Ann. Inst. Statist. Math.
32 (1980), Part B, 465-480

COMPARISONS OF MODELS FOR ESTIMATION OF SAFE DOSES USING MEASURES OF THE HEAVINESS OF TAIL OF A DISTRIBUTION

TAKEMI YANAGIMOTO AND DAVID G. HOEL

(Received Mar. 1, 1980; revised July 30, 1980)

Summary

Procedures to estimate a dose of which the incidence probability is very small (e.g. 10^{-6}) have been developed to evaluate the safety of chemical compounds. To compare models for estimation of safe doses quantitatively, a measure of the heaviness of tail of a distribution and a measure of tail at the origin are introduced. These measures have a theoretical basis for the comparison of tail behavior between distributions. Using the two measures, a tail ordering is defined to present a criterion for the comparison of models and is discussed for the probit, the logit, the Weibull, the (generalized) multihit, the (generalized) multitarget and the multistage models.

The multistage model is most conservative among them, while the probit model has the reverse property. The Weibull model is more conservative than the logit. The multihit and multitarget models are found to be more sensitive than the Weibull and the logit.

1. Introduction

Recently many statisticians have become interested in the methodology of safety evaluation of chemical compounds. Procedures are being developed which depend upon dose-response relationships where the response of primary interest is usually carcinogenesis. (The reader is referred to [3], [9], [10], [11], [14] for a sample of such procedures.) Most investigators have concentrated on the mathematical and computational techniques for treating data under an assumed, dose-response model. Less attention has been given to the choices of reasonable models which is a difficult or impossible problem without adequate biological guidance.

In this paper we shall look at several of the dose-response functions. We will use the concept of the heaviness of the tail of a distribution which has been of value in reliability theory, robust estimation and nonparametric inference [1], [5], [13]. The heaviness of the tail of the dose-response model distributions will be considered numerically. The results will hopefully provide some insight into model selection.

Two measures of the heaviness of tail of a distribution are introduced. One of them is defined by the standard deviation of a tolerance distribution for log-transformed variable. The other is a measure of the behavior of the distribution at the origin.

We restrict ourselves mainly to six families of distributions which are used as representing tolerance distributions and are listed in Table 1. We do not distinguish among a distribution, the distribution function and a random variable with the distribution for simplicity, unless some confusion would result.

Table 1. Models in study, the corresponding tolerance distributions and the distribution functions

Model	Corresponding tolerance distribution	Distribution function[1)]
Probit model	Lognormal, $\mathcal{L}_N(\nu)$ $(\nu>0)$	$F_N(\nu, x)=\Phi(\nu \log x)$
Logit model	Loglogistic, $\mathcal{L}_L(\beta)$ $(\beta>0)$	$F_L(\beta, x)=x^\beta/(1+x^\beta)$
Weibull model	Weibull, $\mathcal{L}_W(\gamma)$ $(\gamma>0)$	$F_W(\gamma, x)=1-\exp\{-x^\gamma\}$
Multihit model (generalized)	Gamma, $\mathcal{L}_G(k)$ $(k>0)$	$F_G(k, x)=\int_0^x t^{k-1}e^{-t}/\Gamma(k)dt$
Multitarget model (generalized)	Minimum Exponential, $\mathcal{L}_M(\lambda)$ $(\lambda>0)$	$F_M(\lambda, x)=(1-e^{-x})^\lambda$
Multistage	Exponential Polynomial, $\mathcal{L}_P(n, \alpha_1,\cdots, \alpha_n)$ $(\alpha_i\geqq 0$ and $\sum \alpha_i>0)$	$F_P(n, \alpha_1,\cdots,\alpha_n, x) =1-\exp\{-\sum(\alpha_i x)^i\}$

1) A scale parameter is preassigned as 1 for simplity.

The paper is constructed as follows: In Sections 2 and 3, a measure of tail and a measure of tail at the origin are defined and their properties are given. A tail ordering among distributions using the two measures of tail is defined, and tolerance distributions are compared by the ordering in Section 4. The results are examined by calculating the tail probabilities at the origin for various specific distributions in Section 5. Finally, in Section 6 the results are also examined practically using twelve data sets.

2. Measure of tail of a positive random variable

Let X be a random variable and $F(x)$ be its distribution function. Throughout this paper it is assumed that $F(0)=0$ and $F(x)$ has a positive density function $f(x)>0$.

As seen in the probit and the logit analyses, dose data are initially log-transformed. The distribution of $\log X$ (with distribution function $F(e^x)$) is then of interest as well as that of X.

A measure of tail of X is defined by

DEFINITION 1. $\tau(X)=\sqrt{\mathrm{Var}(\log X)}$.

There are other measures of tail such as Gini's coefficient of concentration, the kurtosis and the coefficient of variation. But these measures are inconvenient for our purposes. Here it should be noted that our definition has a theoretical basis. As shown in [14], the heaviness of tail of X and the size of dispersion of $\log X$ are closely related. Let Y be a random variable with the distribution function, $G(x)$. Y is called to have heavier tail than X, iff $G^{-1}(u)/F^{-1}(u)$ is increasing in $1>u>0$, which is denoted by $Y \succ X(\mathcal{T})$. $Y \succ X(\mathcal{T})$, iff $G^{*-1}(u)-G^{*-1}(v) \geqq F^{*-1}(u)-F^{*-1}(v)$ for $1>u>v>0$, where $G^*(x)=G(e^x)$ and $F^*(x)=F(e^x)$. This condition means that $\log Y$ has the larger size of dispersion than $\log X$. Since the standard deviation is a natural measure of the size of dispersion, our definition become to be natural. In addition, $\tau(X)/\log_e 10$ in the probit model is equal to the inverse of the so called slope. Thus $\tau(X)$ in a general model can be considered to be a global measure of the inverse of the slope of the dose-response curve to log-transformed doses.

In many cases the calculation of the measure of tail is not complicated. For several families of distributions, the measures of tail are represented in the following proposition. The proofs are found in literatures for example [7], [8].

PROPOSITION 1. The measure of tail of a random variable X, $\tau(X)$, is represented as follows.

(i) Lognormal distribution: Suppose that X has the distribution function $F_N(\nu, x)$. Then $\tau(X)=1/\nu$.

(ii) Loglogistic distribution: Suppose that X has the distribution function $F_L(\beta, x)$. Then $\tau(X)=\pi/\sqrt{3}\,\beta$.

(iii) Gamma distribution: Suppose that X has the distribution function $F_G(k, x)$. Then

$$\tau(X)=\sqrt{\sum_{n=0}^{\infty} 1/(n+k)^2}=\sqrt{1/k+2/k^2+1/6k^3-\sum_{n=0}^{\infty} 1/\{(n+k)(n+1+k)\}^3}.$$

(iv) Weibull distribution: Suppose that X has the distribution function $F_W(\gamma, x)$. Then $\tau(X)=\pi/\sqrt{6}\,\gamma$. Especially, the measure of tail of the exponential distribution is $\pi/\sqrt{6}$.

The measures of tail of the minimum exponential and the exponential polynomial distributions can not be represented explicitly, though a recursive formula for $\tau(F_M(\lambda, x))$ can be obtained only when λ is a positive integer [14].

Table 2. Measures of tail of distributions in typical families at various levels of parameters

Parameter	Lognormal $F_N(\nu, x)$	Loglogistic $F_L(\beta, x)$	Weibull $F_W(\gamma, x)$	Gamma $F_G(k, x)$	Minimum Ex. $F_M(\lambda, x)$
0.5	2.	3.628	2.565	2.221	2.210
0.6	1.667	3.023	2.138	1.907	1.897
0.7	1.429	2.591	1.832	1.683	1.676
0.8	1.25	2.267	1.603	1.516	1.511
0.9	1.111	2.015	1.425	1.387	1.384
1	1.	1.814	1.283	1.283	1.283
2	.5	.907	.641	.803	.827
3	.333	.605	.428	.628	.670
4	.25	.453	.321	.533	.587
5	.2	.363	.257	.470	.534
6	.167	.302	.214	.426	.497
7	.143	.259	.183	.392	.469
8	.125	.227	.160	.365	.447
9	.111	.202	.143	.343	.429
10	.1	.181	.128	.324	.414

Table 2 presents the numerical values of $\tau(X)$ for the distributions with the parameter of each distribution ranging over values 0.5(0.1)1 (1)10. Figure 1 illustrates the values for the parameters between 0.5 and 5.0. These calculations show that $\tau(F_G(k))$ and $\tau(F_M(\lambda))$ are close to each other when their parameters k and λ are identified.

The parameters of each of the above distributions can be regarded as a " tail parameter ", though they are usually called a shape parameter. The nomenclature of the shape parameter is too general. Parameters in the lognormal, the loglogistic and the Weibull distributions should be referred to a power parameter [15]. In each, the scale parameter has been assigned the value 1, since it is a nuisance for comparison purposes.

The notion of the heaviness of tail is popular in the various fields of statistics. X is called to be increasing hazard (failure) rate average, iff the exponential distribution function, $1-e^{-x}$ has heavier tail than $F(x)$. The following proposition presents two general properties of the measure of tail which are useful for subsequent discussions.

PROPOSITION 2. (i) $Y \succ X(\mathcal{T})$ implies $\tau(Y) \geqq \tau(X)$. Thus within the families discussed here except for the exponential polynomial distributions the measure of tail is decreasing in its parameter.

(ii) Suppose $\tau(X) < \infty$. Then $\tau(X^\alpha) = |\alpha| \tau(X)$ for any α.

Recently, the multistage model was discussed in [3] and has attracted many researchers' attention. The corresponding distribution to the multistage model is the exponential polynomial distribution, which

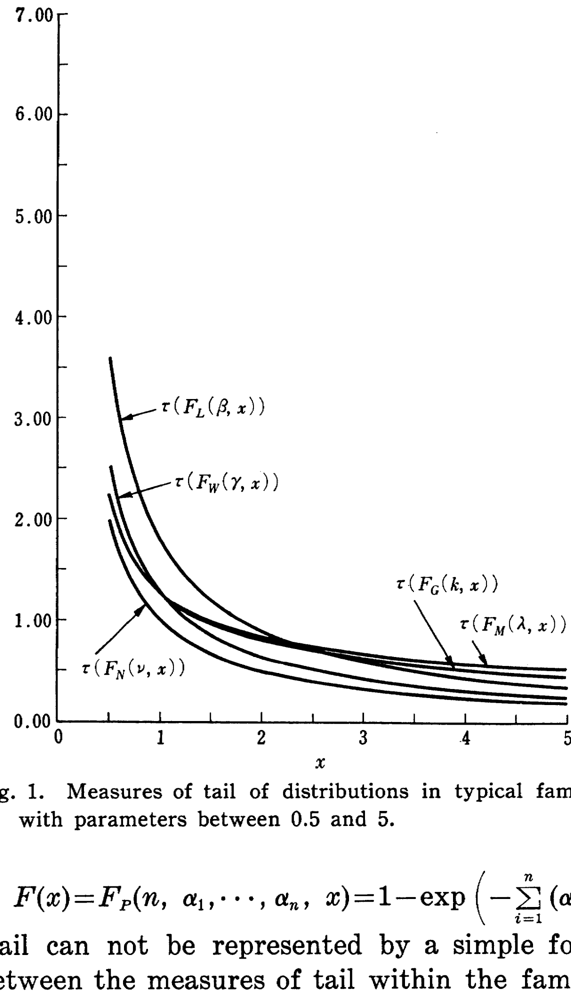

Fig. 1. Measures of tail of distributions in typical families with parameters between 0.5 and 5.

is given by $F(x)=F_P(n, \alpha_1, \cdots, \alpha_n, x)=1-\exp\left(-\sum_{i=1}^{n}(\alpha_i x)^i\right)$, and the measure of tail can not be represented by a simple form. However, inequalities between the measures of tail within the family of exponential polynomial distributions are available.

PROPOSITION 3. Suppose

$$F(x)=F_P(n, \alpha_1, \cdots, \alpha_n, x)=1-\exp\left(-\sum_{i=1}^{n}(\alpha_i x)^i\right),$$

and let m represent the minimum integer i such that $\alpha_i>0$. Then $\pi/(\sqrt{6}\cdot n)\leqq\tau(X)\leqq\pi/(\sqrt{6}\cdot m)$. Especially, the measure of tail of an exponential polynomial distribution is equal to or less than $\pi/\sqrt{6}$.

PROOF. The proof follows from Theorem 1 in [16] and Proposition 2.

3. A measure of tail at the origin of a distribution

Since a risk is assigned to be a very small value, such as 10^{-6} or

10^{-8}, an estimator of the corresponding dose level depends heavily on the behavior of the tolerance distribution at the origin. To evaluate it quantitatively, the following limit discussed in [16] is helpful.

DEFINITION 2. The measure of tail at the origin, $\tau_0(F(x))$ is defined by

$$\tau_0(F(x)) = \lim_{u \to 0} u / \{f(F^{-1}(u))F^{-1}(u)\} = \lim_{x \to 0} F(x) / \{f(x) \cdot x\} .$$

PROPOSITION 4. Suppose that a density function $f(x)$ is continuous and that the following two statements hold:

(i) $\lim F(x)/\{xf(x)\} = \alpha$.

(ii) There exist a number β and a positive number c such that $\lim_{x \to 0} F(x)/x^\beta = c$.

Then $\beta = 1/\alpha$.

PROOF. This is easily shown using the theorem of Cauchy's generalized law of the mean.

Remarks. A) The above proposition shows that the measure of tail at the origin of a distribution represents the inverse of the order of $F(x)$ in x at the origin. Definition of the measure of tail at the origin based on the statement (ii) would be more intuitively appealing than ours, Definition 2. However, our definition is consistent with the definition of the heaviness of tail represented by the function $1/\{f(F^{-1}(u))F^{-1}(u)\}$: $G(x)$ has heavier tail than $F(x)$, iff $1/\{g(G^{-1}(u))G^{-1}(u)\} \geqq 1/\{f(F^{-1}(u))F^{-1}(u)\}$ for any u $(0 < u < 1)$ ([16]).

In connection with this, it should also be noted that the conditions (i) and (ii) are slightly different in the sense that the classes of distributions satisfying these conditions are not identical. Suppose that for any small x, $F(x) = x^{1/\alpha}(-\log x)$. Then (i) holds while (ii) does not hold for any $\beta > 0$.

B) To evaluate the behavior of a distribution more precisely, we may introduce another number ζ in Definition 2. Suppose there exist positive numbers ζ and α such that

$$\lim F^\zeta(x) / \{xf(x)\} = \alpha .$$

These numbers are measures of tail at the origin. A larger values of ζ means heavier tail at the origin. It is easily shown that ζ is equal to or larger than 1 for any distribution.

The value of ζ is 2 for $\exp X$, X being a standard Cauchy random variable, which have heavier tail than any of the distributions we are considering in the present paper.

The next proposition presents explicitly the measures of tail at the origin for the families of distributions in Table 1.

PROPOSITION 5. Measures of tail at the origin are represented as follows:

(i) $\tau_0(F_N(\nu, x))=0$,
(ii) $\tau_0(F_L(\beta, x))=1/\beta$,
(iii) $\tau_0(F_W(\gamma, x))=1/\gamma$,
(iv) $\tau_0(F_G(k, x))=1/k$,
(v) $\tau_0(F_M(\lambda, x))=1/\lambda$ and
(vi) $\tau_0(F_P(n, \alpha_1, \cdots, \alpha_n, x))=1/m$,

where m is the minimum integer such that $\alpha_i>0$.

Example 1. When we take the spontaneous incidence rate into account, Abbott's correction is used, that is for a tolerance distribution function $F_0(x)$ and a spontaneous incidence rate P_0 the dose-response curve is presented by $F(x)=P_0+(1-P_0)F_0(x)$. Another correction is obtained as follows: for a spontaneous incidence rate P_0 the dose-response curve is presented by $G(x)=F(x+x_0)$, where $P_0=F(x_0)$, which is called the additive model [6]. In this case the distribution function corresponding to $F_0(x)$ in Abbott's correction is written by $G^*(X)=(F(x+x_0)-P_0)/(1-P_0)$.

Suppose that $F(x)$ is differentiable at x_0. Then the measure of tail at the origin of $G^*(x)$ is 1, which is equal to that of the exponential distribution. Numerical examples were presented in [6].

4. Tail ordering among distributions for comparisons of models

Let $\{F(\alpha, \beta, x)\}$ and $\{G(\alpha', \beta', x)\}$ be two families of tolerance distribution functions, where α and α' are scale parameters and β and β' represent tail parameters. Suppose that both tolerance distribution functions are fitted to a common data by the maximum likelihood method and that the estimated distribution functions are $F(\hat{\alpha}, \hat{\beta}, x)$ and $G(\hat{\alpha}', \hat{\beta}', x)$. Then we can guess that $\tau(F(\hat{\alpha}, \hat{\beta}, x))$ and $\tau(G(\hat{\alpha}', \hat{\beta}', x))$ are close to each other. In fact the parameters are estimated so that the estimated distributions and the empirical distribution induced by the data are close to each other. Thus values of the global measure of $F(\hat{\alpha}, \hat{\beta}, x)$ and $G(\hat{\alpha}', \hat{\beta}', x)$ are usually close to each other. And we should recall that the measure, τ is the standard deviation of the log-transformed variable by Definition 1. This guess will be examined using practical data sets in Section 6.

On the other hand, we use $F^{-1}(\hat{\alpha}, \hat{\beta}, p)$ and $G^{-1}(\hat{\alpha}', \hat{\beta}', p)$ for small values of p such as 10^{-6} or 10^{-8} as point estimators of a virtually safe doses. These values depend heavily on the behaviors of tail at the origin of both distribution functions, whose measures are $\tau_0(F(\hat{\alpha}, \hat{\beta}, x))$ and $\tau_0(G(\hat{\alpha}', \hat{\beta}', x))$.

The above speculation derives a tail ordering among distribution functions which permit us to compare models for estimation of safe doses. We again suppress the scale parameter by giving it the value one.

DEFINITION 3. Let $\{F(\beta, x)\}$ and $\{G(\beta', x)\}$ be two families of distribution functions. $\{G(\beta', x)\}$ has the larger measure of tail at the origin under a common measure of tail than $\{F(\beta, x)\}$ for an interval I, iff for any value $c \in I$, there exist β_0 and β_0' such that $c=\tau(G(\beta_0', x)) =\tau(F(\beta_0, x))$ and $\tau_0(G(\beta_0', x)) \geqq \tau_0(F(\beta_0, x))$.

From the above definition a model with $G(\beta', x)$ as the tolerance distribution function is considered to be more conservative than another model with $F(\beta, x)$, if $G(\beta', x)$ has the larger measure of tail at the origin under a common measure of tail than $F(\beta, x)$.

Next we present tail orders among distribution functions in Table 1 exactly. Proposition 5 shows that measures of tail at the origin are the inverse of parameters except for the lognormal and the exponential polynomial distributions. Since that of the lognormal distribution function is zero, it has the smallest measure of tail at the origin under a common measure of tail for $(0, \infty)$. Contrarily, it is easily shown that the exponential polynomial distribution function has the larger measure of tail at the origin under a common measure of tail than the Weibull for $(\pi/\sqrt{6}, \infty)$.

In addition the measure of tail is decreasing in each parameter as shown in Section 2. Thus the tail orders among distribution functions are transferred from those of measures of tail of distribution functions with a common parameter. These are summarized as follows.

PROPOSITION 6. Let $\{F(\beta, x)\}$ and $\{G(\beta', x)\}$ be two families of distribution functions. Suppose that $F(\beta, x)$ and $G(\beta, x)$ have $1/\beta$ as their common measure of tail at the origin and that the measure of tail of $F(\beta, x)$ (or $G(\beta, x)$) is strictly decreasing in β. Then

(i) Suppose $\tau(F(\beta, x)) > \tau(G(\beta, x))$ for any β. Then $\tau_0(F(\beta', x)) < \tau_0(G(\beta, x))$ for any β and β' such that $\tau(G(\beta', x)) = \tau(F(\beta, x))$.

(ii) Suppose there exists β_0 such that $\tau(F(\beta, x)) > \tau(G(\beta, x))$ for $\beta < \beta_0$, $\tau(F(\beta_0, x)) = \tau(G(\beta_0, x)) = c_0$ and $\tau(F(\beta, x)) < \tau(G(\beta, x))$ for $\beta > \beta_0$. Then $\tau_0(F(\beta, x)) \leqq \tau_0(G(\beta', x))$ for any β and β' such that $\tau(F(\beta, x)) = \tau(G(\beta', x)) \in (0, c_0]$, and $\tau_0(F(\beta, x)) \geqq \tau_0(G(\beta', x))$ for any β and β' such that $\tau(F(\beta, x)) = \tau(G(\beta', x)) \in [c_0, \infty)$.

Using Proposition 1 the orders of measures of tail among distribution functions are easily derived. But to determine the orders of measures of tail between the minimum exponential distribution and other distribution functions we need numerical computation, since it is impossible to represent the measure of tail with a simple form for the minimum exponential distribution. Extensive computations derive the following results: For $\beta\leqq 1$ it holds that $\tau(F_M(\beta))\leqq\min\{\tau(F_G(\beta)), \tau(F_L(\beta)), \tau(F_W(\beta))\}$, for $2.234\geqq\beta\geqq 1$ it holds that $\max\{\tau(F_W(\beta)), \tau(F_G(\beta))\}\leqq\tau(F_M(\beta)) \leqq\tau(F_L(\beta))$ and for $\beta\geqq 2.234$ it holds that $\max\{\tau(F_G(\beta)), \tau(F_L(\beta)), \tau(F_W(\beta))\} \leqq\tau(F_M(\beta))$.

If we can assume the above results analytically, then we have the following

PROPOSITION 7. (i) Suppose a common measure τ of tail of distribution functions is in $[\pi/\sqrt{6}, \infty)$. Then it holds that

$$\tau_0(F_M(\lambda, x))\geqq\tau_0(F_G(k, x))\geqq\tau_0(F_W(\gamma, x))\geqq\tau_0(F_L(\beta, x))\geqq\tau_0(F_N(\nu, x)) ,$$

where the value $\pi/\sqrt{6}$ is obtained by $\pi/\sqrt{6}=\tau(F_W(1, x))=\tau(F_G(1, x))= \tau(F_M(1, x))$.

(ii) Suppose a common measure of tail is in $[.809, \pi/\sqrt{6}]$. Then it holds that

$$\begin{aligned}\tau_0(F_P(n, \alpha_1,\cdots, \alpha_n, x))&\geqq\tau_0(F_W(\gamma, x))\geqq\tau_0(F_G(k, x))\geqq\tau_0(F_M(\lambda, x))\\ &\geqq\tau_0(F_L(\beta, x))\geqq\tau_0(F_N(\nu, x)) ,\end{aligned}$$

where the value .809 is obtained by $.809=\tau(F_M(2.243, x))=\tau(F_L(2.243, x))$.

(iii) Suppose a common measure of tail is in [.664, .809]. Then it holds that

$$\begin{aligned}\tau_0(F_P(n, \alpha_1,\cdots, \alpha_n, x))&\geqq\tau_0(F_W(\gamma, x))\geqq\tau_0(F_G(k, x))\geqq\tau_0(F_L(\beta, x))\\ &\geqq\tau_0(F_M(\lambda, x))\geqq\tau_0(F_N(\nu, x)) ,\end{aligned}$$

where the value .664 is obtained by $.664=\tau(F_G(2.730, x))=\tau(F_L(2.730, x))$.

(iv) Suppose a common measure of tail is in (0, .664]. Then it holds that

$$\begin{aligned}\tau_0(F_P(n, \alpha_1,\cdots, \alpha_n, x))&\geqq\tau_0(F_W(\gamma, x))\geqq\tau_0(F_L(\beta, x))\geqq\tau_0(F_G(k, x))\\ &\geqq\tau_0(F_M(\lambda, x))\geqq\tau_0(F_N(\nu, x)) .\end{aligned}$$

The proof follows directly from Propositions 5 and 6. In Proposition 7(i) the exponential polynomial distribution $F_P(n, \alpha_1,\cdots, \alpha_n, x)$ is is absent, since the measure τ of this distribution must be less than or equal to $\pi/\sqrt{6}$.

According to the criterion of Definition 3, the multistage model is most conservative, and the probit has the reverse property. The Weibull model is more conservative than the logit. The multihit model as well as the multitarget is conservative when the parameter is small, say less than 1, but is not conservative when the parameter is large, say greater than 3.

5. pth quantile under a common measure of tail

Though the results in the previous section are simple and clear, they are of the limiting case. When a common measure of tail is assigned exactly, the behavior of tail of a distribution can be obtained by calculating the pth quantile. Tables 3-1, 2, 3 and 4 give the pth quantiles of the distribution functions for $p=0.1$, 0.01, 10^{-3}, 10^{-6}, and 10^{-8} in the case that the common measure of tail takes values $\log_e 10$, $\pi/\sqrt{6}$, $\pi/\sqrt{30}$ and $\pi/\sqrt{300}$. Here $\tau=\log_e 10$ is attained for $\Phi(\log_{10} x)$,

Table 3-1. pth quantile of distributions whose measures of the heaviness of tail is common: Case $\tau=\log 10$

	Corresponding parameter	p 0.1	0.01	10^{-3}	10^{-6}	10^{-8}
Lognormal (ν)	0.434	.0523	$.472\times10^{-2}$	$.812\times10^{-3}$	$.176\times10^{-4}$	$.244\times10^{-5}$
Loglogistic (β)	0.788	.0615	$.293\times10^{-2}$	$.157\times10^{-3}$	$.242\times10^{-7}$	$.698\times10^{-10}$
Weibull (γ)	0.557	.0176	$.259\times10^{-3}$	$.412\times10^{-5}$	$.169\times10^{-10}$	$.434\times10^{-14}$
Gamma (k)	0.479	$.640\times10^{-2}$	$.523\times10^{-4}$	$.430\times10^{-6}$	$.238\times10^{-12}$	$.160\times10^{-16}$
Minimum Ex. (λ)	0.476	$.801\times10^{-2}$	$.637\times10^{-4}$	$.508\times10^{-6}$	$.258\times10^{-12}$	$.164\times10^{-16}$

Table 3-2. Case $\tau=\pi/\sqrt{6}$

	Corresponding parameter	p 0.1	0.01	10^{-3}	10^{-6}	10^{-8}
Lognormal (ν)	1.283	.193	$.506\times10^{-1}$	$.190\times10^{-1}$	$.225\times10^{-3}$	$.748\times10^{-3}$
Loglogistic (β)	1.414	.211	$.388\times10^{-1}$	$.757\times10^{-2}$	$.291\times10^{-4}$	$.220\times10^{-5}$
Weibull (γ)	1	.105	$.101\times10^{-1}$	$.100\times10^{-2}$	$.100\times10^{-5}$	$.100\times10^{-7}$
Gamma (k)	1			as above		
Minimum Ex. (λ)	1			as above		

Table 3-3. Case $\tau=\pi/\sqrt{30}$

	Corresponding parameter	p 0.1	0.01	10^{-3}	10^{-6}	10^{-8}
Lognormal (ν)	1.743	.479	.263	.170	$.655\times10^{-1}$	$.400\times10^{-1}$
Loglogistic (β)	3.162	.499	.234	.113	$.127\times10^{-1}$	$.295\times10^{-2}$
Weibull (γ)	2.236	.366	.128	$.455\times10^{-1}$	$.207\times10^{-2}$	$.264\times10^{-3}$
Gamma (k)	3.513	1.425	.624	.302	$.399\times10^{-1}$	$.107\times10^{-1}$
Minimum Ex. (λ)	4.211	.865	.408	.215	$.383\times10^{-1}$	$.127\times10^{-1}$

Table 3-4. Case $\tau=\pi/\sqrt{300}$

	Corresponding parameter	p				
		0.1	0.01	10^{-3}	10^{-6}	10^{-8}
Lognormal (ν)	5.513	.793	.656	.571	.422	.361
Loglogistic (β)	10.000	.803	.632	.501	.251	.158
Weibull (γ)	7.071	.727	.521	.377	.142	.0739
Gamma (k)	30.894	24.02	19.45	16.51	11.23	9.039
Minimum Ex. (λ)	295.255	4.858	4.168	3.767	3.085	2.805

which was assumed in [9] as a conservative tolerance distribution function. The measure of tail of the exponential distribution is $\pi/\sqrt{6}$.

Chand and Hoel [2] presented the pth quantiles in the case $\tau=\log_e 10$ for several families including the lognormal, the loglogistic and the Weibull distributions. Numbers corresponding to the loglogistic distribution in Table 1 in [1] should be corrected by using Table 3-1.

Characteristics of the families of distributions obtained in Section 4 are clearly noticeable in Table 3. Largest values of the 10^{-8}th quantiles relative to those of the .1th are attained by the lognormal distribution. The 10^{-8}th quantiles of the gamma and the minimum exponential distribution functions are very small in the case $\tau=\log_e 10$, but much larger than those of the Weibull distribution in the case $\tau=\pi/\sqrt{300}$.

6. Application to practical data sets

We shall apply our models to a common data set in order to examine whether or not estimated measures of tail of the tolerance distributions are close to each other and also to examine whether or not orderings of measures of tail at the origin under a common measure of tail among the tolerance distributions characterize the models for estimation of safe doses. The multitarget model is omitted here, since it is similar to the multihit.

Each of the models except the multistage model includes three parameters by adding a scale parameter and a parameter representing the spontaneous incidence rate. In each of the models, the parameters are estimated using the maximum likelihood method from each of the data sets in [11] and [12], which are given in Table 4.

The measure of tail of fitted distribution functions and the measures of tail at the origin are listed in Table 5. Here the estimated spontaneous incidence rates are disregarded.

The exponential polynomial distribution is fitted by using the program "global" by Deal [4]. The estimated parameters of the gamma distribution are recalculated. Numbers corresponding to the data sets

Table 4 [11]. Twelve sets of toxic response data

Data set No.	Substance tested	Levels of dose (unit)	No. of animals tested	No. of animals w/toxic response
1	Vinyl chloride	0 (ppm.)	58	0
		50	59	1
		250	59	4
		500	59	7
		2500	59	13
		6000	60	13
2	Methylmercury chloride	0 (mg./kg. over 3 mg./kg).	10	0
		5.4	6	1
		17.0	8	3
		43.0	7	3
		71.0	8	7
		173.0	8	6
3	Span oil	0 (% in diet)	10	1
		5	10	1
		10	10	4
		15	10	4
		20	10	5
4	DDT	0 (ppm.)	111	4
		2	105	4
		10	124	4
		50	104	13
		250	90	60
5	Dimethyl-nitrosamine	0 (ppm.)	29	0
		2	18	0
		5	62	4
		10	5	2
		20	23	15
6	Dieldrin	0.00 (ppm.)	156	17
		1.25	60	11
		2.50	58	25
		5.00	60	44
7	2, 3, 7, 8-tetra-chloro-dibenzo-p-dioxin	0 (ppm.)	24	0
		0.125	38	0
		0.25	33	1
		0.5	31	3
		1.0	10	3
8	2, 4, 5-trichloro-phenoxyacetic acid	0 (mg./kg.)	284	44
		25	140	14
		50	161	37
		100	110	32
		150	58	32
9	Benzopyrene	6 (mg./wk. painted on skin)	300	0
		12	300	4
		24	300	27
		48	300	99
10	Ethylene thiourea	0 (ppm.).	72	15
		5	75	23
		25	73	13
		125	73	16
		250	69	31
		500	70	63

Table 4. (Continued)

Data set No.	Substance tested	Levels of dose (unit)	No. of animals tested	No. of animals w/toxic response
11	Ethylene thiourea	0 (mg./kg.)	167	0
		5	132	0
		10	138	1
		20	81	14
		40	178	142
		80	24	24
12	Botulinum toxin-type A	.010 (ng. in	30	0
		.015 0.05 ml.	30	0
		.020 sodium	30	0
		.024 phosphate	30	0
		.027 buffer)	30	0
		.030	30	4
		.034	30	11
		.037	30	10
		.040	30	16
		.045	30	26
		.050	30	26

Table 5. Estimated measures of tail in five typical models for twelve data sets

Data set No.	Tolerance distribution	Estimated measure of tail	Estimated measure of tail at the origin
1	Lognormal	3.895	0
	Loglogistic	3.810	2.101
	Weibull	2.925	2.281
	Gamma	2.658	2.460
	Exp. Poly.	1.283	1
2	Lognormal	2.154	0
	Loglogistic	2.024	1.116
	Weibull	2.193	1.710
	Gamma	2.418	2.207
	Exp. Poly.	1.283	1
3	Lognormal	1.015	0
	Loglogistic	1.115	.615
	Weibull	.948	.739
	Gamma	.927	.626
	Exp. Poly.	1.003	1
4	Lognormal	.934	0
	Loglogistic	.992	.547
	Weibull	.849	.662
	Gamma	.859	.556
	Exp. Poly.	.870	1

Table 5. (Continued)

Data set No.	Tolerance distribution	Estimated measure of tail	Estimated measure of tail at the origin
	Lognormal	.718	0
	Loglogistic	.749	.413
5	Weibull	.641	.500
	Gamma	.655	.358
	Exp. Poly.	.641	.5
	Lognormal	.744	0
	Loglogistic	.815	.449
6	Weibull	.772	.601
	Gamma	.740	.438
	Exp. Poly.	.825	1
	Lognormal	.928	0
	Loglogistic	.843	.465
7	Weibull	.638	.498
	Gamma	.709	.409
	Exp. Poly.	.641	.5
	Lognormal	.788	0
	Loglogistic	.730	.403
8	Weibull	.568	.443
	Gamma	.639	.343
	Exp. Poly.	.556	.5
	Lognormal	.763	0
	Loglogistic	.712	.393
9	Weibull	.544	.424
	Gamma	.614	.320
	Exp. Poly.	.532	.5
	Lognormal	.405	0
	Loglogistic	.437	.243
10	Weibull	.464	.362
	Gamma	.419	.162
	Exp. Poly.	.465	.5
	Lognormal	.404	0
	Loglogistic	.412	.227
11	Weibull	.378	.295
	Gamma	.387	.139
	Exp. Poly.	.428	.333
	Lognormal	.194	0
	Loglogistic	.204	.112
12	Weibull	.212	.165
	Gamma	.193	.037
	Exp. Poly.	.214	.167

No. 4, 5 and 10 in [11] should be corrected by using Table 5.

To data sets No. 1 and 2 the multistage model is ill fitted and the estimated measures of tail are $\pi/\sqrt{6}$ which are much less than those in other models. Here we should recall the fact that $\tau(F_P(n, \alpha_1, \cdots, \alpha_n, x)) \leqq \pi/\sqrt{6}$. This means the model is inapplicable to these two data sets. Thus we will disregard the calculated values corresponding to these cases in the following consideration.

From Table 5, we find that, for all data sets, the estimated measures of tail of a group of the gamma, the Weibull and the exponential polynomial distributions are close to each other, and the same is seen for those of another group of the lognormal and the loglogistic distributions, where the values for the latter group are apt to be larger than the former. Especially, the differences between them seem to be a bit large in the case of data sets No. 1, 7, 8 and 9. In spite of these differences we can conclude roughly that the estimated measures of tail are stable, even though the assumed tolerance distributions are different.

Next we turn to the estimated measures of tail at the origin. As stated above the estimated measures of tail are close to each other. Thus these measures are regarded as the estimated measures of tail at the origin under a common measure of tail.

Throughout all the data sets the maximum measures are attained by the exponential polynomial distribution and the minimum by the lognormal. The estimated measures of the Weibull distribution are larger than those of the loglogistic. As for the gamma distribution the order of the estimated measures are changeable. In fact for the data sets No. 1 and 2 the estimated measures are larger than those of the Weibull distribution. For the data sets No. 3 and 4 the estimated measures are smaller than those of the Weibull distribution and larger than those for the loglogistic. For the other data sets the estimated measures are smaller than those for the loglogistic distribution. These findings correspond to the results in Proposition 7.

Acknowledgements

The authors wish to express their thanks to Dr. M. Sibuya for his suggestions and stimulating discussions and to Ms. Y. Sakamoto for her patient work in numerical computations. Their thanks extend to referees for their valuable suggestions for improving the original version.

THE INSTITUTE OF STATISTICAL MATHEMATICS
NATIONAL INSTITUTE OF ENVIRONMENTAL HEALTH SCIENCES

References

[1] Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference under Order Statistics*, John Wiley & Sons Inc., New York.

[2] Chand, N. and Hoel, D. G. (1974). A comparison of models for determining safe levels of environmental agents, *Reliability and Biometry*, *SIAM*, 681-700.

[3] Crump, K. S., Guess, H. A. and Deal, K. L. (1977). Confidence intervals and test of hypotheses concerning dose-response relations inferred from animal carcinogenicity data, *Biometrics*, **33**, 437-451.

[4] Deal, K. Global. A fortran program to analyze dichotomous animal carcinogenicity data, unpublished.

[5] Gastwirth, J. L. (1970). On robust tests, *Nonparametric Techniques in Statistical Inference* (ed. M. L. Puri), University Press, Cambridge, 89-101.

[6] Hoel, D. G. The incorporation of background in dose-response models, unpublished.

[7] Johnson, N. L. and Kotz, S. (1970). *Continuous Univariate Distribution*—1, *Distributions in Statistics*, John Wiley & Sons Inc., New York.

[8] Johnson, N. L. and Kotz, S. (1970). *Continuous Univariate Distribution*—2, *Distributions in Statistics*, John Wiley & Sons Inc., New York.

[9] Mantel, N. and Bryan, W. R. (1961). "Safety" testing of carcinogenic agents, *J. Nat. Cancer Inst.*, **27**, 455-470.

[10] Peto, R. and Lee, P. (1973). Weibull distributions for continuous carcinogenesis experiments, *Biometrics*, **29**, 457-470.

[11] Rai, K. and van Ryzin, J. (1979). Risk assessment of toxic environmental substances using a generalized multi-hit dose response model, *Energy and Health*, *SIMS Conf.*, SIAM Press, Philadelphia.

[12] Scientific Committee, Food Safety Council (1978). Proposed system for food safety assessment, (Chapter 11).

[13] Takeuchi, K. (1973). On robust estimation—II, *Jap. J. Appl. Statist.*, **2**, 69-93 (in Japanese).

[14] White, J. S. (1969). The moments of log-Weibull statistics, *Technometrics*, **11**, 373-386.

[15] Yanagimoto, T. and Sibuya, M. (1976). Isotonic tests for spread and tail, *Ann. Inst. Statist. Math.*, **28**, A, 329-342.

[16] Yanagimoto, T. and Sibuya, M. Comparison of tails of distributions in models for estimating safe doses, *Ann. Inst. Statist. Math.*, **32**, A, 325-340.