Multistage Estimation of a Negative Exponential Location Under a Modified Linex Loss Function: Illustrations in Health Studies

Sudeep R. Bapat

(Joint work with Prof. Nitis Mukhopadhyay)

University of Connecticut

June 20, 2016
Outline

1. Introduction
2. The underlying problem
3. Preliminaries
4. Stein type two stage procedure
5. A modified two stage procedure
6. Purely sequential procedure
7. Simulations
8. Real data analysis
9. Conclusions
10. Future ideas
11. References
The highlight of today’s talk involves multistage estimation of a negative exponential location with an unknown scale \( \sigma \), using a modified Linex loss function.

The negative exponential distribution has found its use widely in many reliability experiments such as to depict the failure times of electrical components, complex equipment etc. Another area where it is prominently used is in clinical trials, such as to study the behavior of tumor systems in animals, analysis of survival data in cancer research etc. One may refer to Johnson and Kotz (1970), Bain (1978) for several such examples.
The problem in hand

Again, our goal is to estimate the location parameter of a negative exponential distribution with an unknown scale \( \sigma \), under a variant of the Linex loss function.

We will make use of the usual Stein type two stage, a modified two stage and a purely sequential methodology for estimation purposes.

We will start off with some preliminaries before building the actual stopping rules.
After the Linex loss had been introduced by Varian in 1975, not much had been covered under the sequential estimation field making use of this loss. Chattopadhyay (1998) developed some theory using Linex but was under a normal distribution. He proposed strategies to estimate a normal location and evaluated the asymptotic risk function.

The sequential estimation of a negative exponential location was dealt under a different loss function, majorly the squared error loss. A notable review paper is by Mukhopadhyay in 1988. Again, Chattopadhyay (2000) developed some theory connecting a negative exponential and Linex loss.
The negative exponential distribution

We consider the following negative exponential density:

$$f(x; \mu, \sigma) = \frac{1}{\sigma} \exp \left( -\frac{x - \mu}{\sigma} \right) I(x > \mu)$$  \hspace{1cm} (1)

where $\mu \in (-\infty, \infty)$ and $\sigma \in (0, \infty)$ are unknown parameters. $I(.)$ denotes the indicator function of $(.)$ which in this case takes the value $1(or \ 0)$ whenever $x > (or \leq) 0$.

The parameter $\mu$ if positive is interpreted as the minimum guarantee time or the threshold of the distribution in the sense that no failure will occur before $\mu$. 
The Linex loss function

We will deal with a form of the loss function known as the Linex loss. It would be the key to our estimation of the negative exponential location. It has been used widely after it was introduced by Varian in 1975.

The customary form of this loss is as follows:

\[ L_n(\hat{\mu}_n, \mu) = \exp\{a(\hat{\mu}_n - \mu)\} - a(\hat{\mu}_n - \mu) - 1 \]  

(2)

where \( \hat{\mu}_n \) is a generic estimator of \( \mu \), and \( a \) is a fixed constant.

It is supposed to address estimation error by penalizing overestimation and underestimation unequally in cases where overestimation is deemed more(less) serious than underestimation whenever \( a > (<) 0 \).

For the problem in hand, we will use a modified form of the above loss function which will be elaborated next.
Examples of Linex curves

\[ a = -1 \quad a = 1 \quad a = 0.5 \]
Some practical scenarios

- Construction of a dam: Underestimation of height of a dam is much more serious than an overestimation. We can thus hope to use a Linex loss function with $a < 0$.

- Reliability example: If we assume $X$ to be the lifetime of a certain equipment, then the reliability function is given by $R(t) = P[X > t]$. An overestimation in reliability would be more serious than underestimation.
Preliminaries: The modified Linex loss

Having recorded a random sample $X_1, X_2, \ldots, X_n$ from a negative exponential distribution (1), one would estimate $\mu$ by the customary MLE estimator $X_{n:1}$.

Now we will modify the loss given in (2) and use that for the problem in hand. This is to eliminate the role of $\sigma$ which is unknown. Thus the modified Linex loss is as follows:

$$L_n = \exp \left( \frac{a(X_{n:1} - \mu)}{\sigma} \right) - \frac{a(X_{n:1} - \mu)}{\sigma} - 1$$

(3)

where all the constants are as defined before.
The usual Linex loss function is given by:

\[ \exp(a(X_{n:1} - \mu)) - a(X_{n:1} - \mu) - 1, \ n \geq 1 \]

The associated risk function requires one to evaluate \( E[\exp(a(X_{n:1} - \mu))] \) which can be alternatively expressed as, \( E[\exp\left(\frac{a\sigma}{n}Y\right)] \), where \( Y \sim Exp(1) \).

However this expectation will exist provided \( n > a\sigma \). Since \( \sigma \) is unknown, there is no way to guarantee that the sample size will indeed exceed \( a\sigma \) when \( a > 0 \).
The risk function

The risk function is obtained by taking expectations across (3) as follows:

\[
Risk_n \equiv E[L_n] = E \left[ \exp \left( \frac{a(X_{n:1} - \mu)}{\sigma} \right) - \frac{a(X_{n:1} - \mu)}{\sigma} - 1 \right]
\]

\[
= \left(1 - \frac{a}{n}\right)^{-1} - \frac{a}{n} - 1, \text{ for } n > a
\]

which upon expanding gives,

\[
Risk_n = \frac{a^2}{n^2} + o\left(\frac{1}{n^2}\right)
\]
We now introduce a function for cost per observation whose exact nature would depend upon the problem in hand. In the current situation we propose the following cost:

$$Cost_n = cn\sigma^{-k}$$

with fixed and known $c(>0)$ and $k(>0)$. The idea is that the above cost will increase (decrease) with decreasing (increasing) $\sigma$. 
We now wish to bound the associated "risk" from above, where we interpret the "risk" as *risk per unit cost* (RPUC) given by:

\[ RPUC_n = \frac{Risk_n}{Cost_n} = \frac{a^2}{n^2} \cdot \frac{\sigma^k}{cn} + o(n^{-3}) \]

We thus fix a suitable constant \( \omega(>0) \) and require that \( RPUC_n \leq \omega \) for all \( \mu, \sigma \). This will lead us to determine the *optimum fixed sample size* \( n^* \) as follows:

\[ n^* = \left( \frac{a^2}{cw} \right)^{1/3} \sigma^{k/3} \]  \hspace{1cm} (5)

The magnitude of this \( n^* \) is unknown even though it has a known expression. We hence resort to multistage bounded risk estimation strategies which will be elaborated next.
Stein type two stage procedure

The usual two stage procedure as introduced by Stein in 1945 is as follows: Stage 1 consists of collecting a pilot sample \(X_1, X_2, \ldots, X_m\) of size \(m(>a, \geq 2)\) from which we construct the stopping rule as:

\[
N = \max \left\{ m, \left[ \left( \frac{a^2}{cw} \right)^{1/3} \hat{\sigma}_m^{k/3} \right] + 1 \right\}
\]

which is based on value of \(n^*\) given earlier. Here, \(\hat{\sigma}_m\) is the UMVUE of \(\sigma\). If \(N = m\), we will not sample at the second stage. However if \(N > m\), we sample the difference \(N - m\) at the second stage by recording an additional set of observations \(X_{m+1}, X_{m+2}, \ldots, X_N\).

The form of the risk function involving \(N\) is as follows:

\[
RPUC_N = \frac{n^*^3 w}{a^2} \left\{ E \left[ \frac{1}{N} \left( 1 - \frac{a}{N} \right)^{-1} \right] - E \left( \frac{a}{N^2} \right) - E \left( \frac{1}{N} \right) \right\}
\]

We propose to estimate \(\mu\) by the smallest order statistic \(X_{N:1}\).
A modified two stage procedure

First order approximations, especially the ratios $N/n^*$ or $E(N/n^*)$ do not converge to 1 under the Stein type two stage strategy. We hence resort to a modification along the lines of Mukhopadhyay and Duggan (1997) as follows:

We introduce a known lower bound $\sigma_L (> 0)$ such that $0 < \sigma_L < \sigma$.

We then fix an integer $m_0 (> max\{1, a\})$ and gather a pilot of size $m$

given by:

$$m = m(w) = max \{m_0, \left\lfloor \left( \frac{a^2\sigma_kL}{cw} \right)^{1/3} \right\rfloor + 1\}$$

The estimator of $n^*$ given by $N$ is as defined earlier in (6) and we again propose to estimate $\mu$ by $X_{N:1}$. The advantage of this modification will be clear next.
Some results

Following are some results based on the stopping variable $N$ associated with the modified two stage methodology.

### First order approximations

as $\omega \to 0$,

1. $\frac{N}{n^*} \xrightarrow{} 1$ w.p.1.
2. $E\left(\frac{N}{n^*}\right) \xrightarrow{} 1$
3. $E\left(\frac{n^*}{N}\right) \xrightarrow{} 1$

### Second order approximation

as $\omega \to 0$,

$$\psi + O(\omega^{1/2}) \leq E(N) - n^* \leq \psi + 1 + O(\omega^{1/2})$$

where $\psi$ and $\sigma_0^2$ are suitable constants defined in terms of $k, \sigma$ and $\sigma_L$. 

Sudeep R. Bapat (University of Connecticut)  
IWAP Toronto  
June 20, 2016 17 / 28
Purely sequential procedure

Here, we sample sequentially rather than in batches and thus it proves to provide more accurate inferences as compared to the previous methodologies.

The values of $n^*$ and Risk remain same as before. However there is a change in the stopping rule $N$ as follows:

We fix $m( > \max\{1, a\})$ as pilot and collect a sample $X_1, X_2, \ldots, X_m$. We hence record one additional observation at every stage according to the following rule:

$$N = \inf \left\{ n \geq m : n \geq \left( \frac{a^2}{cw} \right)^{1/3} \hat{\sigma}_{n/3}^k \right\}$$

Once again we estimate $\mu$ by $X_{N:1}$ and some basic results follow.
Apart from the first order approximations converging to 1, we make use of non-linear renewal theory as per Woodroofe (1977) to derive some higher order approximations involving $N$ which are as follows:

**Second order approximations**

as $\omega \to 0$,

$$E \left[ \left( \frac{N}{n^*} \right)^t \right] = 1 + \left\{ t\eta + t + \frac{1}{2} t(t - 1)p \right\} \frac{1}{n^*} + o \left( \frac{1}{n^*} \right)$$

if,

a) $m > \frac{(3-t)k}{3} + 1$ for $t \in (-\infty, 2) - \{-1, 1\}$.

b) $m > \frac{k}{3} + 1$ for $t = 1$ and $t \geq 2$.

c) $m > \frac{2k}{3} + 1$ for $t = -1$.

where $p$ and $\eta$ are predefined known constants.
Asymptotic distribution and Risk

As before, an asymptotic distribution involving $N$ is given by:

$$V \equiv n^{*-1/2} (N - n^*) \xrightarrow{L} N(0, \sigma_1^2)$$

where $\sigma_1^2 = \frac{k^2}{9}$.

The second order risk efficiency has a concise formula and is given by:

$$\omega^{-1} \text{RPUC}_N = 1 + \frac{1}{n^*} (6p + a - 3\eta - 3) + o\left(\frac{1}{n^*}\right)$$

when $m > \frac{7k}{3} + 1$. 
### Table 4.1
Simulation results from 10000 replications of the modified two stage procedure with $\mu = 5$, $\sigma = 10$, $a = 1$ and $c = 0.1$. The values of $m_0$ corresponding each $k$ are 4 and 6.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$n^*$</th>
<th>$\omega$</th>
<th>$\bar{x}<em>{\text{min}}, s(\bar{x}</em>{\text{min}})$</th>
<th>$\bar{n}, s(\bar{n})$</th>
<th>$\bar{n}/n^*$</th>
<th>$R/\omega, s(R/\omega)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>30</td>
<td>$3.7 \times 10^{-2}$</td>
<td>5.3410 (0.0035)</td>
<td>30.35 (0.0554)</td>
<td>1.0119</td>
<td>1.2423 (0.0082)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>$3.7 \times 10^{-5}$</td>
<td>5.0230 (0.0003)</td>
<td>300.26 (0.1731)</td>
<td>1.0003</td>
<td>1.0210 (0.0017)</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>3.7</td>
<td>5.5033 (0.0050)</td>
<td>31.63 (0.1684)</td>
<td>1.0546</td>
<td>1.6420 (0.0908)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>$3.7 \times 10^{-3}$</td>
<td>5.0336 (0.0003)</td>
<td>302.45 (0.5159)</td>
<td>1.0081</td>
<td>1.0987 (0.0066)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\sigma_L = 3$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>$3.7 \times 10^{-2}$</td>
<td>5.3312 (0.0033)</td>
<td>30.32 (0.0469)</td>
<td>1.0109</td>
<td>1.1678 (0.0061)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>$3.7 \times 10^{-5}$</td>
<td>5.0340 (0.0003)</td>
<td>300.46 (0.1468)</td>
<td>1.0015</td>
<td>1.0132 (0.0015)</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>3.7</td>
<td>5.3681 (0.0041)</td>
<td>31.20 (0.1222)</td>
<td>1.0401</td>
<td>1.3429 (0.0343)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>$3.7 \times 10^{-3}$</td>
<td>5.0338 (0.0003)</td>
<td>301.47 (0.3687)</td>
<td>1.0049</td>
<td>1.0837 (0.0043)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\sigma_L = 5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>$3.7 \times 10^{-2}$</td>
<td>5.3078 (0.0030)</td>
<td>32.78 (0.0162)</td>
<td>1.0926</td>
<td>0.8005 (0.0009)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>$3.7 \times 10^{-5}$</td>
<td>5.0311 (0.0003)</td>
<td>320.20 (0.0127)</td>
<td>1.0673</td>
<td>0.8250 (0.0002)</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>3.7</td>
<td>5.2792 (0.0028)</td>
<td>36.21 (0.0305)</td>
<td>1.2070</td>
<td>0.6026 (0.0010)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>$3.7 \times 10^{-3}$</td>
<td>5.0291 (0.0002)</td>
<td>341.34 (0.0240)</td>
<td>1.1378</td>
<td>0.6810 (0.0005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\sigma_L = 11$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Purely sequential procedure

Here, \( \lambda = n^{*-1} (6p + a - 3\eta - 3) \) and we expect \( \bar{R}/\omega \approx 1 + \lambda \).
The data-set in question is the **infant mortality data** from Leinhardt and Wasserman (1979). Treating this data as the population, we found $\mu = 9.6$ and $\sigma = 80.24$.

### Infant mortality data

**Table 4.3.** Analysis of infant mortality rate using modified two stage procedure with $a = 1, c = 0.1, \sigma_L = 40$

<table>
<thead>
<tr>
<th>$n^*$</th>
<th>$m_0$</th>
<th>$k$</th>
<th>$\omega$</th>
<th>$\mu$</th>
<th>$N$</th>
<th>$N/n^*$</th>
<th>$\overline{R}/\omega$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>4</td>
<td>1</td>
<td>0.01253</td>
<td>9.6</td>
<td>41</td>
<td>1.025</td>
<td>0.9518</td>
</tr>
<tr>
<td>60</td>
<td>7</td>
<td>3</td>
<td>23.9176</td>
<td>10.1</td>
<td>64</td>
<td>1.066</td>
<td>0.8370</td>
</tr>
<tr>
<td>70</td>
<td>10</td>
<td>5</td>
<td>96975.15</td>
<td>9.8</td>
<td>62</td>
<td>0.8857</td>
<td>1.4627</td>
</tr>
</tbody>
</table>

**Table 4.4.** Analysis of infant mortality rate using purely sequential procedure with $a = 1, c = 0.1$

<table>
<thead>
<tr>
<th>$n^*$</th>
<th>$m$</th>
<th>$k$</th>
<th>$\omega$</th>
<th>$\mu$</th>
<th>$N$</th>
<th>$N/n^*$</th>
<th>$\overline{R}/\omega$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>5</td>
<td>1</td>
<td>0.01253</td>
<td>12.8</td>
<td>39</td>
<td>0.975</td>
<td>1.0144</td>
<td>0.0171</td>
</tr>
<tr>
<td>60</td>
<td>9</td>
<td>3</td>
<td>23.9176</td>
<td>10.1</td>
<td>52</td>
<td>0.8667</td>
<td>1.5663</td>
<td>0.1294</td>
</tr>
<tr>
<td>70</td>
<td>15</td>
<td>5</td>
<td>96975.15</td>
<td>10.1</td>
<td>70</td>
<td>1</td>
<td>1.1073</td>
<td>0.3063</td>
</tr>
</tbody>
</table>
Some concluding thoughts

As seen from the proposed multistage estimation strategies, the modified two stage sampling scheme proves to have an edge over the usual Stein type two stage methodology. This is in the sense that the first and second-order approximations as well as the Risk function converge to 1. The sequential procedure proves to be better than the modified two stage strategy but one must balance any logistical concerns.

These thoughts were also seen to be validated by simulated data and a real data from Health studies.
Some future ideas

- One can look at the multistage estimation strategies for a two sample case, wherein one can estimate the difference in locations of two independent negative exponential populations under an appropriate Linex loss.
- A three stage or an accelerated sequential rule could be incorporated instead of the given methodologies.
- The field of change point problems is well developed. We may explore its relevance under the present scenario.


This work is in print and the details are:


http://dx.doi.org/10.1080/07474946.2016.1165532.
THANK YOU !!