is the confidence limits for the true probability that the submarine will be killed, given that the aircraft has detected it?

The result is:

TABLE E.1  
AN EXAMPLE :  
PRODUCT OF PROBABILITIES

	THIS METHOD	EXACT VALUE [Ref.E.1]	ERROR IN %
UPPER 10%	0.54270669	0.54224843	+ 0.085
MEDIAN	0.35647715	0.35666951	- 0.054
LOWER 10%	0.19459118	0.19460653	- 0.0079

The error consists of both computational error and error due to the approximation.

E.2 Cumulative Probability Curve [Referring to Sect. 2.1.2 of main text]

Sample tracks have been collected of a submarine and a sonar-fitted surface ship trying to make a sonar detection of the submarine (Fig. E.1). A dot means a detection. From this sample

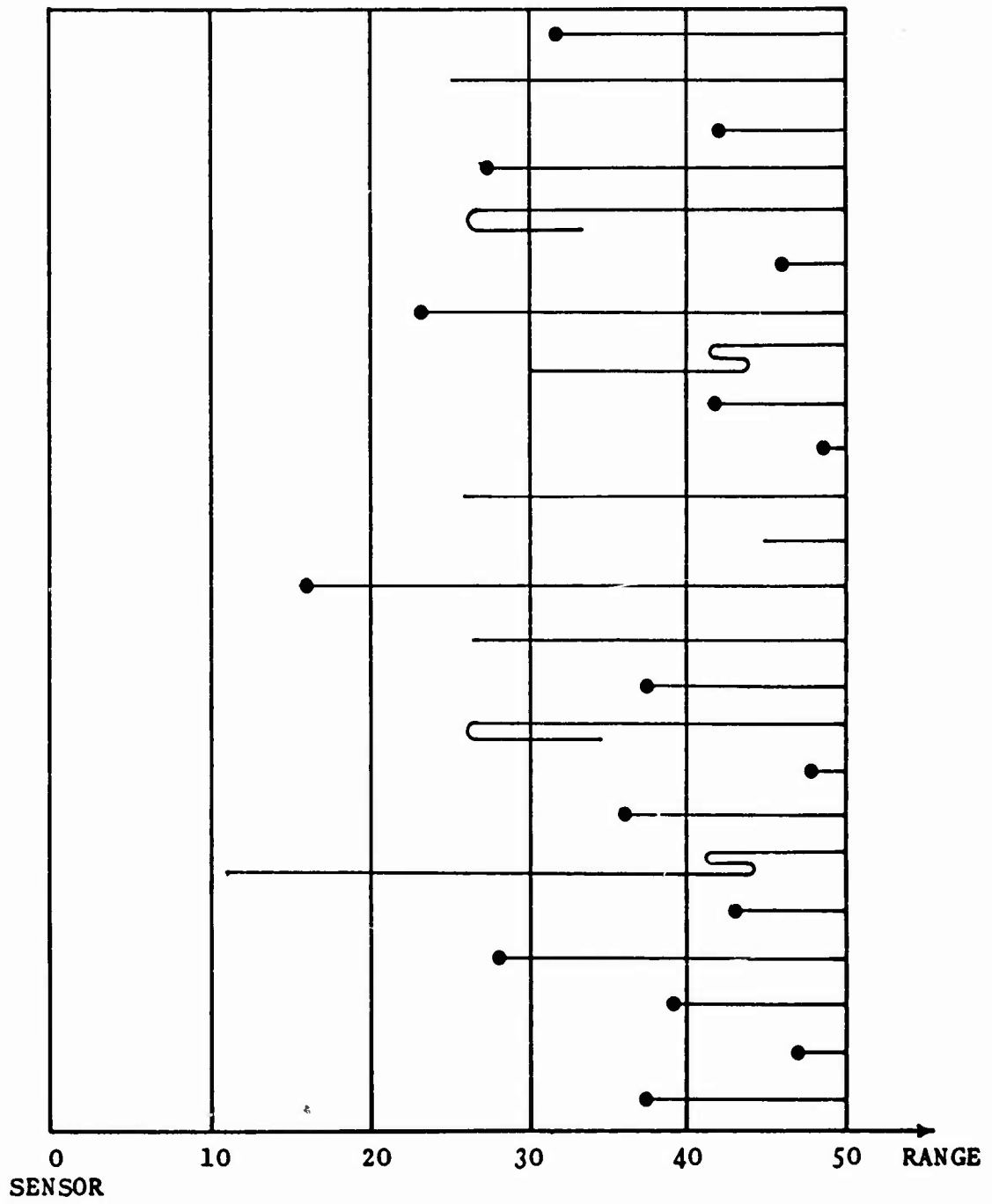


FIG. E.1 A SAMPLE OF DETECTIONS AND OPPORTUNITIES



we want confidence limits for the true probability of the ship detecting the submarine outside a given range. This example could as well have treated radar, visual or ECM detections or other kinds of detections. The range is divided into brackets containing a constant number of opportunities  $n_i$ .

Figure E.2 gives a general picture of each interval.

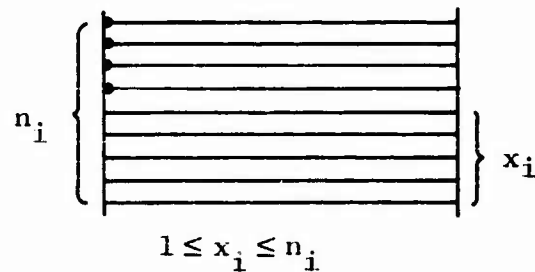


FIG E 2 INTERVAL

Table E.2 is produced from Fig. E.1, leaving out intervals with no detections.

TABLE E.2

AN EXAMPLE :  
CUMULATIVE PROBABILITY CURVE

DETECTION(s) AT RANGE	$n_i$	$x_i$
16.0	2	1
23.0	3	2
27.0	11	10
28.0	12	11
31.5	14	13
36.0	13	12
37.3	15	13
39.0	16	15
41.4	21	20
41.7	22	21
42.8	23	22
45.5	21	20
46.8	22	21
47.5	23	22
48.2	24	23

By using the method explained in Sect. 2.1.2 of the main text on the data in Table E.2 we obtain Fig. E.3. Figure E.3 also marks with crosses the confidence intervals corresponding to a division of the range into brackets of 10 units starting with range 0. Without relation to the sample such a division is arbitrary and creates difficulties in choosing an appropriate value for the number of opportunities.

#### REFERENCE

- E.1 SPRINGER, H.D. & THOMPSON, W.E. Bayesian confidence limits for the product of  $N$  binomial parameters. *Biometrics* 53, 1966: 611-613.

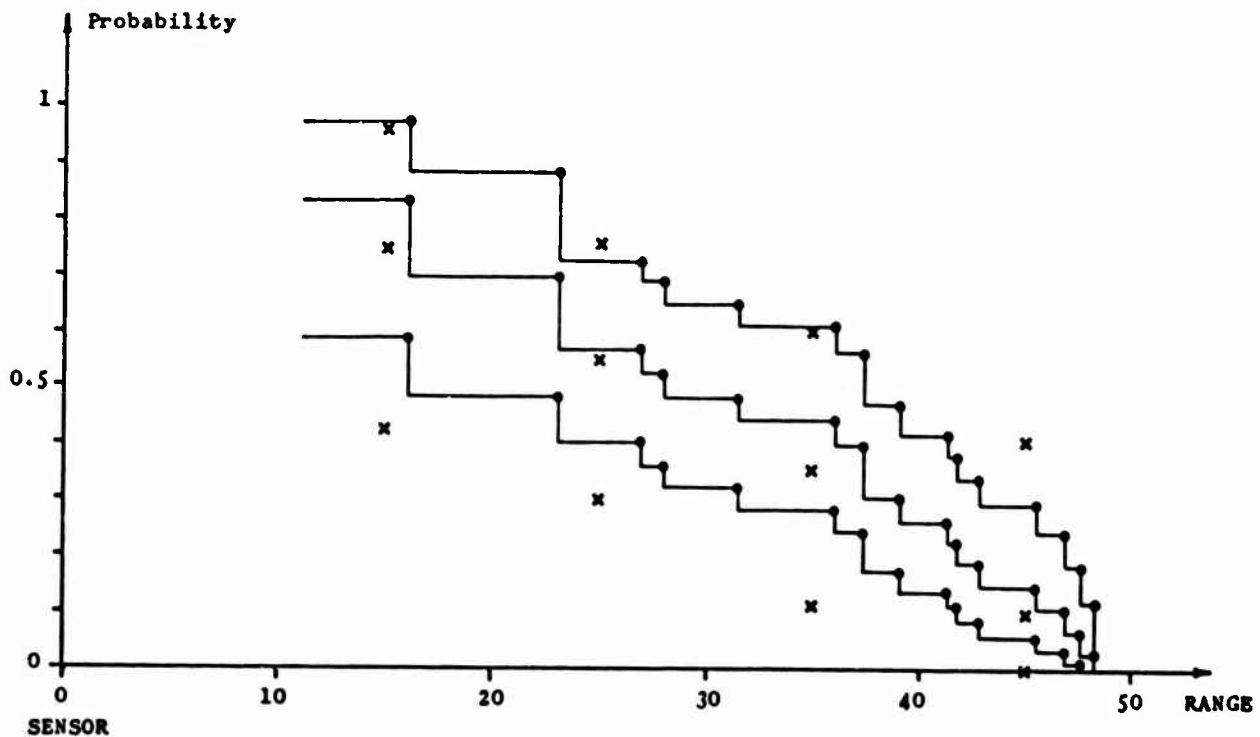


FIG. E.3 CONFIDENCE LIMITS

## APPENDIX F

### CONFIDENCE LIMITS OF A FUTURE OBSERVED VALUE OF PRODUCTS OF PROBABILITIES, BASED ON A SAMPLE ALREADY TAKEN

#### F.1 Product of Probabilities

Usually an estimate for the value of product  $C$  is wanted. Instead of choosing a suitable estimator (mean, median, mode etc.) based on the distribution of  $C$  a much simpler (but not as satisfactory) procedure is often followed, namely, that of choosing a suitable estimate for each of the factors  $C_i$  in the product  $C$  and then forming the product. If  $R_i$  successes are observed out of  $m_i$  opportunities for each factor  $i$ , the maximum likelihood estimator for the true probability of success  $c_i$  is  $R_i/m_i$ .

Let

$$C_i = R_i/m_i, \quad 1 \leq R_i \leq m_i$$

$R_i$  is a random variable

$m_i$  is a non-random variable.

We are interested in finding confidence limits for the product

$$\prod_{i=1}^N R_i/m_i$$

given that we have observed  $x_i$  successes out of  $n_i$  opportunities for each factor  $i$  (as in Sect. 2.1.1 of the main text).

The  $m_i$ 's could be a future sample to be collected and  $(x_i, n_i)$  could be a sample already obtained. If  $z_i$  successes are observed out of  $m_i$  outcomes in the future sample, then the product

$$z = \prod_{i=1}^N z_i/m_i$$

lies inside the confidence limits for the product

$$\prod_{i=1}^N R_i/m_i$$

in a certain fraction of cases (dependent on the confidence level) under the hypothesis that the underlying process (generating successes and failures) is the same for the sample obtained  $(x_i, n_i)$  and for the future sample  $(z_i, m_i)$ .

Therefore if the product  $z$  lies outside the confidence limits it is an indication that the underlying process has changed between the sample  $(x_i, n_i)$  and  $(z_i, m_i)$ .

The inequality  $1 \leq R_i \leq m_i$  means that we assume continuing the experiments until we have at least one success ( $R_i \geq 1$ ) for each factor  $i$ . ( $R_i/m_i$  is then no more the maximum likelihood estimator). The reason is to overcome the singularity at  $R_i=0$ . This assumption is not necessary for the sample  $(x_i, n_i)$ . Our knowledge about the true probability of success,  $c_i$ , is in the form of the distribution for  $c_i$  (the beta distribution)

$$(n_i + 1) \cdot \binom{n_i}{x_i} \cdot c_i^{x_i} \cdot (1 - c_i)^{n_i - x_i} \quad [\text{Eq. F.1}]$$

The probability that  $R_i$  will take the value  $\alpha$  is

$$\binom{m_i - 1}{\alpha - 1} \cdot c_i^{\alpha - 1} \cdot (1 - c_i)^{m_i - \alpha}$$

Therefore

$$I_i(s, 0) = \int_0^1 \sum_{\alpha=1}^{m_i} \left( \ln \frac{\alpha}{m_i} \right)^s \cdot \binom{m_i - 1}{\alpha - 1} \cdot c_i^{\alpha - 1} \cdot (1 - c_i)^{m_i - \alpha} \cdot (n_i + 1) \cdot \binom{n_i}{x_i} \cdot c_i^{x_i} \cdot (1 - c_i)^{n_i - x_i} \cdot dc_i$$

$$I_i(s, 0) = \frac{n_i + 1}{m_i + n_i} \cdot \binom{n_i}{x_i} \cdot \sum_{\alpha=1}^{m_i - 1} \frac{\binom{m_i - 1}{\alpha - 1}}{\binom{m_i + n_i - 1}{\alpha + x_i - 1}} \cdot \left( \ln \frac{\alpha}{m_i} \right)^s$$

Of interest is the case  $m_i = n_i$ , which is implemented in the computer program.

The case  $m_i = n_i$  means that a future sample is to be collected with exactly the same sample sizes as the sample already obtained. Confidence limits for the product

$$\prod_{i=1}^N R_i / n_i \quad [\text{Eq. F.2}]$$

can then be used as a measure of uncertainty for the product

$$\prod_{i=1}^N x_i / n_i$$

already obtained.

Therefore confidence limits for Eq. F.2 is, in the following, called "limits for an observed value of the product".

### EXAMPLE

The same example is used as in Appendix E. The upper and lower 10% limits for the observed product

$$\frac{8}{10} \times \frac{7}{9} \times \frac{3}{4}$$

is given in Fig. F.1, together with confidence limits for the true probability computed in Appendix E.

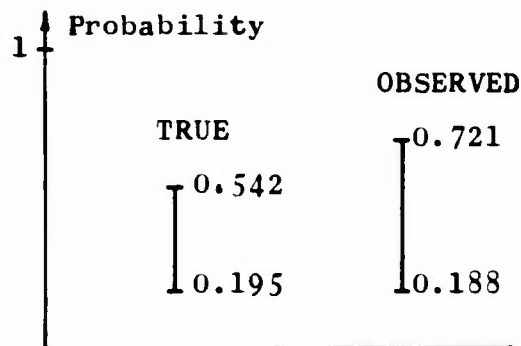


FIG. F.1 LIMITS FOR AN OBSERVED VALUE

### F.2 Cumulative Probability Curve

Calling  $R_i$  number of failures ( $R_i \geq 1$ ) and replacing Eq. F.1 by Eq. 18 of the main text

$$I_i(s, 0) = \int_0^1 \sum_{x=1}^{m_i} \left( \ln \frac{x}{m_i} \right)^s \binom{m_i-1}{x-1} \cdot c_i^{x-1} \cdot (1-c_i)^{m_i-x} \cdot \frac{c_i^{x_i-1} \cdot (1-c_i)^{n_i-x_i}}{Q(x_i, n_i) \cdot \ln c_i} \cdot dc_i,$$

or

$$I_i(s, 0) = \frac{1}{Q(x_i, n_i)} \cdot \sum_{x=1}^{m_i-1} \left( \ln \frac{x}{m_i} \right)^s \cdot \binom{m_i-1}{x-1} \cdot Q(x+x_i-1, m_i+n_i-1).$$

Again the case  $m_i = n_i$  is of interest.

Difficulties have been encountered in computing  $Q(x+x_i-1, m_i+n_i-1)$  accurately for larger  $N$  ( $N > 20$ ), therefore another expression for  $Q$  giving more accuracy than Eq. 19 of the main text has been used in the computer program. A listing of the computer program and test examples are available from SACLANTCEN.

INITIAL DISTRIBUTION

	Copies		Copies
<u>MINISTRIES OF DEFENCE</u>		<u>SCNR FOR SACLANTCEN</u>	
MOD Belgium	2	SCNR Belgium	1
DND Canada	10	SCNR Canada	1
CHOD Denmark	8	SCNR Denmark	1
MOD France	8	SCNR Germany	1
MOD Germany	15	SCNR Greece	1
MOD Greece	11	SCNR Italy	1
MOD Italy	10	SCNR Netherlands	1
MOD Netherlands	12	SCNR Norway	1
CHOD Norway	10	SCNR Portugal	1
MOD Portugal	5	SCNR Turkey	1
MOD Turkey	5	SCNR U.K.	1
MOD U.K.	16	SCNR U.S.	2
SECDEF U.S.	60		
 <u>NATO AUTHORITIES</u>		 <u>NATIONAL LIAISON OFFICERS</u>	
Defence Planning Committee	3	NLC Denmark	1
NAMILCOM	2	NLO Italy	1
SACLANT	10	NLO U.K.	1
SACLANTREPEUR	1	NLO U.S.	1
CINCWESTLANT/COMOCEANLANT	1		
COMIBERLANT	1	<u>NLR TO SACLANT</u>	
CINCEASTLANT	1	NLR Belgium	1
COMSUBACLANT	1	NLR Canada	1
COMCANLANT	1	NLR Germany	1
COMMAIREASTLANT	1	NLR Greece	1
SACEUR	2	NLR Italy	1
CINCNORTH	1	NLR Norway	1
CINC SOUTH	1	NLR Portugal	1
COMNAVSOUTH	1	NLR Turkey	1
COMSTRIKFORSOUTH	1		
COMEDCENT	1	ESRO/ELDO Doc. Service	1
COMSUBMED	1		
COMMARARMED	1		
COMTWOATAF	1		
CINCHAN	1		
		Total initial distribution	231
		SACLANTCEN Library	10
		Stock	<u>39</u>
		Total Number of Copies	280