

APPROXIMATIVE EVALUATION OF THE DISTRIBUTION FUNCTION OF AGGREGATE CLAIMS¹

BY T. PENTIKAINEN

Helsinki

ABSTRACT

A formula, originally presented by HALDANE (1938)², for the evaluation of the distribution of aggregate claims is examined and compared with some other approaches. The idea is to apply a symmetrizing transformation to the original variable in order to make it susceptible to be approximated by the normal distribution.

KEYWORDS

Aggregate claim distribution; approximate evaluation; NP method; Haldane approximation; Wilson–Hilferty approximation

1. INTRODUCTION

1.1. A problem frequently faced in application-orientated risk theory is the numerical evaluation of the distribution function $F(X)$ of the aggregate amount X of claims. It is conventionally expressed by the formula

$$(1.1) \quad F(X) = \sum_{k=0}^{\infty} p_k S^{k*}(X),$$

where p_k is the probability that the number of claims is equal to k , S is the distribution function of the individual claim sizes and S^{k*} its k th convolution (see BPP, p. 51; this and similar quotations in the sequel refer to the Risk Theory book by BEARD, PENTIKAINEN and PESONEN, 1984).

The claim numbers are often assumed to follow a simple Poisson distribution (in order to define notations some well-known basic formulae are recapitulated)

$$(1.2) \quad p_k(n) = e^{-n} n^k / k!$$

where n , the Poisson parameter, is the expected number of claims. A more general approach is to adopt the mixed Poisson distribution (BPP, p.33):

$$(1.3) \quad p_k(n) = \int_0^{\infty} e^{-nq} \frac{(nq)^k}{k!} dH(q).$$

¹ Presented originally at the Risk Theory Seminar in Oberwolfach 1984 and in an extended form at the Risk Theory Seminar of the American Risk and Insurance Association in Nashville 1985

² J. B. S. HALDANE (1892–1964) first studied mathematics and later became Professor of Biometrics at University College, London, before moving to India in 1957

The distribution function H introduces the so-called structure variation of the claim number probabilities into the model, indicating the time variation of the risk exposure, or the heterogeneity of the risks inside the collective, or both. A popular H is the gamma function resulting in the negative binomial distribution (BPP, p. 40):

$$(1.4) \quad p_k(n) = \binom{h+k-1}{k} \left(\frac{h}{n+h}\right)^h \left(\frac{n}{n+h}\right)^k$$

where h is a shape parameter. This alternative is usually called the *Polya case* in distinction from the *Poisson case* (1.2).

Note that for the approximation formulae only the lowest moments of H are necessary. Therefore, it is sufficient merely to estimate (or assume) them, not bothering about the analytic formulation of this function.

1.2. Need to have approximation methods. The construction (1.1) is unfortunately so intricate that the direct computation of F is tractable only in special cases, even though the recently developed recursive methods (ADELSON, PANJER, GERBER, JEWELL & SUNDT, description of the method and references see BPP, Section 3.8) as well as the Fourier transformation technique (BERTRAM, 1981) have made major progress in solving this problem. Nevertheless, they have not removed the need also to have rapid and reasonably comfortable, even though approximate, approaches. This is due to the fact that the number of the computation steps needed for the recursive calculation grows quite massive in cases where the risk portfolio is large (as most insurer's portfolios are) and/or when the claim size distribution has a long tail. This can be a major handicap, in particular in a sophisticated problem complex such as the analysis of long-term processes, simulations, etc. where the computation of F is needed frequently, say 1000, 10000 or 100000 times for one single procedure. Then the problem is, above all, to minimize the computation time in terms of milliseconds (rather than in seconds!) to make the usual present day personal computers operational.

Approximate methods can also have the merit of providing an analytic, often perspicuous, relationship between the main variables controlling the processes.

2. SOME EARLIER APPROACHES

2.1. Normal approximation. A classic approach, based on the central limit theorem, is to approximate $F(X)$ by the normal distribution:

$$(2.1) \quad F(X) \approx N(x)$$

where, denoting the mean and the standard deviation of X by m_X and σ_X

$$(2.2) \quad x = (X - m_X) / \sigma_X.$$

This expression is asymptotically correct in the Poisson case but not generally, e.g. not in the Polya case. Its major weakness is that it may crudely underestimate the risk of large aggregate claims (see BPP, p.105). This is due to the fact that N as a symmetric function cannot successfully approximate any distribution which is notably skew.

2.2. Symmetrization. A way of overcoming the weakness of the normal approximation is to transform the original variable X into an auxiliary variable y_{norm} by using a suitably chosen function v

$$(2.3) \quad y_{\text{norm}} = v(X)$$

so that it makes the distribution (at least approximately) symmetric. Then providing that y_{norm} is standardized to have a zero mean and standard deviation unity, one can expect that it can be satisfactorily approximated by the normal distribution:

$$(2.4) \quad F(X) = \tilde{F}(y_{\text{norm}}) = N(y_{\text{norm}}).$$

Depending on the choice of the transformation v a family of approximation methods is constituted including those dealt with in this paper. An analysis of some of these transformations can be found for example in BOX and COX (1964).

2.3. NP approximation is obtained assuming $v^{-1}(y_{\text{norm}})$ as a polynomial:

$$(2.5) \quad x = (X - m_X)/\sigma_X = y_{\text{norm}} + \gamma(y_{\text{norm}}^2 - 1)/6$$

where $\gamma = \gamma_X$ is the skewness of the original distribution (BPP, Chap. 3.11).

The transformation (2.5) is applicable only for the long tail $X > m_X$ of the distribution and therefore needs a modifying extension (BPP, Chap. 3.11) resulting in a three-piece formula.

2.4. Other methods. There are a number of approaches based on the principle of replacing the original distribution by some suitable approximating function, which is conveniently computable. Most of them are obtained by equating the lowest moments, as is also the case in the above items 2.2 and 2.3. For instance the three-parameter gamma function $\Gamma(ax + b, c)$ (BOHMAN and ESSCHER, 1964) or the Pearson functions (LAU, 1984) are suggested. OSCHWALD (1984) has recently presented an analogous transformation to (2.5) using the gamma function instead of the normal function.

Unfortunately the range of applicability of most of these methods has been examined only by means of very few (and often "easy") examples, as yet. So far as is known, those methods which meet the demand of reasonable convenience and the requirement for computation speed, do not have the accuracy or the other merits which would not prevent the approaches to be dealt with in the sequel to be competitive. Further studies would be desirable, but are, however, beyond the scope of this work.

One of the known approximations is still worthy of special mention. ESSCHER (1932) introduced a method which makes use of the whole range of the claim size distribution, not only of some of its moments. BOHMAN and ESSCHER (1964) gave a number of tests, which proved that the merits of this method may not be very superior to those of the gamma approximation. However, a recent (unpublished) work of PUSA (1985) seems to indicate a good fit also in some cases where the other methods fail. A drawback of the Esscher approach is that it

employs an auxiliary variable h and the relevant quantities F and X are available only as functions of this variable. To get the matching F and X it is necessary first to find the corresponding h , which seems to need an iteration process impairing the speed of the method.

2.5. Moment problem. One should keep in mind the fact that most of the methods referred to above give, when applied to the mixed compound Poisson function, the same approximating function to all those original distributions, which have the same three (or four) lowest moments determined by the mean, standard deviation, skewness (and the kurtosis) of the claim size distribution. However, these moments do not fully determine the flow a function, hence there is “a funnel of doubt” [$F_1(x)$, $F_u(x)$] inside which the values of the original distribution functions $F(x)$ are positioned for varying x values. F_1 and F_u are the upper and the lower envelope curve, respectively, of all those original distribution functions which fulfil the specifications of the problem setting concerned. If the funnel is large, then there are always cases which markedly deviate from the approximating values, whichever of the methods is used.

GOOVAERTS & KAAS (1986) have recently provided a meritorious method of evaluating the range of the variation subject to the condition that the lowest moments are fixed and the variable x is limited to some interval, e.g. $[0, b]$. Some examples showed that the range in which the permitted $F(x)$ values may be positioned is rather wide, in fact reducing the prospects of finding suitable approximations based on the moments. Fortunately, this result does not wreck the prospects of finding reasonably useful approximations, if the basic condition is taken of fixing a sequence of the moments of *the claim size function* S (not of the aggregate distribution as Goovaerts and Kaas propose) and of limiting the claim size Z (not the aggregate X) to some finite interval. In fact, this is the proper problem setting for the NP method as well as for the methods to be discussed in the sequel. We will return to the moment problem in Section 6.9.

3. HALDANE APPROXIMATION

3.1. The idea. The approach we are going to deal with was originally presented by WILSON and HILFERTY (1931) for an approximate evaluation of the gamma function. HALDANE (1938) extended it to the function classes which have suitably convergent sequences of cumulants. In what follows we apply the method to the mixed compound Poisson function specified in Section 1.1, even though the most part of the derivation is valid more generally.

The method makes use of the symmetrization as described in Section 2.2 above. Haldane first adopted a power expression

$$y = (X/m_x)^h$$

where h is an auxiliary parameter. Then (truncated) expansions are derived for the mean m_y , standard deviation σ_y and skewness γ_y of y . The symmetrization is achieved by assigning for the auxiliary parameter h a value which equates the

skewness γ_y with zero. Putting

$$(3.1) \quad y_{\text{norm}} = (y - m_y) / \sigma_y = [(X/m_x)^h - m_y] / \sigma_y$$

the transformation aimed at, corresponding to v in (2.3), is obtained.

The derivation of the Haldane formulae is notably laborious, even though rather elementary methods only are needed. Therefore, we shall not give more than some intermediate expressions in the chain of treatments in order to provide a conception of how the results are found, the more so because the Haldane derivations do not result in any strictly rigorous estimates for the accuracy of the approximations nor for clear rules of their applicability, but rather only justify the expectation that in a certain environment the procedure may lead to acceptable outcomes. In fact, the discussion about the appropriateness of the approach is mainly based on tests where a number of distributions are calculated exactly and in parallel, by using the approximations that will be presented in the subsequent sections. Readers who are mainly interested in the practical results may well skip over to Section 3.4, at least at the first reading.

3.2. Derivation of the formula. We aim to preserve, as far as possible, the original procedures and notations of Haldane even though some of the results, e.g. the value of the parameter h , could be obtained by more straightforward ways.

The technique to be assumed operates partially the so-called cumulants, which are, as is well known from textbooks on statistics (e.g. KENDALL and STUART, 1979, Section 3.12), the coefficients in the expansion of the cumulant-generating function $\log \varphi(it)$ in the terms of (it) : $\kappa_1(it) + \kappa_2(it)^2/2! + \dots$, where φ is the characteristic function and i the imaginary unit. For the convenience of the reader, we recall the connecting equations between the lowest cumulants and the more commonly applied central moments, denoted by $\mu_i (i = 1, 2, \dots)$:

$$\kappa_1 = m_x, \kappa_2 = \mu_2, \kappa_3 = \mu_3, \kappa_4 = \mu_4 - 3\mu_2^2.$$

Another auxiliary variable is introduced:

$$(3.2) \quad x' = X - \kappa_1$$

and substituted into (3.1) after which expansions for the moments of y about zero, denoted by β_r , are obtained as follows

$$(3.3) \quad \beta_r = Ey^r = E\left\{\left[\left(1 + \frac{x'}{\kappa_1}\right)^h\right]^r\right\} = E\left\{\left(1 + \frac{x'}{\kappa_1}\right)^{rh}\right\}$$

$$= 1 + f(r, 2) \frac{\mu_2}{2! \kappa_1^2} + f(r, 3) \frac{\mu_3}{3! \kappa_1^3} + f(r, 4) \frac{\mu_4}{4! \kappa_1^4} + \dots$$

where, for brevity,

$$(3.4) \quad f(r, j) = rh(rh - 1)(rh - 2) \dots (rh - j + 1)$$

and the Taylor series of the power function

$$(1 + x)^h = 1 + hx + h(h - 1)x^2/2! + \dots$$

was applied. The μ 's are moments of x' and, according to (3.2), central moments of the original variable X as well. It is now useful to adopt the cumulants κ of X as the current characteristics instead of the moments μ , because they give an essentially better convergence behaviour to the expansions. By using the relations referred to above between these sets of characteristics, equation (3.3) is transformed as follows:

$$(3.5) \quad \beta_r = 1 + f(r, 2) \frac{\kappa_2}{2\kappa_1^2} + f(r, 3) \frac{\kappa_3}{6\kappa_1^3} + f(r, 4) \frac{3\kappa_2^2 + \kappa_4}{24\kappa_1^4} \\ + f(r, 5) \frac{\kappa_2\kappa_3}{12\kappa_1^5} + f(r, 6) \frac{\kappa_2^3}{48\kappa_1^6} + \dots$$

The sequence

$$(3.6) \quad \rho_i = \kappa_i / \kappa_1^i$$

is assumed to be reasonably convergent when i increases (in the Poisson case ρ_i is of order n^{-i+1} , c.f. Section 5.1).

Finally, the central moments of y are the characteristics which are necessary for the transformation aimed at. They are calculated by means of the well-known general relationship between the central moments and the moments about zero ($\mu_2 = \beta_2 - \beta_1^2$, etc.). After elementary but quite tedious operations, and observing that the expression (3.5) can be expressed in terms of the cumulant ratios ρ_i , the following expansions result:

$$(3.7) \quad m_y = Ey = 1 + \frac{1}{2}h(h-1)\rho_2 + \frac{1}{24}h(h-1)(h-2)[4\rho_3 + 3(h-3)\rho_2^2] \\ + \frac{1}{48}h(h-1)(h-2)(h-3)[2\rho_4 + 4(h-4)\rho_2\rho_3 + (h-4)(h-5)\rho_2^3]$$

$$\mu_2(y) = \kappa_2(y) = \sigma_y^2 = h^2\rho_2 + \frac{1}{2}h^2(h-1)[2\rho_3 + (3h-5)\rho_2^2] \\ + \frac{1}{12}h^2(h-1)[(7h-11)\rho_4 + 4(h-2)(7h-12)\rho_2\rho_3 + 2(h-2)(7h^2-30h+32)\rho_2^3]$$

$$\mu_3(y) = \kappa_3(y) = h^3[\rho_3 + 3(h-1)\rho_2^2] \\ + \frac{1}{2}h^3(h-1)[3\rho_4 + 3(7h-10)\rho_2\rho_3 + (17h^2-55h+44)\rho_2^3]$$

$$\kappa_4(y) = \mu_4(y) - 3\mu_2^2(y) = h^4[\rho_4 + 12(h-1)\rho_3\rho_2 + (h-1)(16h-20)\rho_2^3].$$

Terms only having the order three or less were accepted. (I am grateful to my colleague Mr. H. Simberg for the correct κ_4 .)

We are now enabled to fix the parameter h . For the aimed at symmetrization of the transformed distribution its skewness should be made to vanish, or, what is the same, $\kappa_3(y)$ should be equal to 0. For the sake of computational convenience, only its leading term will be equated to zero. Hence

$$(3.8) \quad h = 1 - \frac{1}{3}\rho_3/\rho_2^2 = 1 - \frac{\kappa_1\kappa_3}{3\kappa_2^2}.$$

3.3. Transformation A. Haldane now states that, when the above value of h is

substituted into equations (3.7), the standardized variable

$$(3.9) \quad y_{\text{norm}} = (y - m_y) / \sigma_y$$

is "almost normally distributed" with a zero mean and standard deviation unity. Haldane calls this formula "Transformation A" as distinguished from another transformation to be dealt with shortly.

3.4. For the risk theory applications, where the mixed compound Poisson function is to be approximated, it is convenient to take the basic characteristics (see BPP, p.54):

$$(3.10) \quad \begin{aligned} \text{Mean} &= m_X = x_1 \\ \text{Standard deviation} &= \sigma_X = \sqrt{x_2} \\ \text{Skewness} &= \gamma_X = x_3 / \sigma_X^3 \end{aligned}$$

of the original distribution as the entries of the calculations. Accepting the terms of the order of at most two and introducing as an auxiliary quantity

$$(3.11) \quad s = \sigma_X / m_X$$

the Haldane approach A can be written in an operative form

$$(3.12) \quad \begin{aligned} h &= 1 - \frac{1}{3} \gamma_X / s \\ m_y &= 1 - \frac{1}{2} h(1-h) \left[1 - \frac{1}{4} (2-h)(1-3h)s^2 \right] s^2 \\ \sigma_y &= hs \sqrt{\left[1 - \frac{1}{2} (1-h)(1-3h)s^2 \right]} \\ y_{\text{norm}} &= [(1 + xs)^h - m_y] / \sigma_y \\ F(X) &= N(y_{\text{norm}}). \end{aligned}$$

The lower case x refers to the standardized variable (2.2). The third degree quantities x_3 (and γ_3) were eliminated for computational convenience by using (3.8).

Note that the moment $\mu_4(y)$ was not needed in this context. It was derived in (3.7) because it will be useful in later sections.

The above formulas are fairly comfortable for computer programming. Examples will be given later and the applicability discussed, but before that we will make some further remarks and present an extended version of the transformation.

3.5. Negative h values. Haldane limited the range of validity to positive h values only. Examples show that the formula also works in cases where h turns negative. However, negative values seem to appear in the area where the skewness is excessive and the goodness of fit is unsatisfactory.

3.6. The case $h = 0$. Special attention is called to the case when $h \rightarrow 0$. Then y_{norm} has the limit

$$(3.13) \quad y_{\text{norm}} = [\ln(1 + xs) + \frac{1}{2} s^2 - \frac{1}{4} s^4] / [s \sqrt{(1 - \frac{1}{2} s^2)}].$$

3.7. Wilson and Hilferty applied, as mentioned already, essentially the same transformation as Haldane for the evaluation of the gamma function and arrived at a constant value of $1/3$ for the parameter h . On the other hand we know that the original distribution can be approximated by the gamma function which is obtained by equating the mean, standard deviation and skewness with the corresponding characteristics of the distribution to be approximated. Hence we can expect that the Wilson–Hilferty formula may also approximate the original distribution. The result can be written as follows (see BPP, p.71)

$$(3.14) \quad \begin{aligned} F(X) &= \Gamma(\alpha + x\sqrt{\alpha}, \alpha) \\ &\approx N[c_1 + c_2(x + c_3)^{1/3}] \end{aligned}$$

where x is the standardized variable (2.2) and

$$c_1 = \frac{\gamma}{6} - \frac{6}{\gamma} \quad c_2 = 3\left(\frac{2}{\gamma}\right)^{2/3}; \quad c_3 = \frac{2}{\gamma}.$$

This formula is very comfortable for computer programming as is also its inverse. Therefore, it is tested in parallel with the Haldane and NP approximations in what follows.

3.8. A link to the NP formula can be found by expanding y_{norm} in (3.12) in terms of (xs) as a Taylor series and then expressing h by means of γ :

$$(3.15) \quad \begin{aligned} y_{\text{norm}} &= [(1 + xs)^h - m_y] / \sigma_y \\ &= [1 + hxs + \frac{1}{2}h(h-1)x^2s^2 + \dots - 1 - \frac{1}{2}h(h-1)s^2 - \dots] / (hs - \dots) \\ &= x + \frac{1}{2}(h-1)s(x^2 - 1) + \dots \\ &= x - \frac{1}{6}\gamma(x^2 - 1) + \dots \end{aligned}$$

But this is just what is also obtained if y in (2.5) is expanded in the terms of x (see BPP, p. 117, eq. (3.11.14)). Hence it can be expected that the Haldane and the NP formulas are close to each other at least in the area of the best convergence. This will be confirmed by examples given later.

4. HALDANE'S TRANSFORMATION B

Haldane also experimented with another formula, which is derived introducing *two parameters h and g* (instead of only one, h , above). They are assigned values which minimize both the skewness and the kurtosis of the transformed variable. The new parameter g is chosen so that the original variable X is first transformed to another variable which has g as its mean and consequently also as x_1 :

$$(4.1) \quad X'' = X + g - x_1.$$

Then the transformation $y = (X/m_X)^h$ is replaced by

$$(4.2) \quad y = (X''/E[X''])^h = [1 + (X - x_1)/g]^h.$$

The transformation (4.1) does not affect higher cumulants than κ_1 . Hence all the results are still valid, if κ_1 is replaced by g everywhere.

The parameters h and g are now determined by equating the leading terms of κ_3 and κ_4 . This implies that

$$(4.3) \quad g = \frac{12\kappa_2^2\kappa_3}{20\kappa_3^2 - 9\kappa_2\kappa_4}, \quad h = \frac{16\kappa_3^2 - 9\kappa_2\kappa_4}{20\kappa_3^2 - 9\kappa_2\kappa_4}.$$

By using the characteristics (3.10) and, in addition, the kurtosis γ_{2X} ($= \kappa_4/\sigma^4 = \mu_4/\sigma^4 - 3$) of the original variable X and further introducing, for brevity, the auxiliary coefficients

$$(4.4) \quad b = \frac{5}{3}\gamma_X - \frac{3}{4}\gamma_{2X}/\gamma_X; \quad c = \frac{4}{3}\gamma_X - \frac{3}{4}\gamma_{2X}/\gamma_X$$

the *Haldane transformation* B can be written as follows

$$(4.5) \quad \begin{aligned} h &= c/b \\ m_y &= 1 - \frac{1}{2}c(b - c)[1 + \frac{1}{4}(2b - c)(3c - b)] \\ \sigma_y &= |c| \sqrt{[1 + \frac{1}{2}(b - c)(3c - b)]} \\ y_{\text{norm}} &= [(1 + bx)^h - m_y]/\sigma_y \\ F(X) &= N(y_{\text{norm}}) \end{aligned}$$

where the lower case x again refers to the standardized aggregate claim (2.2).

Haldane made some reservations concerning the applicability of this transformation, mainly providing for the positivity of h . If the denominators of (4.3) are vanishing, the formulas become invalid.

5. ON THE APPLICABILITY OF THE HALDANE EXPANSION

5.1. General conditions for convergence of the expansions. Haldane assumed that the variable X is inherent from a collective, the risk volume of which can be described by a parameter n . In our risk theory applications, n can be just the expected number of claims as provided in Section 1.1. Furthermore, Haldane assumed that the cumulants κ_i for $i = 1, 2, 3$ and 4 are of the order n when n grows large and that cumulants for $i > 4$ are of the order n^{i-4} or less. Haldane states that the expansions concerned are asymptotically convergent when n is large enough, i.e. the transformations can be applied in large collectives.

In the Poisson case the cumulants are $\kappa_i = na_i$, a_i being the i th moment about zero of the claim size variable. Hence, the Haldane conditions are satisfied in so far as the moments a_i are finite. On the other hand the asymptotic behaviour of the Polya case does not fulfil the conditions. Moreover, the volume parameter n is always finite, often rather small, in practical applications. Then the Haldane criterion does not suit, because the convergence of the relevant expansions may be poor, even though they may asymptotically converge. Furthermore, one should appreciate that the convergence of the expansion itself does not guarantee full accuracy because there are other, deeper, aspects involved, e.g. those which

we already discussed preliminarily in Section 2.5 about the moment problem. Fortunately, and unexpectedly, the tests given in the following sections seem to prove that the approximation outcomes are also fairly satisfactory in numerous cases where the convergence criterion would suggest failure. Therefore, we do not feel that it is useful to explore the problem of the convergence of the Haldane expansions other than that a convergence indicator will be introduced in Section 5.3 below. Otherwise the original paper is again referred to.

5.2. Measures of deviation from normality were suggested by Haldane, making use of the residue skewness and kurtosis which remain as they are reduced by the symmetrization procedure. Hence (cf. (3.7))

$$(5.1) \quad \begin{aligned} \text{Measure 1} &= \gamma' = \kappa_3(y)/\sigma_y^3 \\ \text{Measure 2} &= \gamma_2' = \kappa_4(y)/\sigma_y^4 \end{aligned}$$

where the value (3.8) is to be assigned to h .

The same formulas are valid also for the transformation B, when h is taken from (4.3) and κ_1 is replaced by g ; it is also obtained from (4.3).

These measures will be illustrated by examples in Section 6.

5.3. Cumulant ratios. A crucial condition for the convergence of the expansions (3.7) is a rapid convergence of the sequence of the cumulant ratios ρ_i , defined by (3.6). Therefore, the author experimented with the indicator

$$(5.2) \quad \vartheta = \rho_4/\rho_3$$

as an alternative measure for applicability. If ϑ is small, it implies that the higher cumulant ratios can be expected to be negligible. Values of ϑ are given in the context of test examples and an overall view is provided by Figure A.7 (Appendix 2).

6. EXAMPLES

6.1. Tests. The approximation methods dealt with in the previous sections are tested by calculating a great number of numerical examples on the one hand by using the exact recursive formula, and on the other hand the NP, Wilson–Hilferty (briefly WH), Haldane-A (HA) and the Haldane-B (HAb) approaches. Both the Poisson case and the Polya case, having differing shape parameters h , were examined experimentally. The claim size distribution was the truncated Pareto or log-normal or their mixtures or could also be freely chosen (and given manually to the computer). Because the recursive technique (see details in BPP, Section 3.8) is applicable merely for discrete distributions, the claim sizes were discretized permitting only integer values $Z = 1 + id$ ($i = 0, 1, 2, \dots, I$) where d and I are freely eligible positive integer parameters.

The tested distributions, 54 in total, were chosen to cover broadly the area that is usually applied in risk theory considerations, and also to provide comparisons between the approaches. Regrettably, it is not possible to print all the data.

Typical cases only were picked for the tables and diagrams given in Appendixes 1 and 2. A comprehensive collection of the data will be deposited in the Library of the Actuarial Society of Finland (Address: Bulevardi 28, 00120, Helsinki 12). Copies are available upon request.

6.2. Appendix 1 exhibits exact F values and the approximated ones in parallel. Numerical values for the convergence criteria proposed above are also given in the side column of each distribution box of each of the tables. Discussion of the outcomes will be deferred to Section 7.

6.3. Figures A.1 and A.2 (Appendix 2) graphically present two of the distributions of Appendix 1 to provide a clearer illustration. The deviations between the relevant curves are so slight that they are scarcely discernible in cases where the distribution is not markedly skew. Therefore, the tails of Figure A.1(a) are plotted in a magnified scale in Figure A.1(b).

6.4. The effect of discretization. The deviations between the exact and the approximated values are partially due to the fact that the approximating functions always, more or less, deviate from the exact one and partially to the fact that the "exact" F is discrete but the approximating functions are continuous. This is clearly seen in Figure A.1(b). In order to eliminate the effect of this discrepancy from the tabulated outcomes, such as given in Appendix 1, the discrete F curves were replaced by a broken line which connected the midpoints of the upward steps. If this kind of smoothing of the discrete results is not made the comparison deviations depend on where, for the purpose of comparison, the selected values of the x variable are positioned on the x -axis. As seen in Figure A.1(b) the effect may be larger than the "genuine" deviations are, and depends on whether the test point happens to fall immediately before or after a step.

Information about the steps of F is provided in the last columns of the tables of Appendix 1 where the half of the step height ($= dF\%$) is given. It proved to be mostly larger than the approximation errors in the preceding columns as long as the skewness remained moderate. It depends on the actual relevant problem setting as to whether or not it should be regarded as appropriate to add both the errors.

Note that the discretization results in inaccuracy also if a continuous original claim size distribution, such the Pareto one, is replaced by a step function. This feature was discussed in BPP (Section 3.8c). Because both the exact method and the approximating methods were based on the same discretized claim size distribution, this inaccuracy did not appear in our tests.

6.5. Figure A.2 represents an extreme case where the skewness is large. Then all the approximations turn out irregular and lose their applicability particularly at the tails of the curves.

6.6. Figure A.3 attempts to provide a summarizing survey over a sample of distributions. The relative errors ($dNP\%$, $dWH\%$, etc. in Appendix 1) are

grouped according to different skewness ranges and argument values $x = -2, 2$ and 3 respectively and then shown in the specified diagram boxes. For example all the relative errors of the Haldane-A approximation for $x = 3$ and the skewness less than 0.3 are placed upon the "point" HA-a in the top-most right-hand box. The points inherent from the same distribution are connected by the lines in between.

This figure is meant to provide a rapid visual comparison of the tested approaches. A narrow bundle of the connecting lines indicates a good overall fit of the formula concerned.

6.7. The measures of deviation (5.1) are investigated in Figure A.4. This and the remaining figures are limited to the Haldane-A method only.

The tested cases for $x = -2$ and $x = 3$, respectively, are displayed in the diagram by using the measures 1 and 2 as coordinates. The relative errors $|dHA\%|$ are indicated by symbols, as shown in the figure.

As expected, the fit is good for small measure values. Another useful observation is that the measure values are well correlated, i.e. the points are clustered at a straight line. This suggests that it is sufficient to use only one of the measures, preferably the measure 1.

6.8. Convergence properties are studied in Figures A.5, A.6 and A.7. The tested cases were first placed in Figure A.5 by using the standardized x and the skewness as coordinates. The tests were made only for a sequence of discrete x values $= -2, -1.5, -1, \dots, 4$. For clarity, the points, such as in Figure A.4, were not plotted in the final diagram, but the zones where the errors $|dHA\%|$ having some specified magnitudes are positioned were used instead. For example, in the area below the zone boundary designated by 1 only cases that have $|dHA\%|$ less than 1% are found, and below the 3-boundary only cases having $|dHA\%|$ less than 3% , etc. More exactly, the points of the boundary numbered by N ($N = 1, 3, 5, 10$ or 25) were determined according to that sample case for which the relevant x -value had the error $|dHA\%| \geq N\%$ and the lowest skewness. Note that cases having $|dHA\%| < N\%$ may be also found above the N -boundary, even though they are mainly clustered below.

The fit is good as long as the skewness is relatively small. This is a well-known feature about, for example, the applicability of the NP method (BPP, Section 3.11c). Note that $1 - F$ is very small for $x > 3.5$ if the skewness is not excessive. Hence the poor relative accuracy in the lower right-hand corner is seldom harmful in applications.

The somewhat zigzag course of the zone boundaries is due to fact that the goodness of fit is sensitive to the selection of tested distributions. Of course, if another set of distributions were chosen, a more or less differing course for the boundaries would result. However, the number of tests, 54, was already so large and the selections so variable that it is not likely that any very essential differences would appear.

Figure A.6 represents the dependence of the error $|dHA\%|$ on the measure 1

using the same display technique as Figure A.5. This diagram contains the same information as Figure A.4, but in another shape and extended to more x values.

Finally, Figure A.7 describes the effect of the cumulant ratio convergence depicted by the same technique as applied in Figures A.5 and A.6 and by using the ratio ϑ (cf. (5.2)) as a measure candidate.

6.9. The moment problem that was discussed in Section 2.5 was explored by varying the claim size distribution subject to the conditions that its mean, standard deviation and skewness:

- (6.2) (1) m_Z , σ_Z , and γ_Z are fixed,
 (2) the claim sizes are limited to integer values $1, \dots, Z_{\max}$ and
 (3) the claim number distribution $p_k(n)$ is fixed.

These conditions determine a family of the mixed compound distribution functions that all are approximated by one and the same NP, WH or HA. We illustrated the problem in Section 2.5 by saying that the functions to be approximated and fulfilling the conditions (6.2) are spread in a more or less wide "funnel of doubt" confined by the upper and lower envelope curve. If the funnel is broad for the relevant argument values x , this implies that there is no single curve which could approximate well all of the original curves, i.e. the approximation problem based on the characteristics of (6.2) has no satisfactory solution despite the method used. Unfortunately, evaluation of the envelope curves proved intractable. However, in order to get a grasp of the magnitude of the funnel at the tails of the distribution a lower limit was experimented with as follows.

Some of the test cases presented in Appendix I were chosen as examples. The values of the mixed compound Poisson function F were then calculated for those two distributions that fulfil the conditions (6.2) and have maximal and minimal kurtosis respectively, or what is an equivalent provision, maximal or minimal fourth moment of the claim size distribution. Then also the kurtosis of the aggregate claim distribution is maximized and minimized respectively. It can be reasonably expected that these are the extreme distributions at the tails among all those permitted by the conditions.

It proved that, in the exemplified cases, these distributions were the most dangerous and least dangerous respectively in the meaning defined by GOOVAERTS *et al.* (1984), Section 4.4 (suitably choosing their limit constant β).

Table 1 exhibits two examples, one connected to case 3 of Table A.1 and another connected to case 7. Among all of the distributions having the same characteristics as the selected case those two that have the minimal and maximal kurtosis respectively were sought, and then the exact F was calculated for them also.

It proved that the funnel of doubt is very narrow as long as the skewness is moderate and the individual claim sizes have a reasonably low upper limit. This confirms the earlier experience that the mixed compound Poisson distribution is robust under these provisos. On the other hand the funnel is rather large for large skewness values. This confirms the fact that there cannot be any approximation

TABLE I
 $F(x)$ resp. $1 - F(x)$

x	-2	2	3	4	
Min	0 01612	0 02881	0 00309	0 00021	
Case 3	0.01620	0 02880	0 00312	0 00022	$\gamma_X = 0.24$
Max	0.01632	0 02879	0 00318	0 00023	
Min	0 00105	0 04327	0 01142	0 00243	
Case 7	0 00330	0 03758	0 01253	0.00417	$\gamma_X = 1.08$
Max	0 00507	0 03089	0 01236	0 00571	

based on the three lowest characteristics (6.2) which would fit in all cases and for the whole relevant range of the variable x .

Note that the three curves representing the parallel distributions, as given in the table, intersect each other. That means, for example, that the most dangerous curve is most dangerous only for rather high values of x . The requested funnel of doubt is obviously constituted as an area between the envelope curves in a rather complicated way. Further study of this interesting problem was deferred to a later date.

The smoothing mentioned above in 6.4 has also some effect on the breadth of the funnel, although not an essential one. If for example the height of the step is regarded in the numbers of Table 1 for the case 3 and $x = 4$, the minimum and the maximum should be replaced by 0.00020 and 0.00024 respectively. Similarly, the numbers corresponding to $x = 4$ of the latter example should be replaced by 0.00238 and 0.00580.

7. DISCUSSION

7.1. On accuracy. It would be, of course, highly desirable to find ways to determine rigorously the accuracy of the proposed approximations. Unfortunately, this has not been tractable as yet. Therefore, we have to collect experience by testing various distributions. If a method turns out to have consistently acceptable accuracy in numerous and relevant areas of application well covered by the tests, then the use of the method may be justified in practical calculations. The Figures A.4–A.7 are aimed to provide a survey in concentrated form of the expected accuracy. Three alternative indicators were introduced: the skewness, the Haldane's measure and the convergence of the cumulants. Obviously the skewness is most convenient, because it has to be calculated as one of the entries of the approximation calculations.

7.2. How accurate should the method be? In deeming the usefulness of the approximation one should also appreciate the fact that in many cases the basic data are highly uncertain. In particular this concerns the structure function H (1.3) and its parameters. These choices may have a great effect on the process to be evaluated. If the initial data are inaccurate, then it is meaningless to demand

essentially greater accuracy from the calculation technique, at least if this can be done only at the cost of greatly complicating the calculations. On the other hand, the collectives concerned are often fairly large and the top risks are cut away by reinsurance. Then the skewness may seldom exceed 0.1 or 0.2 and the inaccuracy involved with the approximation formulas obviously scarcely spoils the outcomes.

The situation is different for problems where long chains of computations are needed, e.g. in the calculation of integrals having F in the integrand. Then one should beware of an accumulation of errors.

7.3. An appropriate tool for simulations. Before proceeding further with the discussion about approximation methods attention is called to an attractive feature of the formula of type (2.4). It can be of special benefit for simulations where random numbers are generated, which are distributed according to the mixed compound Poisson law. This is a problem frequently appearing in advanced model building. The approach is simply first to generate normally distributed random numbers r and then to transform them by the inverse of the symmetrizing function (2.3): $X = v^{-1}(r)$ (see BPP, Section 6.8).

The number of the necessary random numbers can be very great. Then it is important that the inverse transformation v^{-1} is convenient to program and is fast. We proposed the NP formula in BPP (Section 6.8.3). The present experience suggests either the WH or HA-A transformations. In particular the WH formula is very handy (which was already recognized in exercise 6.8.1 of BPP).

7.4. Observations. Appendix 1 and Figure A.3 are the most convenient for the evaluation and comparison of the four tested methods.

If the skewness is moderate, i.e. no more than 0.3, and if an inaccuracy of some ± 2 per cent is tolerable, then all four methods are acceptable. However, the Haldane-B showed, by far, the narrowest range of the relative error, the Haldane-A being obviously the next best.

The situation is greatly different for the skewness values 0.3–1. Then the Haldane-B fails for $x = 3$ (note the different scale for the different lines of Figure A.3!). By the way, a similar observation was also made when the long version of the NP formula was investigated (PENTIKAINEN, 1977). These approaches, which are based on four characteristics, kurtosis included, instead of three characteristics (mean, standard deviation and skewness), proved to have superior accuracy for slightly skewed distributions but do not tolerate markedly skewed cases.

When the skewness exceeds unity, then all of the methods already show great irregularities and soon turn out to be useless. The lower example of Table 1 (Section 6.9) suggests that no method that is based only on the three lowest characteristics can be good for all greatly skewed distributions.

A general observation is that the short tail ($x < 0$) shows considerably worse results than the long tail.

For the reasons referred to above it seems doubtful whether the Haldane

variant B is useful, in particular regarding the fact that it is markedly more complicated than the A variant. Haldane himself also observed that it gives sometimes poorer results than the simpler A formula.

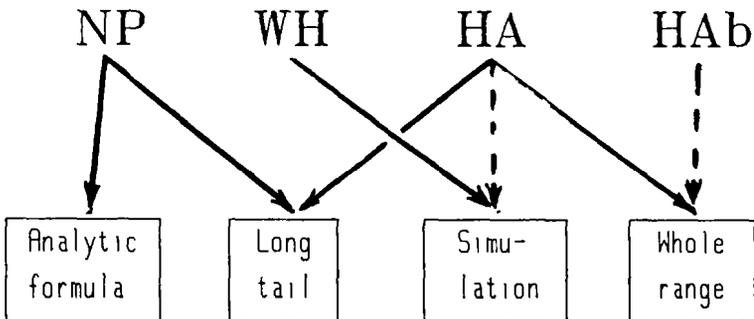
Some rules of thumb are sometimes proposed to guide the use of the approximations (e.g., BPP, 3.11e). Our latest studies do not suggest any such simple rules. Instead, it is much more effective to use Figure A.5 (or Figures A.6 or A.7) as a kind of “map” where the possible accuracy can be evaluated, and just in the environment of concern. If, for example, only positive x values are needed for some particular interval, then the area of applicability is wider than if negative x values are also needed.

Note that even though the general shapes of the error zones in Figures A.5, A.6 and A.7 are similar, it does not imply a full similarity in the test outcomes. For instance, the ϑ indicator would suggest a poor accuracy in case 4 of Appendix 1, but the skewness and Haldane measures still indicate acceptability as seen in the side column of the table.

7.5. The Wilson–Hilferty formula is clearly simpler and also somewhat faster than the Haldane-A. However, its tolerance for medium size and large skewnesses is poorer, as seen from Figure A.3. If the skewness is moderate, this formula may be appropriate at least in cases where very great speed is necessary.

7.6. The NP method has as its special merit the analytic form (2.5) for the long tail. It is of frequent use in many risk theory considerations (see e.g., BPP, Chapter 4). If only the long tail is of concern, then the NP method is the simplest and is also fairly competitive with the other methods concerning the accuracy, with the proviso that the distribution is not very skew.

7.7. In conclusion we summarize our present conception about the usefulness of the studies’ approaches by means of a diagram as follows:



7.8. Finally let us note that the exact and approximate methods complement each other in a happy way. The exact methods (and possibly direct simulations, see BPP, p.239) are most appropriate for small collectives, and the approximate formulas for the large ones.

REFERENCES

- BEARD, R E , PENTIKAINEN, T. and PESONEN, E (1984) *Risk Theory* Chapman and Hall, London (referred as BPP in the text)
- BERTRAM, J. (1981) Numerische Berechnung von Gesamtschadenverteilung. *Blatter der Deutschen Gesellschaft fur Versicherungsmathematik*
- BOHMAN, H and ESSCHER, F (1964) Studies in risk theory with numerical illustrations *Scandinavian Actuarial Journal*
- BOX, G E B and COX, D. R (1964) An analysis of transformations *Journal of the Royal Statistical Society B26*, 211–252
- ESSCHER, F (1932) On the probability function *Scandinavian Actuarial Journal*.
- GOOVAERTS, M J , DE VYLDER, F and HAEZENDONCK, J (1984) *Insurance Premiums* North-Holland, Amsterdam
- GOOVAERTS, M J , and KAAS, R (1986) Best bounds for positive distributions with fixed moments. *Insurance Mathematics & Economics* 5, 1
- HALDANE, J B S (1938) The approximate normalization of a class of frequency distributions *Biometrika* 29, 392–404
- KENDALL M and STUART, A (1979) *The Advanced Theory of Statistics* Vol I Charles Griffin, London
- LAU, H-S (1984) An effective approach for estimating the aggregate loss *Journal of Risk and Insurance*, LI.1
- OSCHWALD, M. (1984) Gamma Power-Entwicklung zur Berechnung der Verteilungsfunktionen des Gesamtschadens *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*
- PENTIKAINEN, T (1977) On the approximation of the total amount of claims *ASTIN Bulletin* 9, 281–289.
- PUSA, O (1985) Tests of the fit of the Esscher and NP formulas, Examination Paper, as yet available in Finnish only
- WILSON, E B and HILFERTY, MARGARET (1931) The distribution of chi-square *Proceedings of National Academy of Science, USA* 17, 684–688

APPENDIX I

EXAMPLES

n = Poisson parameter. $m, \sigma, \gamma, \gamma_2$ are the mean, standard deviation, skewness and kurtosis of the aggregate claim X .

h = Polya parameter.

ha = Haldane parameter (3.8).

r_i ($i = 1, 2$) is the risk indexes a_i/a_j indicating the heterogeneity of the claim size distribution (see BPP, p.54).

F = d.f. of X for $x < 0$ and $1 - F$ for $x > 0$ (1.1).

x = standardized aggregate claim size (2.2).

$dF\%$ = half of the step of F in per cent at the points where the discretized probability mass is concentrated (= $50 \star [F(x+) - F(x-)]$ divided by $F(x+)$ or $1 - F(x-)$).

NP = F approximated by the NP formula, $dNP\%$ its deviation from F in per cent. WH, HA and HAb are the corresponding outcomes for the Wilson–Hilferty, Haldane A and Haldane B formulas

$\text{Ln}(\dots)$, the log normal claim size distribution having the mean, standard deviation and the skewness given in parentheses.

Pareto(.) Pareto claim size distribution with index given in parentheses (see BPP, p.74).

M is the greatest value of the discretized claim size and d is the interval between the consecutive non-zero points (see Section 6.1).

1A and 2A are the measures of deviation defined by (5.1) for the Haldane A and 1B and 2B for the Haldane B.

$\nu = \rho_4/\rho_3 =$ an indicator for the speed of convergence of the cumulant.

Case 1	n	h	m	σ	γ	γ_2	ha	r_2	r_3		
	200 0	0 0	1998 4	145 9	0 080	0 007	0 633	1 07	1 25		
Ln(10, 3, 5)											
$M = 31, d = 1$	x	F	NP	WH	HA	HAb	dNP%	dWH%	dHA%	dHAb%	dF%
1A = -0 0001	-2 0	0 0205	0 0205	0 0205	0 0205	0 0205	-0 1	0 1	-0 0	-0 0	0 9
2A = +0 0009	-1 5	0 0646	0 0646	0 0645	0 0646	0 0645	0 1	-0 0	-0 0	-0 0	0 7
1B = -0 0001	-1 0	0 1586	0 1588	0 1586	0 1587	0 1586	0 1	-0 0	0 0	-0 0	0 5
2B = -0 0000	1 0	0 1586	0 1587	0 1586	0 1587	0 1586	0 0	-0 0	0 0	-0 0	0 5
$\nu = 0 006$	2 0	0 0249	0 0249	0 0249	0 0249	0 0249	0 1	0 0	-0 0	0 0	0 8
	3 0	0 0019	0 0019	0 0019	0 0019	0 0019	0 4	0 5	-0 1	0 0	1 0
	4 0	0 0001	0 0001	0 0001	0 0001	0 0001	1 0	1 7	-0 5	0 0	1 2
Case 2											
	n	h	m	σ	γ	γ_2	ha	r_2	r_3		
	100 0	100 0	787 1	116 9	0 224	0 071	0 497	1 21	1 72		
Mixture											
$M = 31, d = 1$	x	F	NP	WH	HA	HAb	dNP%	dWH%	dHA%	dHAb%	dF%
1A = -0 0008	-2 0	0 0164	0 0163	0 0164	0 0163	0 0164	-0 5	0 3	-0 3	-0 1	1 3
2A = +0 0039	-1 5	0 0599	0 0605	0 0599	0 0599	0 0599	0 9	-0 0	-0 0	-0 0	1 0
1B = -0 0002	-1 0	0 1582	0 1597	0 1580	0 1583	0 1582	0 9	-0 1	0 0	-0 0	0 7
2B = -0 0000	1 0	0 1583	0 1587	0 1581	0 1583	0 1583	0 2	-0 1	0 0	-0 0	0 6
$\nu = 0 048$	2 0	0 0284	0 0286	0 0284	0 0284	0 0284	0 5	0 0	0 0	0 0	0 9
	3 0	0 0029	0 0029	0 0030	0 0029	0 0029	0 4	0 8	-0 4	0 1	1 1
	4 0	0 0002	0 0002	0 0002	0 0002	0 0002	-0 3	2 9	-1 3	0 1	1 3
Case 3											
	n	h	m	σ	γ	γ_2	ha	r_2	r_3		
	100 0	200 0	141 3	19 4	0 238	0 100	0 424	1 39	3 59		
Pareto(3)											
$M = 21, d = 1$	x	F	NP	WH	HA	HAb	dNP%	dWH%	dHA%	dHAb%	dF%
1A = -0 0023	-2 0	0 0162	0 0159	0 0160	0 0160	0 0161	-1 9	-1 1	-1 4	-0 5	8 1
2A = +0 0175	-1 5	0 0595	0 0601	0 0594	0 0594	0 0594	0 9	-0 2	-0 2	-0 2	6 2
1B = -0 0004	-1 0	0 1578	0 1598	0 1580	0 1581	0 1577	1 2	0 1	0 2	-0 1	4 3
2B = -0 0000	1 0	0 1579	0 1587	0 1581	0 1582	0 1578	0 4	0 1	0 1	-0 1	3 6
$\nu = 0 058$	2 0	0 0288	0 0289	0 0288	0 0288	0 0288	0 4	-0 1	-0 1	-0 1	4 7
	3 0	0 0031	0 0031	0 0031	0 0030	0 0031	-2 1	-1 7	-2 3	-0 3	6 0
	4 0	0 0002	0 0002	0 0002	0 0002	0 0002	-8 2	-5 1	-7 2	-0 4	6 4

Case 4	n	h	m	σ	γ	γ_2	ha	r_2	r_3
	10 0	150 0	50 0	17 9	0 463	0 279	0 569	1 21	1 87

Ln(5, 3, 4)	κ	F	NP	WH	HA	HAb	dNP%	dWH%	dHA%	dHAb%	dF%
$M = 21, d = 1$											
1A = -0 0133	-2 0	0 0088	0 0097	0 0094	0 0080	0 0085	10 3	7 4	-8 5	-2 5	14 2
2A = +0 0214	-1 5	0 0506	0 0532	0 0504	0 0499	0 0501	5 3	-0 2	-1 4	-0 9	8 5
1B = -0 0044	-1 0	0 1572	0 1630	0 1558	0 1575	0 1570	3 7	-0 9	0 2	-0 2	5 3
2B = -0 0005	1 0	0 1576	0 1587	0 1566	0 1578	0 1574	0 7	-0 6	0 1	-0 1	3 7
$\nu = 0 215$	2 0	0 0337	0 0343	0 0337	0 0338	0 0338	1 7	-0 2	0 2	0 1	4 8
	3 0	0 0050	0 0051	0 0051	0 0049	0 0050	3 1	3 2	-0 4	0 5	5 6
	4 0	0 0005	0 0006	0 0006	0 0005	0 0005	4 9	11 2	-2 4	1 0	6 2

Case 5	n	h	m	σ	γ	γ_2	ha	r_2	r_3
	100 0	100 0	170 4	34 2	0 593	0 679	0 016	3 03	36 93

Pareto(2)	x	F	NP	WH	HA	HAb	dNP%	dWH%	dHA%	dHAb%	dF%
$M = 61, d = 3$											
1A = -0 0250	-2 0	0 0078	0 0067	0 0059	0 0067	0 0071	-14 1	-25 1	-14 4	-9 1	6 2
2A = +0 0582	-1 5	0 0456	0 0492	0 0443	0 0447	0 0447	7 9	-2 8	-2 0	-1 1	4 5
1B = +0 0196	-1 0	0 1506	0 1658	0 1538	0 1522	0 1510	10 1	2 1	1 1	0 3	2 9
2B = -0 0035	1 0	0 1529	0 1587	0 1554	0 1541	0 1529	3 8	1 6	0 8	0 0	0 9
$\nu = 0 006$	2 0	0 0362	0 0372	0 0361	0 0359	0 0358	2 8	-0 1	-0 6	-0 9	2 3
	3 0	0 0068	0 0064	0 0064	0 0066	0 0069	-4 9	-5 4	-2 1	1 4	2 5
	4 0	0 0011	0 0009	0 0009	0 0010	0 0012	-17 4	-12 2	-0 8	11 9	2 7

Case 6	n	h	m	σ	γ	γ_2	ha	r_2	r_3
	25 0	20 0	47 6	17 5	0 779	0 976	0 297	2 15	13 25

Pareto(2)	κ	F	NP	WH	HA	HAb	dNP%	dWH%	dHA%	dHAb%	dF%
$M = 31, d = 1$											
1A = -0 0092	-2 0	0 0028	0 0035	0 0018	0 0020	0 0022	28 7	-35 4	-26 5	-22 0	22 9
2A = +0 0115	-1 5	0 0354	0 0434	0 0341	0 0344	0 0346	22 6	-3 6	-2 7	-2 2	11 8
1B = +0 0052	-1 0	0 1478	0 1712	0 1495	0 1491	0 1489	15 9	1 2	0 9	0 8	6 5
2B = -0 0002	1 0	0 1526	0 1587	0 1533	0 1530	0 1528	3 9	0 4	0 2	0 1	3 3
$\nu = 0 006$	2 0	0 0394	0 0411	0 0392	0 0391	0 0391	4 3	-0 7	-0 9	-1 0	4 1
	3 0	0 0084	0 0084	0 0083	0 0083	0 0083	1 0	-1 2	-0 8	-0 6	4 3
	4 0	0 0015	0 0014	0 0015	0 0016	0 0016	-5 4	-0 0	1 8	2 9	4 9

Case 7	n	h	m	σ	γ	γ_2	ha	r_2	r_3
	75 0	75 0	122 0	30 8	1 082	2 703	-0 429	3 77	84 37

Pareto(2)	x	F	NP	WH	HA	HAb	dNP%	dWH%	dHA%	dHAb%	dF%
$M = 121, d = 4$											
1A = +0 4175	-2 0	0 0033	0 0009	0 0000	0 0006	0 0011	-73 1	-100 0	-83 1	-67 8	10 0
2A = +0 5406	-1 5	0 0322	0 0342	0 0145	0 0209	0 0204	6 0	-55 2	-35 1	-36 5	5 9
1B = +0 6058	-1 0	0 1357	0 1834	0 1384	0 1307	0 1207	35 2	2 0	-3 7	-11 1	3 8
2B = -0 2405	1 0	0 1376	0 1587	0 1491	0 1419	0 1285	15 3	8 4	3 1	-6 6	2 1
$\nu = 0 630$	2 0	0 0376	0 0470	0 0431	0 0407	0 0379	25 1	14 7	8 3	0 7	1 9
	3 0	0 0125	0 0119	0 0112	0 0116	0 0139	-4 9	-10 4	-7 3	11 3	1 6
	4 0	0 0042	0 0027	0 0028	0 0035	0 0068	-35 5	-34 0	-16 5	63 7	1 9

Case 8	n	h	m	σ	γ	γ_2	ha	r_2	r_3		
	25 0	50 0	42 7	16 6	1 628	5 801	-0 394	3 29	54 61		
Pareto(2)											
$M = 91, d = 3$	x	F	NP	WH	HA	HAb	dNP%	dWH%	dHA%	dHAb%	dF%
1A = +1 4047	-2 0	0 0004	0 0000	0 0000	0 0000	0 0000	-95 8	-100 0	-100 0	-99 4	41 8
2A = +1 0048	-1 5	0 0178	0 0197	0 0000	0 0024	0 0026	10 7	-100 0	-86 5	-85 5	16 2
1B = +2 2843	-1 0	0 1190	0 2180	0 0950	0 0892	0 0702	83 2	-20 1	-25 0	-41 0	8 9
2B = -1 5625	1 0	0 1280	0 1587	0 1399	0 1248	0 0983	24 0	9 3	-2 5	-23 2	3 7
$v = 1 387$	2 0	0 0370	0 0562	0 0473	0 0395	0 0291	51 8	27 7	6 7	-21 4	3 4
	3 0	0 0140	0 0184	0 0158	0 0139	0 0119	31 4	12 6	-0 9	-14 7	2 4
	4 0	0 0069	0 0057	0 0053	0 0054	0 0065	-18 1	-24 1	-21 4	-6 4	1 8

APPENDIX 2

FIGURES

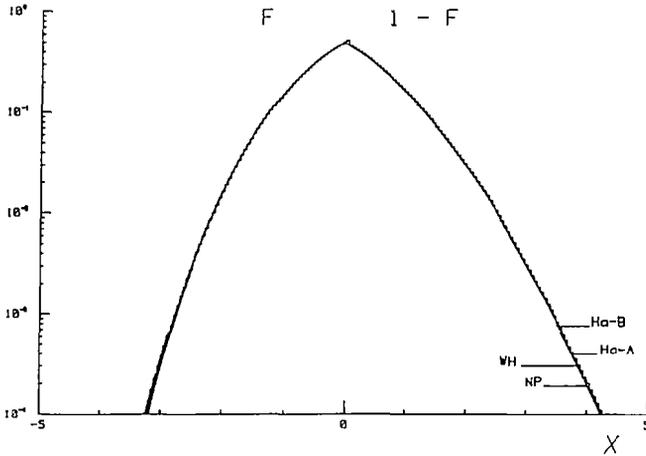


FIGURE A 1(a) The case 3 of Appendix 1 presented as a graph. The step curve represents the exact F . The key parameters are $n = 100, h = 200, \gamma = 0.24$

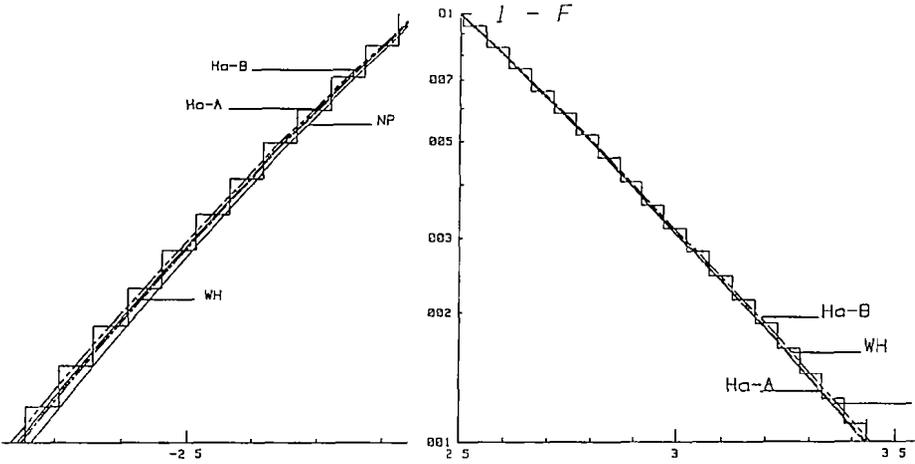


FIGURE A 1(b) The tails of the curves of Figure A 1(a) plotted in a magnified scale

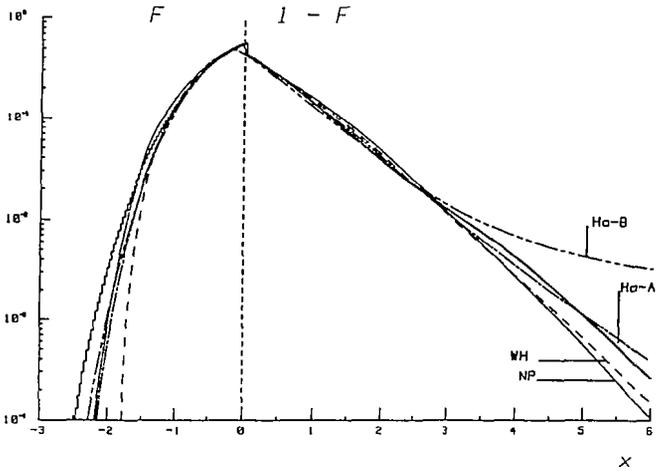


FIGURE A 2 The case 7 of Appendix 1 The key parameters are $n = 75$, $h = 75$, $\gamma = 1.08$

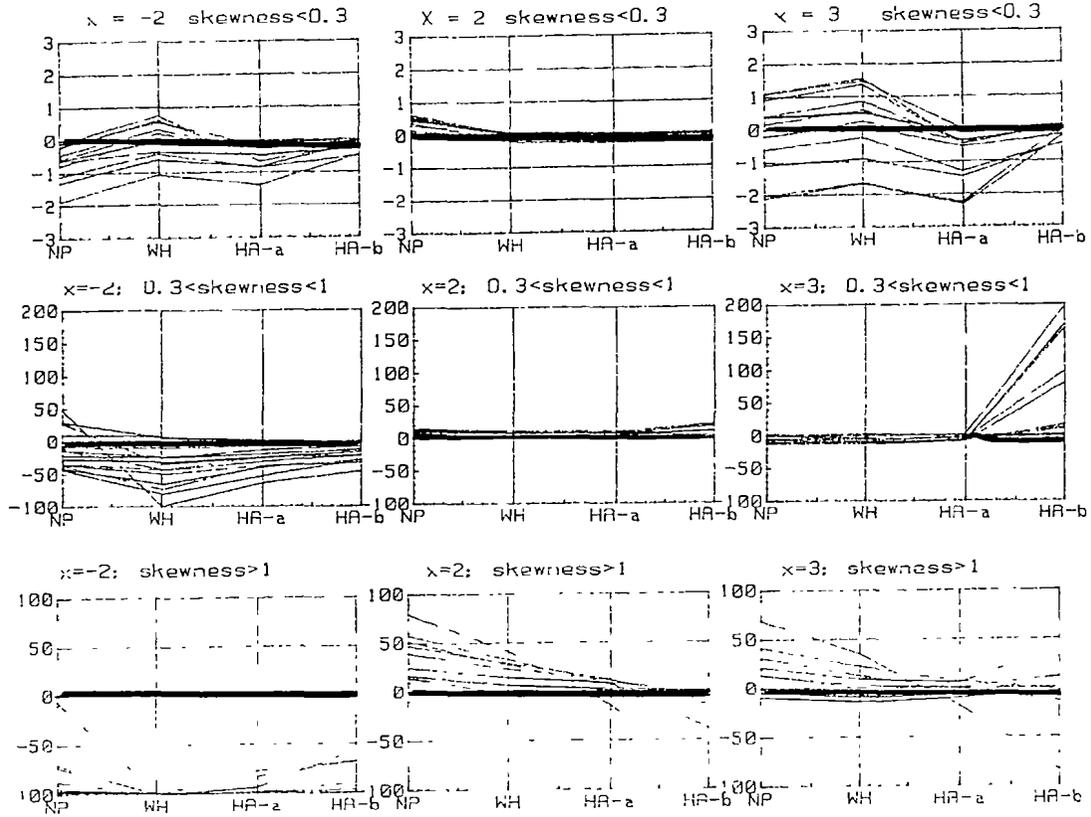


FIGURE A 3. The relative errors $dNP\%$, $dWH\%$, $dHA\%$ and $dHAb\%$ (see Appendix 1) grouped according to specified skewness and x values (see Section 6.5 for further explanations) calculated for a selected sample of distributions

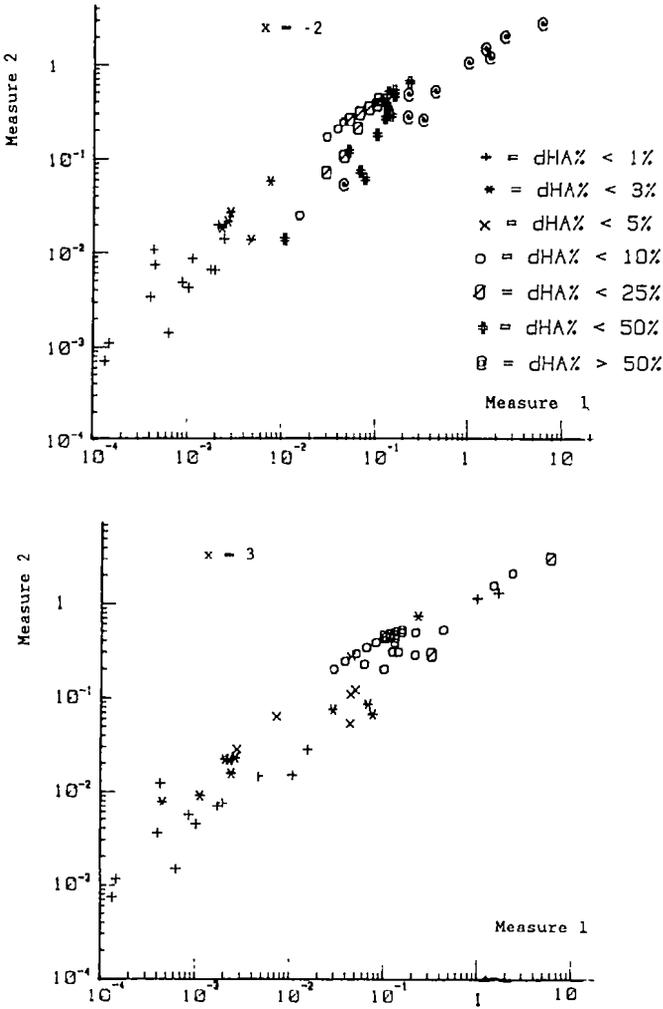


FIGURE A 4 The relative errors dHA% of the Haldane-A formula displayed according to absolute values of the measures 1 and 2 as defined by (5.1) The order of $|dHA\%|$ is indicated by the symbols given in the graph

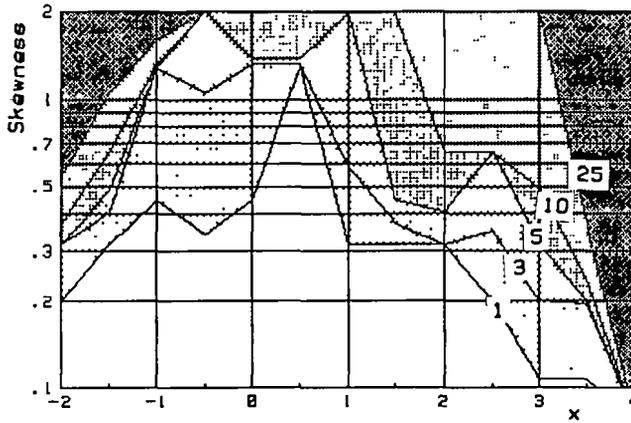


FIGURE A 5 The zones of the relative errors $|dHA\%|$ according to the argument x and the skewness
For explanations see Section 6 8

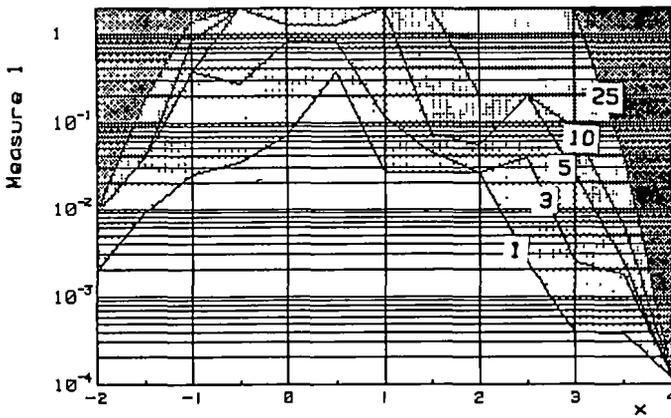


FIGURE A 6 The zones of the relative errors $|dHA\%|$ according to the argument x and the measure I

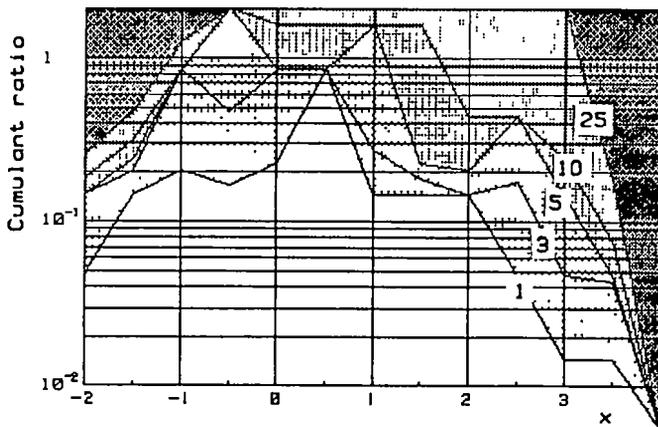


FIGURE A 7 The zones of the relative errors $|dHA \%$ according to the argument x and the cumulant convergence indicator v (see the explanations of Appendix 1)

TEIVO PENTIKAINEN
Kasavuorentie 12 C 9, SF-02700 Kaunainen, Finland.

