1 Introduction

The minimum divergence estimator presented by Basu et al. in 1999 is a family of M-estimators indexed by a parameter ($\alpha \geq 0$) that is believed to control the trade-off between efficiency and robustness. When $\alpha = 0$ the estimator is the familiar maximum likelihood estimator and as α increases the efficiency decreases as robustness increases. However, the authors only provide rough guidelines for selecting an appropriate α value. This paper investigates the selection of α for estimating a location parameter. The model is the Normal distribution with a mean of μ and a standard deviation of σ, $\Phi(\mu, \sigma)$. The true distribution is assumed to lie within a neighborhood of the model. The set of distributions under consideration is

$$\aleph_\rho = \{F : F = (1-\rho)\Phi(\mu,\sigma) + \rho\Phi(\mu+\delta, \varepsilon\sigma), 0 \leq \rho \leq 1, -\infty < \delta < \infty, \varepsilon \geq 0\},$$

the familiar two component normal mixture model. The second component is the contamination and ρ is the amount of contamination. Several methods of selecting α will be considered and their relative merits will be compared. Well, actually there is only one at the present time. Hopefully, more are scheduled for production :-)

2 Asymptotic Minimax Bias (AMB)

To define the bias we use the theory of influence functions for M-estimators. However, the influence function is calculated using a member of $\aleph_\rho$ instead of using $F = (1-\rho)\Phi(\mu,\sigma) + \rho\Delta_x$, where Δ_x is the probability measure which puts mass 1 at the point x. Therefore, the bias (β) is defined as

$$\beta := \frac{\delta\,(1+\alpha)\sqrt{\dfrac{1+\alpha}{1+\alpha\,\varepsilon^2}}\,e^{\left(-1/2\,\frac{\alpha\,\delta^2}{\sigma^2\,(1+\alpha\,\varepsilon^2)}\right)}}{1+\alpha\,\varepsilon^2}.$$

As a prelude to a mean square error analysis, results will be derived from β^2.

$$\beta^2 := \frac{\delta^2\,(1+\alpha)^3\,e^{\left(-\frac{\alpha\,\delta^2}{\sigma^2\,(1+\alpha\,\varepsilon^2)}\right)}}{(1+\alpha\,\varepsilon^2)^3}$$

To obtain the minimax solution, we must first find the maximum β^2 for each α. Considering μ and σ to be fixed unknown quantities, The gradient of β^2 with respect to δ and ε is

$$\nabla := \left[2\,\frac{\delta\,(1+\alpha)^3\,\%1}{(1+\alpha\,\varepsilon^2)^3} - 2\,\frac{\delta^3\,(1+\alpha)^3\,\alpha\,\%1}{(1+\alpha\,\varepsilon^2)^4\,\sigma^2},\right.$$
$$\left. -6\,\frac{\delta^2\,(1+\alpha)^3\,\%1\,\alpha\,\varepsilon}{(1+\alpha\,\varepsilon^2)^4} + 2\,\frac{\delta^4\,(1+\alpha)^3\,\alpha^2\,\varepsilon\,\%1}{(1+\alpha\,\varepsilon^2)^5\,\sigma^2}\right]$$
$$\%1 := e^{\left(-\frac{\alpha\,\delta^2}{\sigma^2\,(1+\alpha\,\varepsilon^2)}\right)}.$$

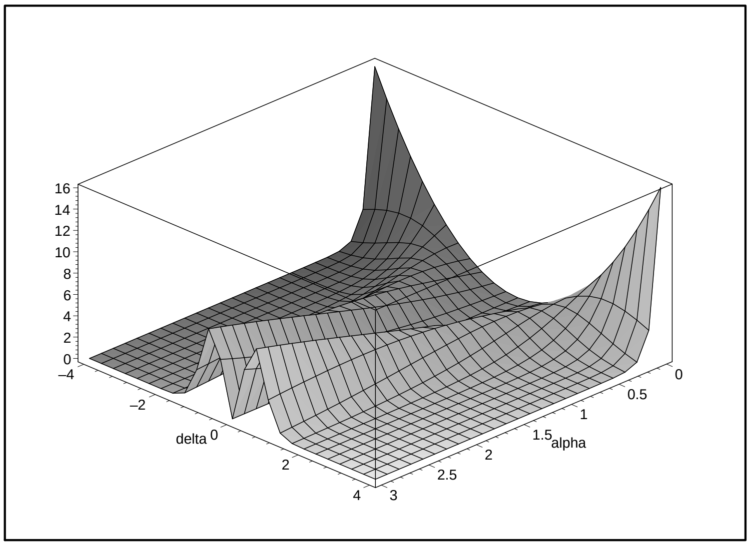

Figure 1: Plot of β^2 setting $\sigma = 1$, $\varepsilon = 0$, $0 \leq \alpha \leq 150$, $-4 \leq \delta \leq 4$

Solving for δ and ε, subject to the constraint $\varepsilon \geq 0$, the three solutions are

$$\{0 \leq \varepsilon,\, \delta = 0\},\ \{\delta = \pm\frac{1}{\sqrt{\alpha}}\sigma,\, \varepsilon = 0\}.$$

The first solution is of no consequence since β^2 is 0 for all α when $\delta = 0$. Since β^2 is symmetrical, the solution with positive δ will be examined. To verify the conditions for a local maximum, take the Hessian of β^2 with respect to δ and ε and substite $\delta = \frac{1}{\sqrt{\alpha}}\sigma$ and $\varepsilon = 0$. The evaluated Hessian is

$$\nabla^2 := \begin{bmatrix} -4\,(1+\alpha)^3\, e^{(-1)} & 0 \\ 0 & -4\,\sigma^2\,(1+\alpha)^3\, e^{(-1)} \end{bmatrix}.$$

Inspection of ∇^2 and of the determinant of ∇^2 shows the solution is a local maximum. Therefore, we have the maximum β^2 for each value of α. Substituting $\delta = \frac{1}{\sqrt{\alpha}}\sigma$ and $\varepsilon = 0$ into β^2, the maximum for each α is

$$\beta^2_{max} := \frac{\sigma^2\,(1+\alpha)^3\, e^{(-1)}}{\alpha}.$$

The last step is to find the α that minimizes β^2_{max}. The derivative of β^2_{max} taken with respect to α is

$$\frac{\partial}{\partial\alpha}\beta^2_{max} := \frac{\sigma^2\,(1+\alpha)^2\, e^{(-1)}\,(-1+2\,\alpha)}{\alpha^2}$$

Solving for α, yields the solutions of -1 and $\frac{1}{2}$. Since $0 \leq \alpha < \infty$, only the second solution will be considered. Taking the second derivative of β^2, $\alpha = \frac{1}{2}$ is found to be the minimum and this is the minimax solution. For the enjoyment of the reader, a plot of β^2 is provided in Figure 1. To confirm the theoretical results,

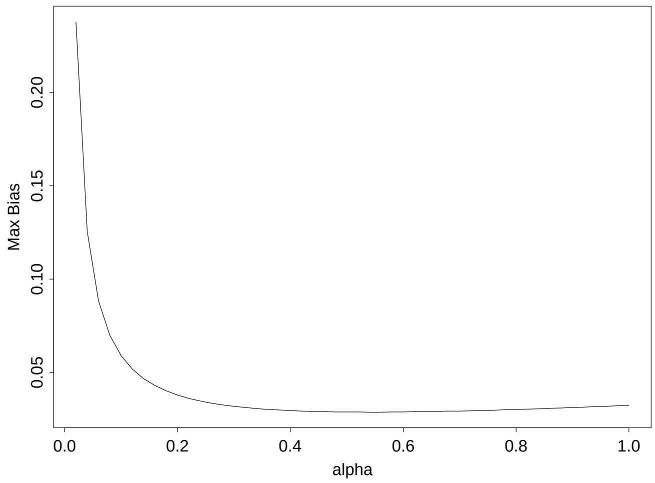

Figure 2: Plot of the simulated maximum bias for $0 \leq \alpha \leq 1$

a simulation was performed. The simulated results are displayed in Figure 2, plotting the maximum bias as a function of α. The data is simulated with a sample size of 1000 and true distribution is $0.90\Phi(0,1)+0.10\Delta_{1/\sqrt{(a)}}$. In Figure 2, the maximum bias is decreasing until $\alpha = 0.5$ and then increases slightly until $\alpha = 1$. The increase is almost negligible and hopefully simultaneous examination of the variance will provide more insight.

Since $\varepsilon = 0$ corresponds to a Δ_x contamination, restricting ε to be strictly greater than 0 yields a true normal mixture model. The AMB for this case is derived in the next section.

3 AMB with $\varepsilon > 0$

The results in this section are similar to the previous section. However, the derivatives are taken only with respect to δ. Taking the derivative of β^2 with respect to δ and solving for δ, the solutions are

$$\delta := \frac{\sqrt{\alpha\,(1+\alpha\,\varepsilon^2)}\,\sigma}{\alpha},\ -\frac{\sqrt{\alpha\,(1+\alpha\,\varepsilon^2)}\,\sigma}{\alpha},\ 0.$$

For the same reasons as above, only the positive solution will be considered. Substituting, $\delta = \frac{\sqrt{\alpha\,(1+\alpha\,\varepsilon^2)}\,\sigma}{\alpha}$ into β^2, taking the derivative with respect to α, and solving for α yields the solutions

$$\alpha := -\frac{1}{3\,\varepsilon^2 - 2},\ -1,\ -1.$$

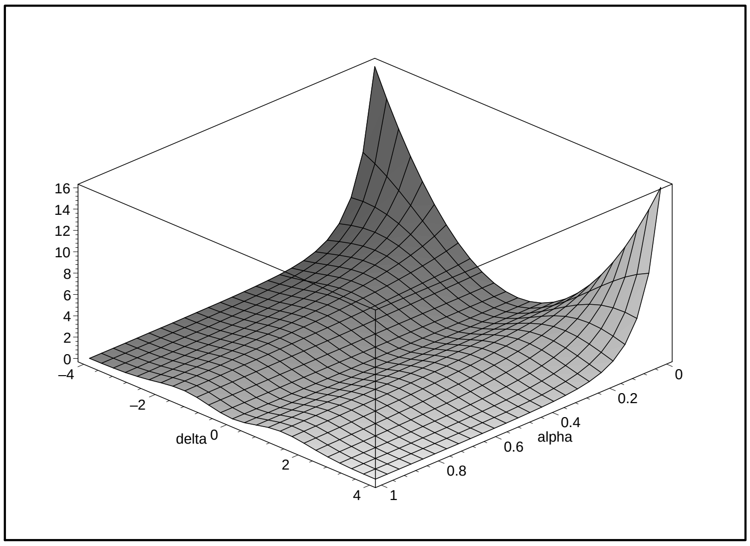

Figure 3: Plot of β^2 setting $\sigma = 1$, $\varepsilon = 1$, $0 \le \alpha \le 1$, $-4 \le \delta \le 4$

The only value of interest is $-\dfrac{1}{3\,\varepsilon^2 - 2}$. Through inspection, one easily obtains the result from the first section and the interesting point is clearly $\varepsilon = \sqrt{\frac{2}{3}}$. When $\varepsilon < \sqrt{\frac{2}{3}}$, a feasible solution is obtained. The case of $\varepsilon \ge \sqrt{\frac{2}{3}}$ must be further investigated. Taking the derivative of β^2 taken with respect to α and substituting $\varepsilon = \sqrt{\frac{2}{3}} + \tau$ $(\tau \ge 0)$ into this expression, yields a value that is always negative. Therefore, when $\varepsilon \ge \sqrt{\frac{2}{3}}$, β^2 is minimized at $\alpha = \infty$. Since, $\sqrt{\frac{2}{3}} \approx 0.80$, the contaminating distribution must inflate the standard-deviation by less than 80% to have a tangible solution. When $\varepsilon = \frac{1}{\sqrt{3}} \approx 0.58$ the AMB solution is $\alpha = 1$, corresponding to the L_2 distance. Figure 3 displays a plot of β^2 with $\varepsilon = 1$.