QMPE: Estimating Lognormal, Wald and Weibull RT distributions with a parameter dependent lower bound.

Andrew Heathcote

School of Behavioural Sciences, University of Newcastle, Aviation Building, University Avenue, Callaghan, 2308 NSW, Australia

Ph: 61-2-49215952, andrew.heathcote @newcastle.edu.au

Scott Brown

Department of Cognitive Sciences, University of California at Irvine, Irvine, CA 92697-5100, USA

Ph: +1 949 824 2051
scottb@uci.edu

Denis Cousineau

Département de psychologie
Université de Montréal
C. P. 6128, succ. Centre-ville
Montréal, Québec,
H3C 3J7, CANADA

Ph: +1 514 343 7981
denis.cousineau@umontreal.ca

We describe and test QMPE, an open source ANSI Fortran90 program for response time distribution estimation[1]. QMPE enables users to estimate parameters for the ex-Gaussian and Gumbel distributions, along with three "shifted" distributions (i.e., distributions with a parameter dependent lower bound), the Lognormal, Wald and Weibull distributions. Estimation can be performed using either the standard maximum likelihood (CML) method, or quantile maximum probability (QMP: Heathcote & Brown, in press). We review the properties of each distribution and theoretical evidence showing that CML estimates fail for some cases with shifted distributions, whereas QMP estimates do not. In cases where CML does not fail, a Monte Carlo investigation showed that QMP estimates were usually as good, and in some cases better, than CML estimates. However, the Monte-Carlo study also uncovered problems that can occur with both CML and QMP estimates, particularly when samples are small and skew is low, highlighting the difficulties of estimating distributions with parameter dependent lower bounds.

This paper describes and tests QMPE (quantile maximum probability estimator), an open source ANSI standard Fortran90 program for estimating the parameters (**Θ**) of continuous density functions *f*(**Θ)** commonly used to model response time (RT) data. QMPE extends Brown and Heathcote's (2003) QMLE software, which fits only the three parameter ex-Gaussian distribution, to four new positively skewed distributions: the two-parameter Gumbel distribution and the three-parameter shifted Lognormal, shifted Wald, and shifted Weibull distributions. The shifted distributions are bounded below, which makes them attractive as models of RT. Like QMLE, QMPE can fit distributions using continuous maximum likelihood (CML) estimation, and quantile maximum probability (QMP^2, Heathcote, Brown & Mewhort, 2002; Heathcote & Brown, in press) estimation. In the next section we describe the distribution functions fit by QMPE. We then describe the estimation methods, and demonstrate that in cases where CML estimation fails for shift distributions, QMP estimates are viable. Finally, we report the results of a Monte Carlo study that compares CML and QMP estimation in small samples.

QMPE Distribution Functions

Like Brown and Heathcote's (2003) QMLE program, QMPE fits the ex-Gaussian distribution, a positively skewed distribution produced by the convolution of a normal and exponential distribution (see Heathcote, 1996 for details). The ex-Gaussian has three parameters, the mean (μ) and standard deviation (σ>0) of the normal component and the mean of the exponential component (τ>0). The parameters have a simple relationship to the first three cumulants, the mean (κ_1), the variance (κ_2) and the third central moment ($\kappa_3 = \int (x-\kappa_1)^3 f(x)dx$) given by:

$$\kappa_1 = \mu + \tau \qquad \kappa_2 = \sigma^2 + \tau^2 \qquad \kappa_3 = 2\tau^3$$

The third central moment is a measure of skew, and can be estimated by the method of moments formula: $\hat{\kappa}_3 = \sum (x_i - \bar{x})^3 / n$. The "Fisher Skew" measure, $\gamma_1 = \kappa_3 / \kappa_2^{3/2}$, is also often used to quantify distribution asymmetry as it a dimensionless quantity which is invariant to scale changes. Figure 1a shows three examples of ex-Gaussian distributions.

Woodworth and Schlosberg (1954) suggested the Lognormal as an empirical approximation to RT distributions. As its name implies, the logarithm of a Lognormally distributed random variable is distributed normally (equivalently, an exponentiated normal random variable has a Lognormal distribution). The Lognormal distribution is the asymptotic distribution of the product of random variables (McClelland's, 1979, cascade model is an example of a multiplicative model, see Ulrich & Miller, 1994, for more details). Hence, the Lognomal distribution can be motivated as an approximation to the finishing time of a series of stages with randomly varying rates (West & Shlesinger, 1990). Breukelen (1995) shows that two well know parallel models are also compatible with the Lognormal distribution.

The Lognormal distribution is bounded below by zero (x>0) and has parameters corresponding to the normal distribution mean (μ) and standard deviation (σ>0). In order to allow for a lower bound greater than zero, we added a shift parameter, θ>0. The corresponding density, which is only defined for θ<x, is:

$$f(x) = \frac{1}{(x-\theta)\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x-\theta)-\mu}{\sigma}\right)^2}$$

Ratcliff and Murdock (1976) reported that the shifted Lognormal provided as good a fit as the ex-Gaussian to RT distribution data from recognition memory experiments.

Figure 1b shows three examples of the Lognormal density. In order to specify the first three cumulants it is convenient to define $\omega = \exp[\sigma^2]$:

$$\kappa_1 = \theta + e^{\mu + \frac{1}{2}\sigma^2} \qquad \kappa_2 = e^{2\mu}\omega(\omega - 1) \qquad \kappa_3 = e^{3\mu}\omega^{\frac{3}{2}}(\omega - 1)^2(\omega + 2)$$

The Wald distribution (Wald, 1947) can be motivated as a model of RT by a continuous approximation to the sequential acquisition of information. Suppose that at each time step identical, normally distributed observations with mean $\mu>0$ are accumulated. A decision to respond is made when the sum exceeds some criterion, $a>0$. Given $\mu<<a$, the number of steps to exceed the criterion is approximately Wald distributed. As described by McGill (1963), in the limit the discrete steps can be replaced by a continuous time variable, resulting in a diffusion process with an exactly Wald distributed stopping time.

The Wald distribution is bounded below by zero ($x>0$). We added a shift parameter, $\theta>0$, to allow for a lower bound greater than zero. The corresponding density, which is only defined for $\theta<x$, is:

$$f(x) = \frac{a}{\sqrt{2\pi(x-\theta)^3}} \exp\left[-\frac{(a - \mu(x-\theta))^2}{2(x-\theta)}\right]$$

Figure 1c shows three examples of the Wald density. The first three cumulants are:

$\kappa_1 = \theta + a/\mu$ $\qquad$ $\kappa_2 = a/\mu^3$ $\qquad$ $\kappa_3 = 3a/\mu^5$

The Wald distribution is also used with a different parameterisation in terms of its (un-shifted) mean, a/μ, and its “dispersion”, $\lambda = a^2$. In this form the distribution is often called the “Inverse Gaussian” distribution[3]. When we tested this parameterisation with QMPE, we found that parameter estimates, particularly for λ, were very biased and inefficient. Hence, QMPE uses the “diffusion” parameterisation, which also has the advantage of parameters being directly interpretable in terms of the information accumulation decision model.

The final two distributions implemented in QMPE, the Weibull and Gumbel, are related in that they both occur as the asymptotic distribution of the minima of samples from sets of random variables (see Weibull, 1951 and Cousineau, Goodman & Shiffrin, 2002, for details): the Weibull arises from the minimum value of samples from random variables that are bounded below by zero, the Gumbel from random variables that are unbounded. Both distributions have two parameters but differ in that the Gumbel distribution is unbounded, whereas the Weibull distribution is bounded below by zero. As for the other distributions that are bounded below, QMPE adds a shift parameter, θ>0, to the Weibull.

The Weibull distribution is a power transformation (with exponent *c*>0) of an exponential random variable (with mean τ>0) as is evident from its density:

$$f(x) = c\tau^{-c}\left(x-\theta\right)^{c-1} e^{-((x-\theta)/\tau)^{c}}$$

As for the other shift distributions, the density is only defined for θ<*x*. Figure 1d shows three examples of the Weibull density. The cumulants of the Weibull are expressed in terms of the incomplete Gamma function ($\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$, $x > 0$)

$$\kappa_1 = \theta + \tau\Gamma\left(c^{-1}+1\right)$$
$$\kappa_2 = \tau^2\left[\Gamma\left(2c^{-1}+1\right) - \Gamma^2\left(c^{-1}+1\right)\right]$$
$$\kappa_3 = \tau^3\left[\Gamma\left(3c^{-1}+1\right) - 3\Gamma\left(c^{-1}+1\right)\Gamma\left(2c^{-1}+1\right) + 2\left[\Gamma\left(c^{-1}+1\right)\right]^3\right]$$

The Gumbel density is also named the Type I extreme value, the double-exponential or the Fisher-Tippett distribution. It has only two parameters, one for location (μ) and one for scale (σ>0).

$$f(x) = \frac{1}{\sigma}\exp\left[\frac{-(x-\mu)}{\sigma} - \exp\left[\frac{-(x-\mu)}{\sigma}\right]\right]$$

Figure 1e shows three examples of the Gumbel density. Its first three cumulants are:

$\kappa_1 = \mu + 0.57722\sigma$, $\kappa_2 = (\sigma\pi)^2/6$, and $\kappa_3 = 2.40412\sigma^3$

The Fisher Skew of the Gumbel is fixed (γ_1=1.13955) so does not have the flexibility to model changes in RT distribution skew. Although the Gumbel cannot be a general model of RT distribution, it was included in QMPE as it might have utility in special cases. Although we briefly report estimation results for the Gumbel our main focus is on estimation of shift distributions.

Likelihood and Shifted Distributions

The likelihood of a sample given a model is the joint probability of the data assuming the model. When the observations making up the data (x_i, i=1…n) are sampled independently, the joint probability is given by the product of the probabilities for each observation. Maximum likelihood methods choose model parameter estimates that maximize the likelihood. However, when applied to continuous distributions, the conventional approach to maximum likelihood estimation maximizes the product of the densities for each observation, $f(x_i,\mathbf{\Theta})$, rather than the product of their probabilities. This procedure is justified using an approximation to the probability of each observation:

$$\Pr(x_i - h_i/2 \le X \le x_i + h_i/2) = \int_{x_i-h_i/2}^{x_i+h_i/2} f(x,\Theta)dx \approx f(x,\Theta)h_i$$

Two assumptions must hold to ensure that maximising the product of the densities ($\prod_{i=1}^{n} f(x_i,\mathbf{\Theta})$) is equivalent to maximising the joint probability. First, the h_i (>0) must be small for the approximation to be accurate, and the approximation can be made to be exact as the h_i tend to zero. Second, the h_i must be independent of Θ, so that the h_i can be ignored in the maximization. We will refer to this approximation as CML. For computational reasons, estimates are usually found by maximising the log-likelihood, that is the sum of the logarithms of the densities, $\sum_{i=1}^{n} \ln f\left(x_i, \Theta\right)$, which is equivalent to maximising their product.

Unfortunately the assumptions underlying the CML approximation can fail in some cases, such as when the distribution's range depends on its parameters, as is the case for shift distributions. When the shift parameter (θ) equals the smallest observation, CML log-likelihood is infinite as the density for the smallest observation x_1 is zero (without loss of generality we assume the observations are ordered, $x_1 \le x_2 \le \ldots \le x_n$). In this case, CML estimates of the other parameters become inconsistent, in the sense that they do not tend to their true values as sample size increases. The singularity associated with $\theta = x_1$ is not a problem for iterative estimation methods if 1) log-likelihood has a local maximum (say at $\theta_L < x_1$) which yields consistent parameter estimates and 2) if the singularity is disconnected (i.e., log-likelihood decreases as θ approaches x_1 from below on the interval $\theta_L \le x < x_1$). The singularity has been shown to be disconnected for the shifted Wald distribution (Cheng & Amin, 1981).

For the shifted Lognormal and Weibull distributions the singularity is not always disconnected, so inconsistent estimates can be obtained when maximising the log-likelihood by iterative methods. This problem, which we will call the "unbounded likelihood problem", occurs particularly for parameter values where the distribution is highly skewed and has a sharply increasing leading edge. For the Lognormal distribution the unbounded likelihood problem is rarely a practical concern, as the increasing region is usually very small (cf. Giesbrecht & Kempthorne, 1976, Table 1), and a local maximum that produces consistent parameter estimates usually exists outside this region. Given good starting point estimates, iterative methods will converge on this local maximum and provide consistent parameter estimates, which are sometimes called “local maximum likelihood” estimates.

For the Weibull distribution, however, the unbounded likelihood problem can cause more severe difficulties when estimating the highly skewed distributions produced by small values of the Weibull shape parameter, c. For $c=1$, the Weibull distribution is equivalent to the exponential distribution, which has a sharp and discontinuous leading edge. For $c<1$ even more skewed distributions with a shape similar to the exponential are obtained. For $c>1$ the distribution becomes less skewed and the increase of the leading edge is more gradual. Symmetry occurs for $c\approx3.6$, and skew then becomes negative as c increases, approaching a lower bound Fisher skew≈-1.14. For $c>2$ for the singularity is disconnected. For $c<2$ the singularity is connected, but when $c>1$ a local maximum usually exists which produces consistent parameter estimates (Cheng & Amin, 1983). When $c<1$ no local minimum exists, so iterative estimation results in a shift estimate $\theta=x_1$ and inconsistent estimates of the other parameters. Similar conditions apply to the shifted Gamma, another distribution commonly used to model RT which also has the exponential distribution as a special case, for exactly the same values of its shape parameter (Cheng & Amin). Although not described here, a newer version of QMPE also fits the shifted Gamma distribution, and its parameter estimates behave similarly to those of the shifted Weibull.

Cheng and Amin (1983), and independently Ranneby (1984), suggested a solution to the unbounded likelihood problem, called the maximum product of spacings (MPS) method. MPS, like CML, obtains parameter estimates by maximising a goodness of fit (objective) function. The MPS objective function is proportional to a special case of the QMP objective function, and so produces identical estimates. As defined by Heathcote et al. (2002), QMP estimates are obtained by maximising the multinomial log-likelihood:

$$\sum_{j=1}^{m} N_j \ln\left(D_j\right), \text{ where } D_j = \int_{\hat{q}_{j-1}}^{\hat{q}_j} f\left(x_i, \Theta\right) dt \qquad (1)$$

The $\hat{q}_j$, j=1…m-1 are quantile estimates, ($\hat{q}_o, \hat{q}_m$) equals the domain of the distribution (which might depend on **Θ**), and each inter-quantile range ($\hat{q}_{j-1}, \hat{q}_j$) contains N_j observations (in general N_j may not be an integer).

The MPS estimator is a special case of QMP[4] where order statistics (i.e., x_i) are used to estimate quantiles, N_j = 1 and $m = n$. Titterington (1985) suggested a modified version of the MPS objective function that is proportional to the QMP1 objective function examined by Heathcote et al. (2002). By QMP1, we mean estimates obtained by maximising (1) and based on $\hat{q}_j = \left(x_j + x_{j+1}\right)/2$, $j = 1 \ldots n$-1, and N_j = 1. Heathcote et al. showed that QMP1 produced more efficient and less biased estimates than CML for the ex-Gaussian distribution. They also examined QMP4 estimates, where the data set is reduced to a set of (n/4)-1 equally spaced quantile estimates (for $n = 4m$, $\hat{q}_j = \left(x_{4j} + x_{4j+1}\right)/2$ and N_j = 4), and found similar estimation performance to CML, despite the fact that QMP4 is clearly not a sufficient estimator (i.e., it does not use all of the information contained in the data set).

Although Titterington (1985) suggested that MPS can be viewed as maximum likelihood estimation for grouped data, it is important to acknowledge that the equivalence is only approximate in finite samples, because (1) does not take into account the error associated with quantile estimates. However, when the range of the distribution is not parameter dependent, MPS and CML are asymptotically equal, as shown by Cheng and Amin (1983), so in this case MPS has all of the asymptotic sufficiency, consistency and efficiency properties of CML. When the range of the distribution is parameter dependent CML and MPS can behave quite differently. Importantly, both the original version of MPS and Titterington's variation (i.e., QMP1) differ from CML in that they are not subject to the unbounded likelihood problem (Cheng & Iles, 1987). Hence, they continue to give consistent and efficient estimates even when CML completely fails. In fact, these estimates can become "super-efficient", in the sense that estimation error decreases as sample size (n) increases at a rate faster than $n^{-1/2}$.

In summary, it is clear that the unbounded likelihood problem can cause CML to completely fail in cases where QMP continues to work well. It might be argued that such cases are of little interest for RT distribution fitting, as RT distributions rarely have a sufficiently sharp leading edge or degree of skew. We are not aware of any systematic investigation on this point, and caution that sampling error may cause the problem to occur in small samples even if the true distribution comes from a parameter region where CML does not fail. In any case, there is little point comparing the estimation performance of CML and QMP in such cases, as QMP will necessarily be superior. The parameters for the Monte Carlo study were chosen to avoid distributions associated with CML failure.

Monte Carlo Study

The Monte Carlo study was modelled after the study reported by Heathcote et al. (2002). It had three aims: 1) to extensively test the QMPE code, 2) to compare the estimation performance of CML and QMP, and 3) to compare estimation performance among the five distributions fit by QMPE. Relatively small sample sizes were used (n=40, 80 or 160) in order to investigate performance under demanding and realistic conditions. QMP estimation was performed both using QMP1 and QMP4.

For each distribution three sets of parameter values, given in Table 1, were used, (Figure 1 illustrates the corresponding densities). The parameters for the ex-Gaussian distribution were the three sets with medium levels of skew used by Heathcote et al. (2002). The choice of parameters for the other distributions was guided by fitting them to large samples from the three ex-Gaussian distributions, so that results are approximately comparable across distributions. As Table 1 shows, this procedure resulted in a fairly good match on means and standard deviations. Fisher Skew varied more between distribution types but covered approximately the same range, except for the Weibull where skew was generally lower, and the Gumbel distribution, where skew is fixed. The smallest value of the Weibull shape parameter investigated (c=1.5) was large enough to avoid CML failure due to the unbounded likelihood problem even in the smallest samples.

Examination of Figure 1 indicates that shift estimation in the least skewed cases (labelled 1 in the figure) of the Lognormal and Wald distributions will be challenging, because they have long thin left tails. Such tails make estimation of the shift parameters difficult because samples near the lower bound are rare, and the sampled values of the first order statistic (x_1) are highly variable. As a result, shift estimates are likely to be biased upward and to be more variable for these cases in the Monte Carlo study, particularly in small samples.

Methods

For each distribution type and the nine combinations of sample size and parameter set, 10000 replicates were fit, with the same samples fit by CML, QMP1 and QMP4. The simulated samples were obtained using random number generators provided by the S-plus statistical package[5], and were rounded to the nearest integer.

QMPE uses the same numerical methods as QMLE (see Brown & Heathcote, 2003, for more details). CML and QMP estimates are obtained by a conjugate gradient optimisation algorithm. This algorithm requires analytic expressions for the gradient of the objective function. However, QMPE requires only analytic gradients for the density; gradients for CML and QMP are automatically computed from the density gradients. Once search is complete, analytic expressions for the Hessian (second derivative matrix) of the density are used to estimate approximate parameter standard errors and correlations (see Brown & Heathcote, 2003, for a proof that these estimates are asymptotically correct for QMP). Although derivative free optimisation methods are available, we have found that analytic gradients greatly speed estimation and that analytic Hessians result in better standard error and correlation estimates.

QMPE automatically obtains starting points for optimisation by substituting method of moments' estimates of cumulants into the equations relating cumulants and parameters. However, this approach fails when sample estimates of skew are negative. In such cases heuristics are used to estimate starting points. For the three distributions with a shift parameter, the heuristic estimates the shift as slightly smaller than the minimum value in the sample (e.g., $\hat{\theta} = p \times x_1$, where $0<p<1$ is an appropriately chosen constant), then solves for the other parameters using the first two moments calculated on $\mathbf{x} - \hat{\theta}$. The heuristic is always used for the Lognormal, which we found rarely works with the full method of moments approach. As it has only two parameters, Gumbel start points are obtained from only the first two cumulants. Users can also supply their own starting points.

Good automatic starting point estimates are essential when large numbers of conditions must be fit, and particularly for the shifted distributions when only the local CML solution is useful. QMPE's start point heuristics were fine tuned throughout the course of the Monte Carlo study. We have also found them to work well in real RT data.

The stopping criteria for optimisation were set at a proportional objective function exit tolerance of 10^{-9}, a proportional parameter change tolerance of 10^{-5}, and the maximum number of search iterations was fixed at 250 (see Brown & Heathcote, 2003, for details on these settings), resulting in parameter estimates accurate to more than four significant figures. For all parameters bounded below by zero QMPE sets the objective function to a low value when the estimate is less than 10^{-9}, which ensures both that the bound is respected and that numerical errors do not occur. For distributions with shift parameters, these parameters were also restricted to less than the sample minimum for CML fits and less than the minimum quantile for QMP fits.

Results

Estimation Failures

Only 50 fits out of the 1.35 million performed failed to produce usable parameter estimates, indicating that the starting point heuristics are robust. As shown in Table 2, estimation of parameter standard errors and correlations failed at a greater (but still low) rate, because the Hessian was not invertible (i.e., not positive-definite). For brevity, Table 2 averages over parameter sets and sample sizes, and omits results for the Gumbel, which never failed. Generally, better performance was obtained with larger samples and for more skewed distributions. QMP4 estimation for the Lognormal stands out as producing many more failures than the other cases, indicating that, if parameter standard error and correlation estimates are required for the Lognormal, higher levels of grouping should be avoided.

Bias and Efficiency

Bias was estimated as the difference between the mean of the Monte Carlo parameter estimates and the true value, with positive values indicating over estimation and negative values indicating underestimation. Efficiency was estimated by the standard deviation (SD) of the parameter estimates. Estimates are described as "consistent" if the magnitude of bias decreases and efficiency increases as sample size increases. Tables 3-5 contain the bias and efficiency estimates for the shift distributions' parameters.

In order to compare estimation performance across distributions, it is useful to recognise that the shift parameters (θ), and the Weibull scale parameter (τ), have the same units as the data. For these parameters relative estimation performance can be judged on the same scale. Estimation performance can also be judged for all parameters as a proportion of their true values, which are given in Table 1.

The results for the ex-Gaussian distribution replicated Heathcote et al. (2002) with only a slightly different methodology (i.e., rounded samples from a different random number generator), and are omitted for brevity, as is a detailed discussion of the results for the Gumbel distribution, which were uniformly good in all cases and for all estimation methods[6]. In contrast to the Gumbel and ex-Gaussian estimates, estimates for the shift distributions were very poorly behaved in some cases. That is, estimates were very biased, not always consistent (i.e., bias could increase and efficiency decrease with sample size) and the parameter estimate distributions were not even approximately normal.

The Lognormal estimates were best behaved amongst the shift distributions in terms of consistency and distributions. Parameter estimate distributions were mainly uni-modal, with the exception of CML estimates for the least skewed distribution at the smallest sample size, which had small second modes overestimating θ and underestimating μ. For all estimation methods, the parameter estimate distributions were slightly skewed, particularly for small sample sizes, to the left for θ and to the right for μ and σ.

As shown in Table 3, bias was generally upward for θ and σ and downward for μ. Bias in all parameters was substantial for the least skewed distribution, even for the largest sample, and small for the other two distributions at all sample sizes. The least skewed distribution has a small shift parameter (475) and a long left tail. As is evident from Figure 1, sampled values less than 750 are rare, resulting in a strong upward bias in shift estimates even for larger sample sizes. Estimation efficiency and bias were consistent in all cases, except for a few cases occurring when bias was negligible, likely due to Monte Carlo error. CML was generally the least biased, and QMP1 the most efficient.

For the Wald distribution, parameter estimate distributions were generally uni-modal but could also be heavy tailed. All parameter estimate distributions contained extreme underestimates for shift and overestimates for μ and a, but the main body of the distribution tended to be skewed in the opposite direction, particularly for smaller samples and for CML. As shown by bias values in Table 4, the shift parameter of the least skewed Wald distribution was generally overestimated, whereas for the most skewed distribution it was generally underestimated. For the other parameters the opposite pattern generally applied, underestimation for the least skewed and overestimation for the most skewed distribution. Bias was particularly pronounced in CML estimates for the least skewed distribution. This bias was not due to extreme outliers; the same pattern occurred when the central tendency of the parameter estimate distribution was estimated by its median. Bias was consistent for all but CML estimates for the most skewed distribution and QMP4 estimates for the least skewed distribution, but the inconsistency was relatively small. Efficiency estimates were consistent and CML estimates were clearly more efficient than QMP estimates. However, CML was more biased than QMP1 estimates, particularly for the least skewed distribution. Hence, CML parameter estimate distributions are less variable, but tend to be centred on a biased estimate of the true parameter value.

For the Weibull distribution, bias was substantially smaller for the more skewed Weibull distributions. For these cases, QMP estimates, particularly QMP1 estimates, were less biased than CML estimates. For the least skewed distribution CML estimates were substantially less biased than QMP estimates. However, even for CML bias was substantial for the smallest sample size. For all distributions CML estimates were the most efficient, although the advantage over QMP1 was relatively small for the more skewed distributions, particularly at larger sample sizes. QMP4 was clearly the least efficient method, particularly for smaller sample sizes.

Weibull parameter estimate distributions were almost always bimodal to some degree for the least skewed case, although the second mode tended to disappear as sample size increased. The second mode always underestimated shift (θ) and overestimated the scale (τ) and shape (c) parameters. Estimates fell in the second mode mainly for samples with negative Fisher Skew. The Weibull distribution can have negative Fisher Skew, which slowly approaches a bound of approximately –1.14 for large values of *c*. For example, Fisher Skew values are -0.08, -0.6, -1, -1.1 and -1.13 for *c* = 4, 10, 100, 1000 and 10000 respectively. These results indicate that caution should be exercised when fitting the Weibull to samples with negative Fisher Skew.

Estimates in the deviant mode usually had non-invertible Hessians, indicating that the neighbourhood of the solution is not locally quadratic. When estimates with non-positive-definite Hessians were censored, bias was reduced although not eliminated. Hence, it appears that censoring estimates with ill conditioned Hessians can improve overall estimation performance for the Weibull. An alternative approach is to bound the estimate of *c* during estimation, a strategy that can be easily implemented by modifying and re-compiling the QMPE open-source code. We obtained improved bias and efficiency using an upper bound of 10, but this resulted in even more clearly bimodal parameter estimate distributions with estimates "piling up" against the bound.

Estimating Fisher Skew

Ratcliff and Murdock (1976) used CML to fit the ex-Gaussian distribution in order to estimate RT distribution skew. Ratcliff (1978) pointed out that that estimates of skew based on the method of moments are both inefficient and non-robust. Hence, unrealistically large sample sizes are required for precise estimates and estimates can be greatly distorted by even small levels of outlier contamination. In this section we compare the indirect method of calculating Fisher Skew from CML and QMP1 parameter estimates with direct estimates obtained from the method of moments.

Skew estimates for all three-parameter distributions (i.e., those with variable skew) are shown in Figure 2. Results are expressed as a percentage of the true Fisher Skew value, and absolute values of bias are given in order to make comparison easier.

For the ex-Gaussian distribution both CML and QMP1 estimates were less biased and more efficient than method of moments estimates, with QMP1 having clearly less bias than CML for the least skewed distribution. Efficiency was comparable for CML and QMP1, and clearly better than method of moments except for the least skewed distribution and smallest sample size. For the Lognormal, the method of moments estimates were much less biased than CML and QMP1 for the least skewed distribution, but more biased for the more skewed distributions. Generally, the methods of moments estimates are less efficient than CML and QMP1, except for the smallest sample size and more skewed distributions.

For the Wald distribution, the method of moments estimates were least efficient and CML estimates most efficient. However, CML estimates were the most biased, particularly for the least skewed distribution, with QMP1 the least biased overall. For the least skewed Weibull distribution the method of moments were the least biased and most efficient, although all estimation methods displayed poor efficiency due to bimodal parameter estimate distributions. Efficiency problems may have been exaggerated by the use of a percentage measure, as Fisher Skew was smaller for the Weibull than for other distributions in the least skewed case. Estimation performance was much better for the more skewed distributions with QMP1 the best overall.

The results presented in Figure 2 provide a basis for comparing estimation performance among distributions with variable skew. Clearly CML and QMP1 fits of the ex-Gaussian distribution provide better skew estimates than the method of moments at all levels of skew. QMP1 and CML are also generally superior for the two more skewed cases of the other distributions. For the least skewed case, however, the method of moments generally outperforms QML1 and CML, reflecting the estimation difficulties for this case noted in the last section.

Discussion

The results of the Monte Carlo study confirmed that QMP is generally superior to CML for the two distributions with an unbounded range, the ex-Gaussian and Gumbel. CML and QMP are on a more equal footing for the shift distributions when Fisher Skew is greater than one, although it should be remembered that for more extreme skew CML can fail entirely for the Lognormal and Weibull distributions, due to the unbounded likelihood problem.

No method worked very well for the least skewed Lognormal, Wald and Weibull distributions, particularly for smaller sample sizes. Overestimation of shift for the Lognormal and Wald distributions results from their long thin left tails in the least skewed cases. Overestimation in these cases may be difficult to avoid because information about the shift value is very variable. Hence, QMPE parameter estimates for these distributions should be interpreted with caution when Fisher Skew is less than one.

Underestimation of shift for the Weibull appears to be related to a non-quadratic maximum for both CML and QMP, as indicated by ill-conditioned Hessian estimates. Removal of cases with ill-conditioned Hessians improved performance. Heathcote (in press) noted similar behaviour for CML estimates of Wald distributions in small samples (n = 40), and suggested that estimates with ill-conditioned Hessians should be censored when using his software. However, censoring QMPE Wald estimates with ill-conditioned Hessians did not reduce bias appreciably. In contrast, when Heathcote’s optimisation methods (the Splus *nlminb* algorithm using analytic first and second derivatives) were applied to the Wald data from the present Monte Carlo study, bias was almost eliminated by censoring. Unfortunately reduced bias was bought at the cost of reduced efficiency relative to the QMPE estimates.

Cheng and Iles (1990) provide an explanation for these difficulties, called the "embedded models" problem. They showed that each of the shift distributions fit by QMPE has a special case as shift approaches $-\infty$, which they call an "embedded" distribution. The embedded distributions (normal for all but the Weibull, which has an embedded left skewed Gumbel distribution) have only two parameters, a scale parameter and a location parameter. When the embedded model fits as well as the shift model it indicates that the sample does not contain sufficient information to estimate shift or shape; instead only location and scale can be reliably estimated. Because the embedded model occurs at an infinite parameter value, iterative estimation of the shift model is difficult, and estimates for all of its parameters become unreliable.

The bimodality and underestimation of shift seen in the QMPE Weibull parameter estimates appears to be due to the embedded model problem. Similar problems found here for some QMP4 estimates produced by QMPE and by Heathcote (in press) for CML estimates of the shifted Wald appear to be examples of the embedded model problem. Heathcote's Wald estimates were probably more prone to this problem than QMPE Wald estimates because his fitting routine used analytic Hessians, and so was more sensitive to the non-quadratic maximum produced by the embedded model problem.

General Discussion

In this paper we have described and tested QMPE, an open source ANSI standard Fortran 90 program, which can estimate the parameters of five continuous density functions commonly used to model RT. QMPE can fit these distributions using continuous maximum likelihood (CML), perhaps the most widely used and recommended method of estimating RT distribution (Heathcote, 1996; Van Zandt, 2000). It can also fit these distributions using Heathcote et al.'s (2002) quantile maximum probability (QMP) method.

A Monte Carlo study replicated Heathcote et al.'s (2002) finding that QMP produces less biased and more efficient parameter estimates than CML for the ex-Gaussian distribution. QMP was also found to produce less biased, but also slightly less efficient, parameter estimates than CML for the Gumbel distribution. Overall estimation performance was excellent for both distributions, as might be expected in the idealized situation represented by the Monte Carlo study; fitting the true data generating model to uncontaminated data. Of course, real RT data does not conform to this ideal, but at least the excellent performance under ideal conditions is reassuring for the practical application of QMPE.

The ex-Gaussian and Gumbel distributions have an unbounded range. This might be seen as a disadvantage when they are used to model RT data, because RT data must be bounded below by a positive value. QMPE also fits three "shift" distributions, which have a positive, parameter dependent lower bound, and so seem more promising as models of RT. Unfortunately, estimation performance for the shift distributions was much worse than for the ex-Guassian or Gumbel. One reason for fitting parametric distributions to RT data is to obtain a more reliable estimate of skew than is provided by the methods of moments (Ratcliff, 1978). If QMPE is used for this purpose we suggest that the ex-Gaussian be fit (the Gumbel has fixed skew and so is not useful for this purpose). At least when the ex-Gaussian is an accurate approximation to the data, this approach is more efficient and less biased than the method of moments. The shift distributions, in contrast, did not consistently outperform the method of moments, and for less skewed distributions and small sample sizes they could produce substantially worse skew estimates.

One reason why the ex-Gaussian outperforms the shift distributions might be the parameterisations used by QMPE, and most packages aimed at fitting RT distribution (see Cousineau, Brown & Heathcote, submitted). The ex-Gaussian has its least skewed (Gaussian) form when its exponential parameter approaches zero. The shift distributions have their least skewed form when the shift parameter approaches $-\infty$. In small samples, the least skewed case may often be most appropriate because the data contain mainly information about location and scale. Because this case occurs as the shift parameter diverges, parameter estimates for the shift distributions become unreliable, whereas this does not happen for the ex-Gaussian.

Cheng and Iles (1990) suggested solving this problem, which they dubbed the "embedded model problem", by using a parameterisation of the shift distributions where the least skewed case occurs at a zero rather than infinite value of one parameter (see their Table 1). We have implemented and are tested Cheng and Iles parameterisation in a new version of the QMPE software. However, little improvement in estimation performance was obtained, even after censoring cases with negative shift estimates, or cases where the full model did not fit significantly better than the embedded model. We suggest that users of the existing version of QMPE, and similar software, exercise caution in interpreting parameter estimates from small samples, particularly those with negative skew or Hessians that are not invertible. We are presently investigating the use of hierarchical models to address these difficult cases. Hierarchical models implement "soft bounding" on parameter estimates by assuming that parameters are drawn from a population distribution with an assumed form (see Rouder, Lu, Speckman, Sun & Jiang, in press, for a Bayesian approach to hierarchical Weibull estimation).

The embedded model problem occurs for data with low skew. Highly skewed data can also cause a problem, called the "unbounded likelihood problem", for CML estimation of the shifted Weibull and Lognormal distributions. The problem occurs because the likelihood maximum occurs when the shift parameter equals the minimum observed value. Although this might seem like a plausible estimate of shift, estimates of the remaining parameters are inconsistent. QMP does not suffer from the unbounded likelihood problem and so remains useful for highly skewed data. Hence, when skew is high, we recommend QMP fitting over CML fitting. QMP might also be useful in other contexts, such as fitting mixtures (e.g., Dolan et al., 2002), as they can also be subject to the unbounded likelihood problem (Cheng & Traylor, 1995).

In the course of this investigation we discovered that QMP has a special case called the maximum of product spacings (MPS), which was advocated by Cheng and Amin (1983) and Ranneby (1984) as a means of overcoming the unbounded likelihood problem. Their work proves that MPS has all of the desirable asymptotic properties of CML when CML estimates exist, and continues to work well when CML fails. Further consideration of the MPS and its generalizations (e.g., Ekstrom, 2001) is beyond the scope of the present work. However, this literature places QMP on a firm theoretical footing, not just as an approximation to likelihood, also as a goodness of fit measure that can be derived from information theory (see also Speckman & Rouder, in press; Heathcote & Brown, in press).

References

Brown, S., & Heathcote, A. (2003). QMLE: Fast, robust and efficient estimation of distribution functions based on quantiles. Behavior Research Methods, Instruments and Computers, 35, 485-492.

Breukelen, G. J. P. (1995). Parallel processing models compatible with Lognormally distributed processing times. Journal of Mathematical Psychology, 39, 396-399.

Cheng, R. C. H. & Amin, N. A. K. (1981). Maximum likelihood estimation of parameters in the Inverse Gaussian distribution, with unknown origin. Technometrics, 23, 257-263.

Cheng, R. C. H. & Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. Journal of the Royal Statistical Society, Series B, 45, 394-403.

Cheng, R. C. H. & Iles, T. C. (1987). Corrected maximum likelihood in non-regular problems. Journal of the Royal Statistical Society, Series B, 49, 95-101.

Cheng, R. C. H. & Iles, T. C. (1990). Embedded models in three-parameter distribution estimation. Journal of the Royal Statistical Society, Series B, 52, 135-149.

Cheng, R. C. H. & Traylor, L. (1995). Non-regular maximum likelihood problems. Journal of the Royal Statistical Society, Series B, 57, 3-44.

Colonius, H. (1995). The instance theory of automaticity: Why the Weibull? Psychological Review, 102, 744-750.

Cousineau, D., Brown, S. & Heathcote, A. (submitted). Extending statistics of extremes to distributions varying on position and scale, and implication for race models. Behavior Research Methods, Instruments and Computers

Cousineau, D., Goodman, V. & Shiffrin, R. M. (2002). Extending statistics of extremes to distributions varying on position and scale, and implication for race models. Journal of Mathematical Psychology, 46, 431-454.

Cousineau, D. & Larochelle, S. (1997). PASTIS: A Program for Curve and Distribution Analyses. Behavior Research Methods, Instruments, & Computers, 29: 542-548.

Dolan, C. V., van der Maas, H. L. J. & Molenaar, P. C. M. (2002). A framework for ML estimation of parameters (mixtures of) common reaction time distributions given optional truncation or censoring. Behavior Research Methods, Instruments, & Computers.

Ekstrom, M. (2001). Consistency of generalized maximum spacing estimates. Scandinavian Journal of Statistics, 28, 343-354.

Giesbrecht, F. & Kempthorne, O. (1976). Maximum likelihood estimation in the three-parameter Lognormal distribution, Journal of the Royal Statistical Society, Series B, 38, 257-264.

Gumbel, E. J. (1958). The Statistics of Extremes. New York: Columbia University Press.

Heathcote, A. (in press). Fitting the Wald and Ex-Wald distributions to response time data. Behavioural Research Methods, Instruments, & Computers

Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. Behavioural Research Methods, Instruments, & Computers, 28, 427-445.

Heathcote, A., & Brown, S. (in press). Reply to Speckman and Rouder: A theoretical basis for QML. Psychonomic Bulletin and Review.

Heathcote, A., Brown, S. & Mewhort, D.J.K. (2002). Quantile Maximum Likelihood Estimation of Response Time Distributions. Psychonomic Bulletin and Review, 9, 394-401.

Hyndman, R.J. & Fan, Y. (1996). Sample quantiles in statistical packages. The American Statistician, 50(4) 361-365.

McGill, W. J. (1963). Stochastic latency mechanisms. In R.D. Luce, R. R. Bush, & E. Galanter (Eds.), Handbook of Mathematical Psychology, 193-199.

McClelland, J. L. (1979). On the time relations of mental processes: A framework for analyzing processes in cascade. Psychological Review, 86, 287-330.

Ranneby, B. (1984). The maximum spacing method: an estimation method related to the maximum likelihood method. Scandinavian Journal of Statistics, 11, 93-112.

Ratcliff, R. (1978). A theory of memory retrieval. Psychological Review, 85, 59-108.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. Psychological Review, 83, 190-214.

Rouder, J. N., Lu, J., Speckman, P., Sun, D. & Jiang, Y. (in press). A hierarchical model for estimating response time distributions. Psychonomic Bulletin and Review.

Speckman, P. L. & Rouder, J. N. (in press). A comment on Heathcote, Brown and Mewhort's QMLE estimation method for response time distributions. Psychonomic Bulletin & Review.

Titterington, D. M. (1985). Comment on "Estimating parameters in continuous univariate distributions", Journal of the Royal Statistical Society, Series B, 47, 115-116.

Ulrich, R., & Miller, J. (1994). Effects of outlier exclusion on reaction time analysis. Journal of Experimental Psychology: General, 123, 34-80.

Wald, A. (1947). Sequential analysis. New York: John Wiley and sons.

Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of Applied Mechanic, 18, 292-297.

West, J. & Shlesinger, M. (1990). The noise in natural phenomena. American Scientist, 78, 40-45.

Woodworth, R. S., & Schlosberg, H. (1954). Experimental Psychology, Holt, New York.

Van Zandt, T. (2000). How to fit a response time distribution. Psychonomic Bulletin and Review, 7, 424-465.

Footnotes

---

[1]Source code, Linux and Windows binaries, a manual and sample instruction and data files can be obtained from either of the second and first author's websites: http://oz.ss.uci.edu/ and http://www.newcastle.edu.au/school/behav-sci/ncl/. These supplementary materials have also been accepted for the Psychonomic Society Norms, Stimuli, and Data Archive, available after August 1, 2004, at http://www.psychonomic.org.

[2]Heathcote et al. (2000) described their method as "quantile maximum likelihood" (QML). Speckman and Rouder (in press) pointed out that that QML is not maximum likelihood, and so Heathcote and Brown (in press) renamed the method "quantile maximum probability" (QMP).

[3]This parameterisation is convenient when shift does not have to be estimated, as analytic maximum likelihood estimates are available for both the mean (the arithmetic mean) and $1/\hat{\lambda} = \sum_{i=1}^{n}\left(x_i^{-1} - \hat{\kappa}_1^{-1}\right)/n$. When shift is estimated, iterative methods are required, but computational cost can be reduced using "profile likelihood" (a line search on the shift parameter with other parameters estimated analytically).

[4] We reserve the term QMP for the general procedure that maximizes (1) based on a set of quantile estimates obtained in an unspecified manner. Many different quantile estimators are available (see Hyndman & Fan, 1996, for a review). Although these estimators are asymptotically equivalent they differ in finite samples and so can result in differing parameter estimates in practice. In the QMPE software we implemented the quantile estimator specified in Heathcote et al. (2002), which produces QMP1 and QMP4 estimates as special cases and corresponds in general to Hyndman and Fan's definition 5. Although we have found this estimator works well, QMPE can be used with other quantile estimators by providing the estimated quantile values rather than raw data as input to the program.

[5]By default S-plus has functions to generate samples from all distributions fit by QMPE except the Wald. Wald random deviates were obtained using an S-plus function that uses the Inverse Gaussian parameterisation (available at http://www.statsci.org/s/invgauss.s). Heathcote (in press) provides an S-plus Wald random number generator using the diffusion parameterisation.

[6] All estimation methods were consistent for the Gumbel, with negligible bias and good efficiency (SD<15) for all sample sizes, estimation methods and parameters, with a slight improvement for the less variable distributions. QMP was less biased than CML, but in contrast to findings with the ex-

Gaussian distribution, QMP4 was less biased than QMP1. CML and QMP1 were almost equal in efficiency with QMP4 being slightly less efficient in some cases. The results indicate that all three methods should be useful in practice with samples as small as 40 observations and perhaps less.

Tables

Table 1. Parameters of distributions used in the Monte Carlo study, and associated moment statistics.

	Set	μ	σ	τ	Mean	SD	γ_1
Ex-Gaussian	1	929.289	70.711	70.711	1000	100.000	0.7071
	2	910.557	44.721	89.443	1000	100.000	1.4311
	3	905.132	31.623	94.868	1000	100.000	1.7076
		θ	σ	μ	Mean	SD	γ_1
Lognormal	1	470	0.18	6.25	996.47	95.538	0.5504
	2	745	0.36	5.45	993.34	92.379	1.1674
	3	800	0.48	5.15	993.49	98.488	1.6590
		θ	μ	a	Mean	SD	γ_1
Wald[a]	1	625	0.1886	70.711	1000	102.698	0.8216
	2	725	0.1626	44.721	1000	101.973	1.1124
	3	800	0.1414	28.284	1000	100.000	1.5000
		θ	τ	c	Mean	SD	γ_1
Weibull	1	700	315	3.2	982.13	96.772	0.1064
	2	800	220	2.0	994.97	101.915	0.6311
	3	840	170	1.5	993.47	104.200	1.0720
		μ	σ	-	Mean	SD	γ_1
Gumbel	1	955	85	-	1004.06	109.017	1.1396
	2	955	74	-	997.71	94.909	1.1396
	3	955	68	-	994.25	87.213	1.1396

[a] The samples were generated using the Inverse Gaussian parameterization, with (mean,λ) = (375,5000), (275,2000) and (200,800) for sets 1 to 3 respectively.

Table 2. Percentages of fits with invertible Hessian estimates.

	Ex-Gaussian	Lognormal	Wald	Weibull
CML	98.4	100.0	93.4	98.7
QMP1	98.8	95.2	98.9	96.2
QMP4	96.0	59.0	99.7	94.8

Table 3. Bias (Monte Carlo mean – true value) and efficiency (SD) for Lognormal parameter estimates

			Distribution 1			Distribution 2			Distribution 3		
			40	80	160	40	80	160	40	80	160
BIAS	θ	CML	131.1	128.8	103.1	-2.4	-1.8	-3.4	-5.1	-1.3	-1.5
		QMP1	140.4	124.2	117.3	-4.4	-2.3	-1.4	-9.0	-5.6	-5.0
		QMP4	171.3	142.4	118.8	-3.5	-3.5	-2.3	-16.9	-8.7	-5.1
	σ	CML	0.082	0.071	0.051	0.028	0.016	0.005	0.020	0.013	0.002
		QMP1	0.091	0.072	0.063	0.029	0.015	0.007	0.015	0.004	-0.004
		QMP4	0.126	0.087	0.066	0.054	0.020	0.008	0.029	0.006	-0.002
	μ	CML	-0.355	-0.320	-0.242	-0.053	-0.027	-0.006	-0.028	-0.020	-0.003
		QMP1	-0.373	-0.310	-0.283	-0.037	-0.019	-0.008	0.003	0.011	0.018
		QMP4	-0.479	-0.366	-0.292	-0.075	-0.025	-0.009	0.003	0.015	0.015
SD	θ	CML	120.8	84.8	70.4	75.7	57.6	45.0	54.9	36.4	22.4
		QMP1	108.9	91.3	78.6	73.8	52.7	35.9	52.4	34.7	23.9
		QMP4	116.5	99.3	87.5	91.0	62.2	41.6	75.5	45.1	27.5
	σ	CML	0.091	0.058	0.038	0.135	0.096	0.069	0.156	0.107	0.072
		QMP1	0.088	0.061	0.045	0.128	0.089	0.061	0.150	0.104	0.074
		QMP4	0.128	0.073	0.054	0.182	0.108	0.069	0.212	0.128	0.082
	μ	CML	0.333	0.228	0.168	0.346	0.254	0.191	0.322	0.221	0.146
		QMP1	0.313	0.239	0.193	0.328	0.234	0.162	0.309	0.213	0.152
		QMP4	0.381	0.272	0.223	0.424	0.277	0.186	0.426	0.267	0.171

Table 4. Bias (Monte Carlo mean – true value) and efficiency (SD) for Wald parameter estimates

			Distribution 1			Distribution 2			Distribution 3		
			40	80	160	40	80	160	40	80	160
BIAS	θ	CML	107.2	95.4	87.6	28.6	21.4	16.2	-19.9	-24.7	-29.3
		QMP1	15.0	7.3	3.4	-15.3	-10.9	-6.7	-21.8	-18.4	-17.3
		QMP4	-8.8	11.0	11.0	-42.8	-14.1	-6.9	-40.8	-21.3	-15.5
	μ	CML	-0.039	-0.036	-0.034	-0.009	-0.008	-0.006	0.017	0.017	0.018
		QMP1	-0.013	-0.009	-0.004	0.003	0.003	0.003	0.012	0.010	0.009
		QMP4	-0.006	-0.009	-0.007	0.010	0.003	0.002	0.018	0.011	0.009
	a	CML	-30.64	-28.17	-26.18	-6.93	-5.46	-4.17	6.53	7.26	8.21
		QMP1	-5.45	-2.96	-1.14	5.28	3.80	2.62	6.85	5.27	4.67
		QMP4	6.08	-2.84	-3.19	15.62	4.81	2.58	13.04	6.23	4.37
SD	θ	CML	31.9	26.2	21.7	27.8	20.8	15.0	19.4	13.1	9.9
		QMP1	104.8	86.4	65.7	83.1	59.0	42.9	51.6	29.8	21.1
		QMP4	167.0	101.7	75.6	144.3	74.0	48.5	94.6	42.6	24.9
	μ	CML	0.018	0.012	0.009	0.020	0.013	0.010	0.022	0.015	0.011
		QMP1	0.030	0.024	0.020	0.033	0.025	0.020	0.033	0.022	0.016
		QMP4	0.048	0.031	0.022	0.048	0.029	0.021	0.044	0.026	0.017
	a	CML	7.16	5.18	4.24	7.29	5.00	3.78	5.86	3.86	2.90
		QMP1	28.93	25.00	19.12	23.96	17.12	12.83	15.39	8.85	6.05
		QMP4	53.45	30.81	22.46	46.85	21.92	14.15	31.21	12.92	6.95

Table 5. Bias (Monte Carlo mean – true value) and efficiency (SD) for Weibull parameter estimates

			Distribution 1			Distribution 2			Distribution 3		
			40	80	160	40	80	160	40	80	160
BIAS	θ	CML	-25.8	-9.2	-3.1	4.7	5.2	4.0	6.7	4.4	2.7
		QMP1	-90.6	-46.6	-17.5	-3.6	0.2	1.2	2.2	1.8	1.0
		QMP4	-95.9	-95.0	-44.5	-28.1	-3.8	1.9	-5.0	1.8	1.6
	τ	CML	24.4	8.3	2.6	-6.9	-6.8	-5.0	-9.6	-6.6	-3.9
		QMP1	91.2	47.3	17.9	0.4	-2.0	-2.9	-7.1	-6.2	-5.0
		QMP4	94.9	96.5	45.3	25.5	2.0	-3.5	1.8	-4.6	-4.1
	c	CML	0.47	0.17	0.06	0.01	-0.03	-0.04	-0.06	-0.05	-0.03
		QMP1	1.15	0.58	0.21	0.07	0.01	-0.02	-0.02	-0.03	-0.02
		QMP4	1.25	1.14	0.52	0.40	0.06	-0.02	0.10	-0.02	-0.02
SD	θ	CML	160.6	93.9	47.0	44.9	20.5	11.5	14.4	7.1	4.1
		QMP1	251.7	168.7	78.0	68.5	22.7	13.1	22.2	8.0	4.6
		QMP4	299.8	258.7	171.6	166.8	63.4	20.8	73.3	13.7	6.7
	τ	CML	167.0	98.8	50.7	53.9	27.7	16.6	26.4	16.3	10.9
		QMP1	262.1	175.7	81.6	76.7	30.5	19.3	32.1	17.4	11.7
		QMP4	308.9	268.1	177.9	173.8	68.8	26.0	78.8	21.1	12.4
	c	CML	2.24	1.25	0.62	0.70	0.34	0.20	0.33	0.18	0.11
		QMP1	3.24	2.11	0.97	0.97	0.36	0.22	0.39	0.18	0.12
		QMP4	3.86	3.14	2.02	2.22	0.84	0.30	1.10	0.25	0.14

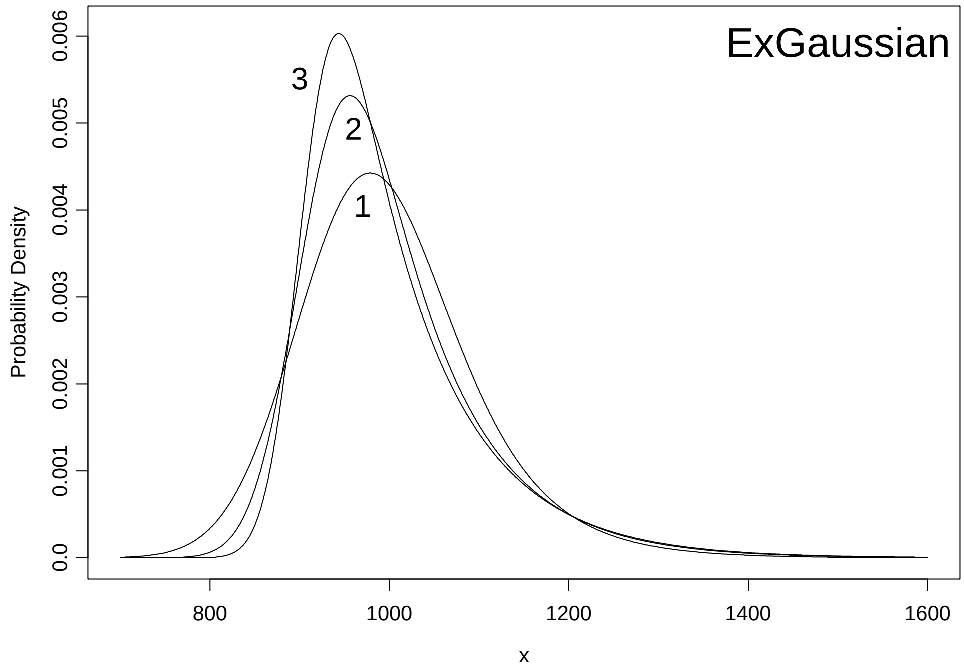
Acknowledgements

Thanks to Professor Doug Mewhort for providing computing resources and to reviewers for their helpful suggestions. This work was supported by Australian Research Council Grants to S. Andrews and A. Heathcote and to A. Heathcote, B. Hayes and D.J.K. Mewhort.

Figure Captions

Figure 1: Distributions used in the Monte Carlo study. Table 1 gives the parameters corresponding to the distributions marked 1, 2 and 3.

Figure 2. Mean absolute bias and efficiency (SD) for Fisher Skew estimates, as a percentage of the true value, estimated using the method of moments, CML and QMP1.



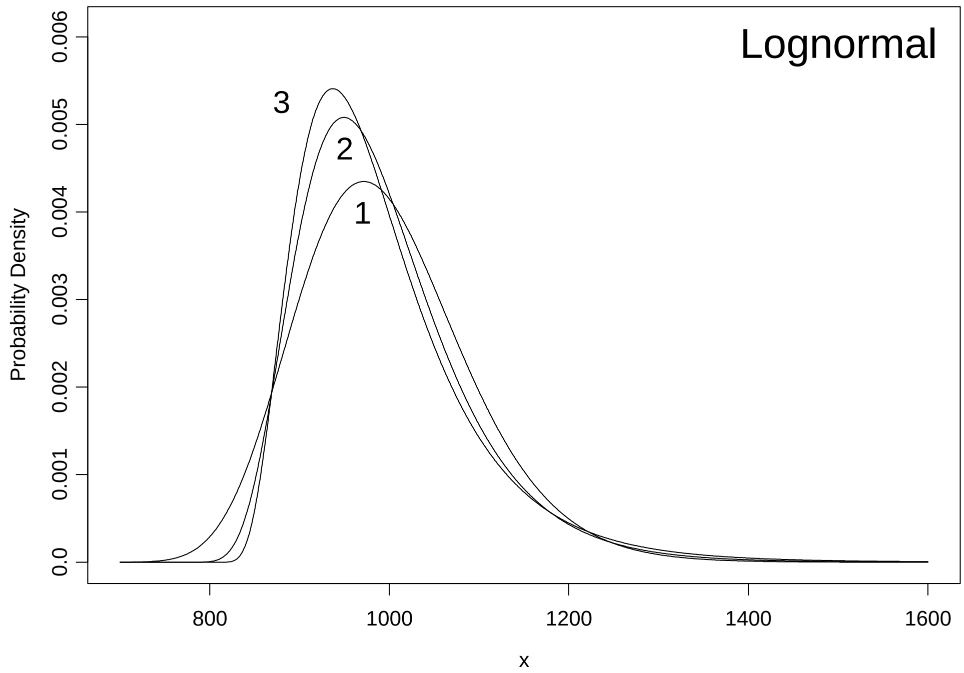


Figure 1 (continues)

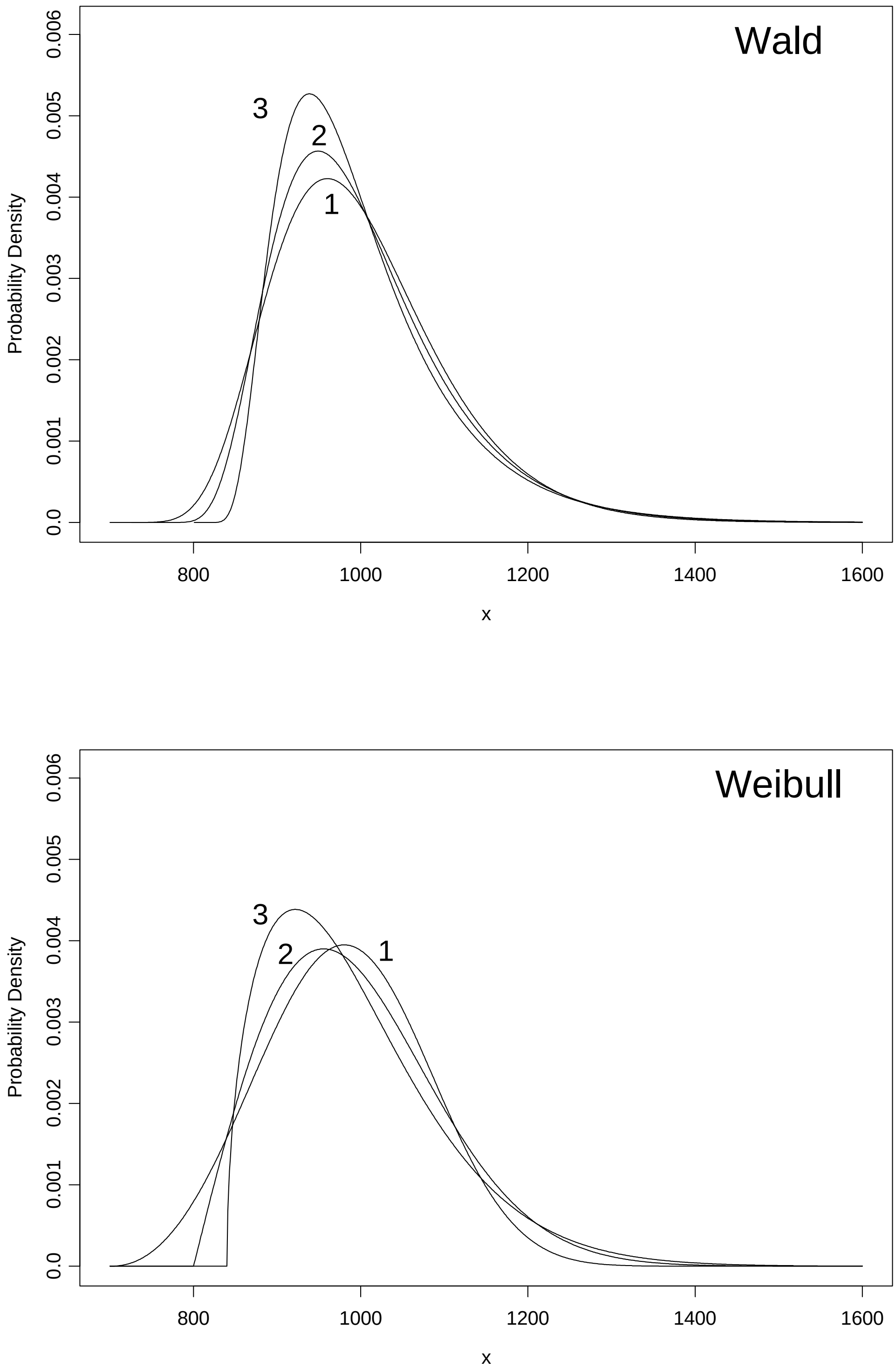


Figure 1 (continued)

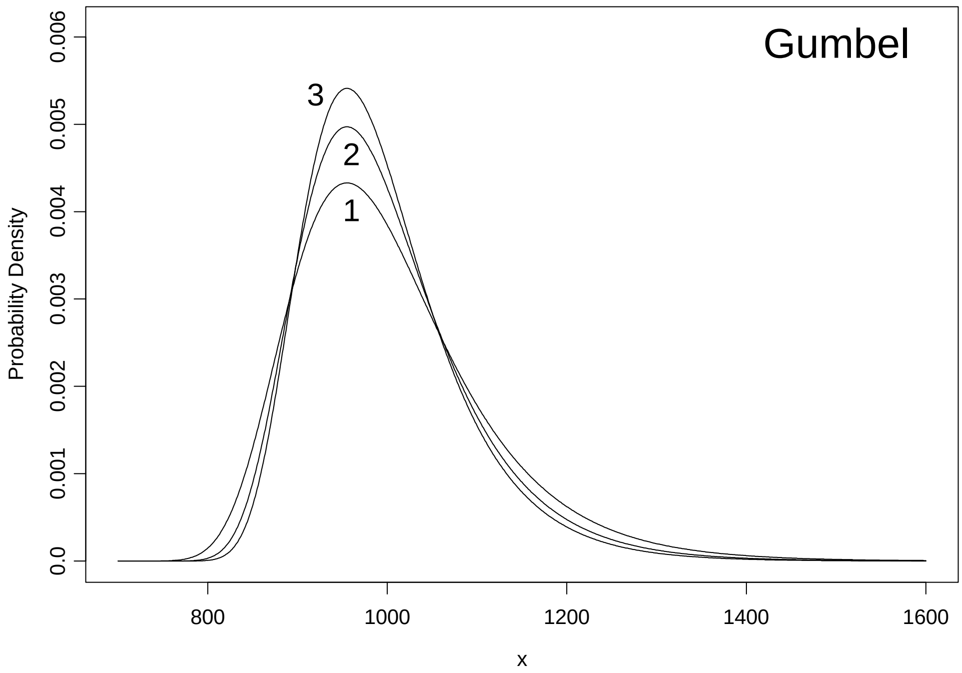
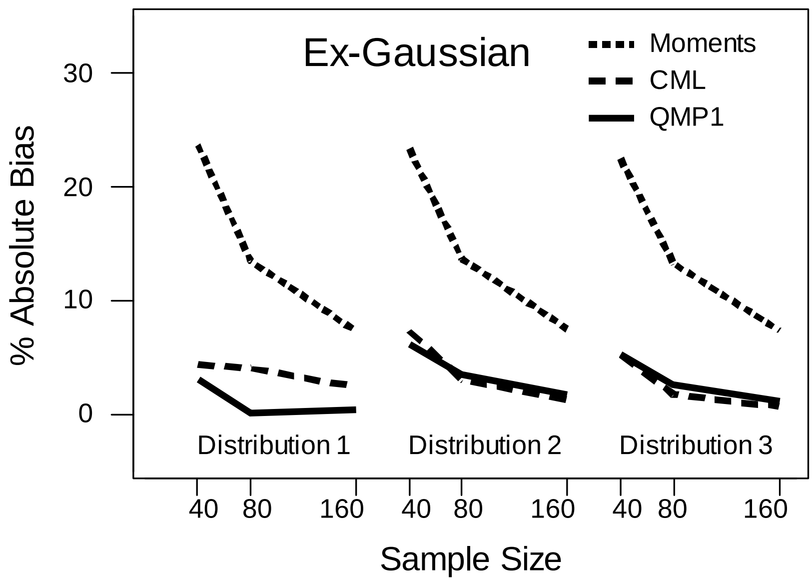

Gumbel
3
2
1
Probability Density
0.0
0.001
0.002
0.003
0.004
0.005
0.006
800
1000
1200
1400
1600
x

Figure 1



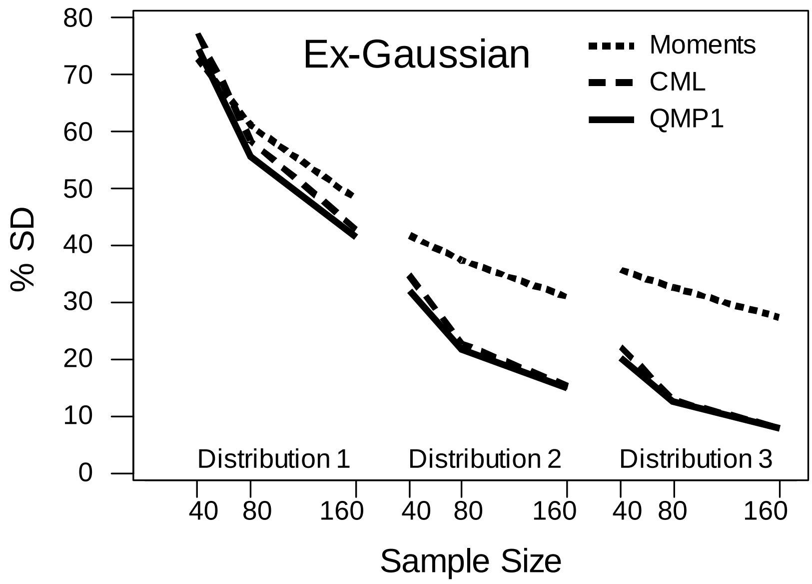
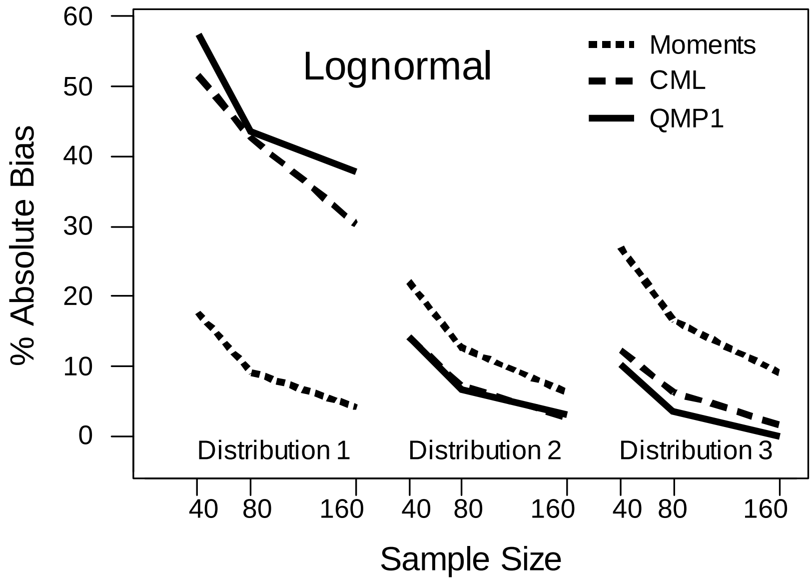


Figure 2 (continues)



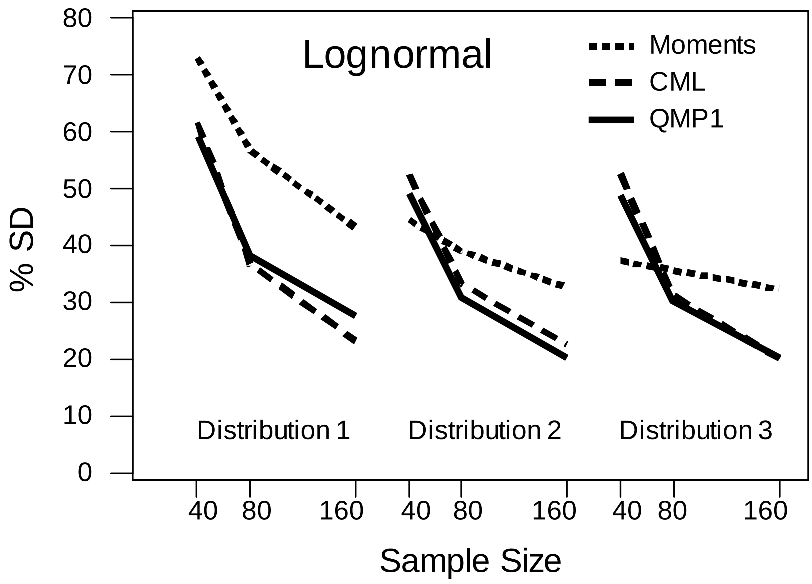
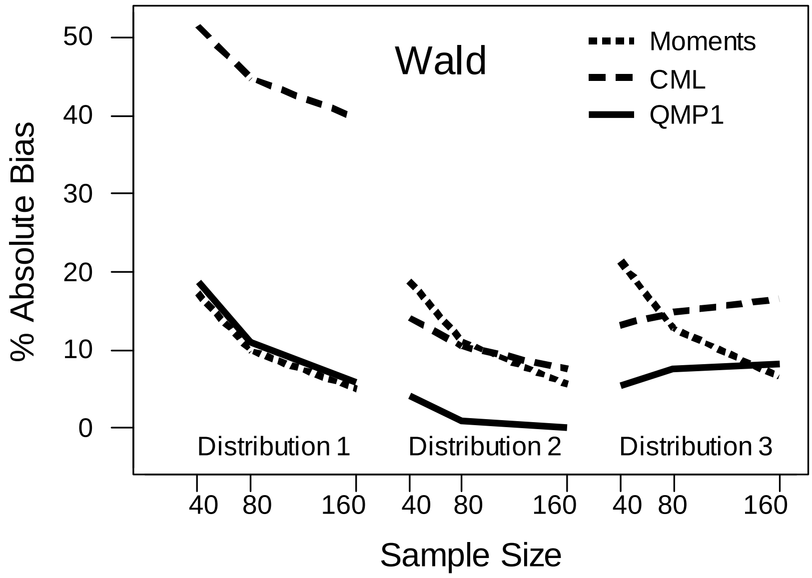


Figure 2 (Continued)



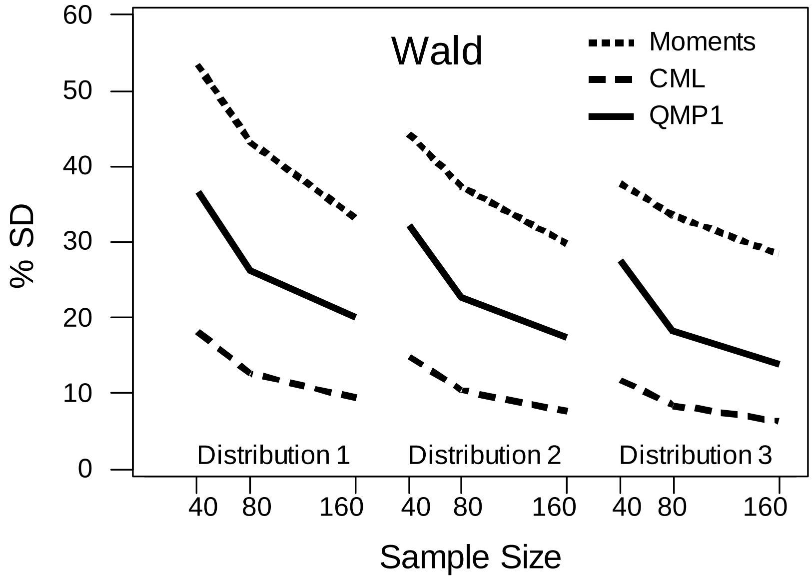
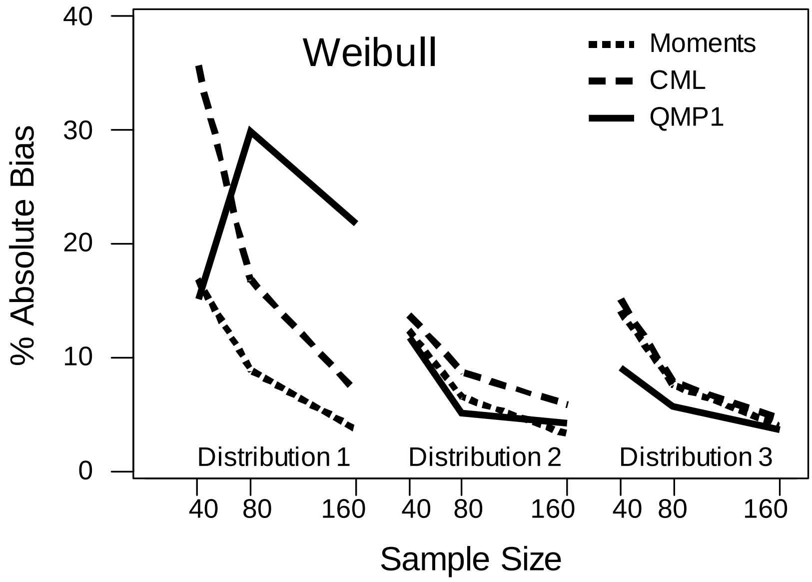


Figure 2 (Continued)


Weibull
Moments
CML
QMP1
% Absolute Bias
40
30
20
10
0
Distribution 1
Distribution 2
Distribution 3
40 80 160 40 80 160 40 80 160
Sample Size

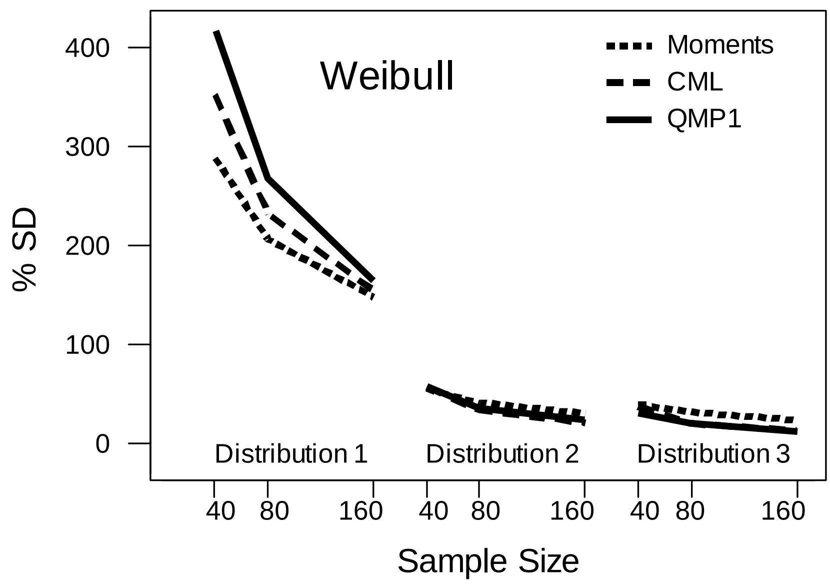

Weibull
Moments
CML
QMP1
% SD
400
300
200
100
0
Distribution 1
Distribution 2
Distribution 3
40 80 160 40 80 160 40 80 160
Sample Size

Figure 2.