Infant Mortality In India: Evaluating Log-Gaussian and Gamma Distributions

Rabindra Nath Das¹, Shankar Dihidar², and Richard R. Verdugo³,*

¹Department of Statistics, The University of Burdwan, Burdwan, W.B., India
²Indian Statistical Institute, 203 B.T. Road, Kolkata-108, India.
³National Education Association, Washington, DC. USA

Abstract: Infant mortality is a serious problem in India. In order to better understand the problem previous research has looked at the topic using sophisticated multivariate models assuming that infant mortality is either Gaussian or Log-Gaussian. In this paper we argue that infant mortality is Log-Gaussian distributed with non-constant variance and that making such an assumption leads to more efficient estimates and a better fit to the data. Using infant mortality data from the National Health Survey in Bihar, India we compare two distributions—Log Gaussian and Gamma—and find that the Log-Gaussian non-constant variance model does indeed lead to more efficient estimates and a better fit to the data.

Keywords: Infant survival time; Joint generalized linear model; Multiplicative model; Non-constant coefficient of variation; structured dispersion.

1. INTRODUCTION

Infant mortality is not only an important factor in population growth; it is also an important measure of economic development. Throughout the world it appears that the infant mortality rate (IMR) has declined since the 1960s. Indeed, between 1960 and 1998 there was a world-wide 55 percent decrease in the IMR [1]. Variations exited, though. In the Sub-Sahara region of Africa, for example, the IMR declined by 34 percent between 1960 and 1989. Nevertheless, the world-wide decline in the IMR has been good news.

The IMR in India has also been the subject of much research. Though India’s IMR has been declining, in 2009, it ranked 143rd in IMR with 55 infant deaths per 1,000 live population.¹ Such a record is troubling not only because of the loss in human life, but because India is an emerging economic super-power, and it’s IMR is clearly not related to its emerging economic status.

Why does India have such a high IMR? What are the factors that contribute to its high IMR? An existing body of research has attempted to answer these and related questions [2-12]. The problem is complex and scholars have used a wide array of factors to better understand the IMR in India. Yet, there are some important concerns over the assumption that IMR data are Gaussian distributed. That is, past research has estimated IMR using multivariate models that assume a Gaussian distribution. Such an assumption may lead to inefficient and biased estimates, and draw researchers into making important conclusions from dubious results if the distribution is not Gaussian [13]. We have two objectives in this paper. Our first objective is to estimate the impact selected risk factors have on the IMR in Bihar, India. Our second objective is to evaluate two models based on different assumptions regarding infant mortality: one based on the assumption that infant mortality is Log-Gaussian distributed, and the other assumes that it is Gamma distributed. Our data are from the Indian Survey of the National Family Health Survey-2 (NFHS-2) conducted in 1998-1999.

II. BACKGROUND

A. Research on Infant Mortality in India

Previous research on India’s IMR begins by setting up a model that contains risk factors in three domains: proximate factors, maternal factors, and household/community factors [2-12]. Proximate factors are those items that involve medical care and non-medical care during the antenatal period, care at birth, and care during the postnatal period. Maternal factors refer to such things as the age, and birth intervals of the mother. And finally, household and community factors refer to such things as sanitation, water supply, and household and community cleanliness. Jain and Visaria [2] found that significant declines in the IMR are possible without improvement in societal economic development. What the authors found instead was that access to a small number of health and maternal services reduced the IMR: reproductive health services, perinatal care, improved breast feeding, immunization, the treatment of diarrhea, and the introduction to

¹Data from the CIA Factbook. See the Wikipedia website: http://en.wikipedia.org/wiki/List_of_countries_by_infant_mortality_rate
supplementary foods. Other research has reported similar findings [6].

In terms of economics, research has shown that economics is not a major factor in reducing the IMR in India, rather it is those non-economic factors, such as maternal and child health interventions [3, 4, 6].

Both sets of studies inform our own study in two ways. First, they assist us in setting up a model for analysis that includes predictors from the three risk-factor domains: proximate, maternal, and household/community. Secondly, it is our sense that previous research may have been in error because they estimated models assuming a Gaussian or Log-Gaussian distribution with constant variance, when, in fact, the distribution may be Gamma distributed and/or with non-constant variance. For example, Myers et al. [13] found that transforming data for stabilization may not, in fact, stabilize the distribution (see also [14] for a discussion). Results from previous research, then, might be inefficient and biased. For example, Fig. (1) presents the actual distribution of the IMR data to be used in our study. Note how the distribution bears a strong resemblance to a Gamma distribution. Estimating these data under the Gaussian assumption could lead to serious errors.

Our study attempts to contribute to this research by estimating two models: one that assumes that the IMR is Log-Gaussian and the other Gamma distributed. We propose to use the simultaneous modeling of the mean and dispersion in infant mortality by using a Joint Generalized Linear Model (JGLM).

B. Log-Normal and Gamma Models with Constant Variance

For heteroscedastic data, the log-transformation is often recommended in stabilizing the variance [15]. It is well-known that if the variance is constant, parameters from the Log-Gaussian and Gamma models have a common interpretation [16]. However, as Das and Lee [14] have shown that the simple log-transformation may not be sufficient in stabilizing the variance, and that a different structured constant dispersion may be required. Further, with structured dispersion it is unlikely that the two models (Log-Gaussian and Gamma) will provide coefficients with a common interpretation [14].

In regression models with multiplicative error, the estimation approach is commonly based on either the Log-Gaussian or the Gamma model [16]. For positive observations it is well-known that the logarithm of the dependent variable leads to a correspondence between multiplicative regression and additive models. The correspondence satisfies the classical linear model’s assumption that the variance of the response (Y) is constant over the entire range of parameter values (homoscedastic).

However, when the variance increases with the mean (heteroscedastic) we might consider a model with a constant coefficient of variation:

\[
\text{Var}(Y) = \sigma^2 \mu_Y^2,
\]

where \( \sigma \) is the coefficient of variation of \( Y \) and \( \mu_Y = \text{E}(Y) \). In generalized linear models [17] the gamma model satisfies the above mean and variance relationship. For small \( \sigma \), the variance-stabilizing transformation, \( Z = \log(Y) \), has approximate moments

\[
\text{E}(Z) = \log \mu_Y - \sigma^2/2 \quad \text{and} \quad \text{Var}(Z) \equiv \sigma^2
\]

If the systematic part of the model is multiplicative on the original scale, and hence additive on the log scale, then

\[ Y_i = \mu_Y \cdot \epsilon_i \quad (i = 1, 2, \ldots, n) \quad (1) \]
with
\[ \eta_i = \log \mu_{Yi} = x_i^\prime \beta + x_i \beta_1 + \ldots + x_{ip} \beta_p \]
and the errors, \( \{C_i\} \)'s, are independent identically distributed (IID) with \( \mathbb{E}(C_i) = 1 \). In GLMs \( \mu_{Yi} \) is the scale parameter and \( \text{Var}(C_i) = \sigma^2 \) is the shape parameter. Consequently,
\[ Z_i = \log Y_i = \mu_{Zi} + \delta_i \quad (i = 1, 2, \ldots, n) \quad (2) \]
with
\[ \mu_{Zi} = \mu_{Yi} + \text{E}(\log(C_i)) + x_i \beta_1 + \ldots + x_{ip} \beta_p \]
and \( \{\delta_i = \log(C_i - E(\log(C_i)))\}'s are IID with \( \mathbb{E}(\delta_i) = 0. \)

Conversely, if \( Y_i \) follows a log-normal distribution, i.e., \( Y_i \sim N(\mu_{Zi}, \sigma^2_{Zi}) \) then
\[ \mu_{Yi} = \mathbb{E}(\exp Z_i) = \exp(\mu_{Zi} + \sigma^2_{Zi}/2) \neq \exp(\mu_{Zi}). \]

With the exception of the intercept term the remaining parameters \( \beta_1, \beta_2, \ldots, \beta_p \) can be estimated either from the constant coefficient of variation model (1) or the linear model for the transformation of the original data to a log scale (2) with a common interpretation [13]. Firth [16] provided a comparison of the efficiencies of the maximum-likelihood (ML) estimators from a Gamma model (the constant coefficient of variation model) when the errors are in fact Log-Gaussian, with those of the Log-Gaussian model when the errors have a Gamma distribution. He concludes that the ML estimators from the Gamma model perform slightly better under reciprocal specification. For small \( \sigma^2 \) it is likely to be difficult to discriminate between the Normal-theory linear models for log \( Y \), and Gamma-theory multiplicative models for \( Y \).

1. Multiplicative Models with Non-Constant Variance

Issues in industrial processes are germane to the issue in our paper. A common problem in an industrial process is to find an operating condition that achieves the target value for the mean of a process characteristic, and simultaneously minimizes the process variability. If \( \sigma^2 \) is not constant, i.e., \( C_i \) is not IID with a common \( \mathbb{E}(\log C_i) \), parameter estimates from one model may have no correspondence with the other under reciprocal specification. For small \( \sigma^2 \) it is likely to be difficult to discriminate between the Normal-theory linear models for log \( Y \), and Gamma-theory multiplicative models for \( Y \).

A. Data

The National Family Health Survey, Bihar 1998-99: India's first National Family Health Survey (NFHS-1) was conducted in 1992-93. The Ministry of Health and Family Welfare (MOHFW) subsequently designated the International Institute for Population Sciences (IIPS) in Mumbai, as the agency to initiate a second survey (NFHS-2), which was conducted in 1998-99. An important objective of (NFHS-2) was to provide State-level and National-level information on fertility, family planning, infant and child mortality, reproductive health, child health, nutrition of women and children, and the quality of health and family welfare services.
Another important objective of the NFHS-2 was to examine the above information in the context of socioeconomic and cultural factors. NFHS-2 used three types of questionnaires: the Household questionnaire, the Woman Questionnaire, the Village Questionnaire. The Woman Questionnaire collected information from ever-married women belonging to the age cohort 15-49 who were residents of the sampled household. Female respondents were asked about their background, the details of births of their children during the preceding three years, and whether they practiced contraception. The Child Questionnaire was designed to record details of antenatal care, details of delivery, breastfeeding, and post partum amenorrhoea, immunization and health care for the two most recent births during the three years preceding the survey. In our analysis, we have taken a random sample (from NFHS-2 data, India) of 139 ever-married women, aged between 15 and 49 who resided in the state of Bihar, India.

### B. VARIABLES

#### 1. Dependent Variable

The dependent variable for our study is the age at which an infant died. Age is measured in months.

### 2. Independent Variables

There are three categories of independent variables to be used in our analysis: proximate, maternal, and household/community. Table 1 presents a description of each set of items and how they are operationalized for our study.

#### IV. FINDINGS

##### A. Descriptive Statistics

Table 2 presents means and standard deviations for all variables to be used in our models. In terms of Household/Community items, these data suggest that the vast majority of the sample live in rural areas (96.4%), and are of the lower caste (50.4%). Moreover, it appears that respondents are mostly not Muslim (15%). Finally, note that about 40 percent of the infants that died are female.

In terms of Proximate items for infants, most infants do not appear to have received their vaccinations. Only about 4 percent received their first BCG, and 5% their first DPT; 4% their second DPT; and 10% their first polio shot, and 9% their second polio shot. In contrast, about 52 percent have
actually been vaccinated for tetanus, and only about two per-
cent of infants received their measles vaccinations.

There are two Maternal items in our study: breast feeding, and mother’s age. A majority of women breast feed their
infants (55%), and the average age of mothers in our study is
26.65 years. We should also point out that the average age of
death among infants, in months, was 3.21 months.

B. Comparing Log-Gaussian and Gamma Models

Table 3 presents GLM results from the Log-Gaussian
Model (LGM) and Gamma Model (GM) with log-link of
infant mortality in Bihar, India.

There are three criteria we apply in conducting our com-
parison: We look at the Akaike information criterion (AICs)
for the best fitting model; we look at the standard errors; and
we also examine graphical analysis. First, we look to see if
the coefficients appear to be statistically different. Contrary
to what we had expected, the models are generally not dif-
ferent in their effects on infant mortality in Bihar, India. The
three exceptions are mother’s age (Mage), place of delivery
(Deliver) and Polio1, where the Gamma model exhibits sig-
nificant effects whereas the Log-Gaussian model does not.

How well do each of the models fit the data? Apparently,
from the information in Table 3, the better fit is the Log-
Gaussian model (LGM). The LGM has an AIC of 499.1,
while the Gamma Model’s (GM) AIC is 532.128.

We also note that the standard errors in the Log-Gaussian
model are only slightly smaller than the Gamma model. The
implication is that the Log-Gaussian model is only slightly
more efficient and bolsters our confidence in using it rather
than the Gamma model.

We next examined the model fit based on graphical
analysis. In Fig. 2(a) and Fig. 2(b), we plot the absolute val-
ues of residuals with respect to fitted values for LGM and
GM, respectively. Both the figures show that variances in-
crease with the increase in the means, indicating that vari-

---

**Table 2. Means and Standard Deviations for all Items in the Analysis**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proximate Items</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tetanus1</td>
<td>.52</td>
<td>.50</td>
</tr>
<tr>
<td>BCG1</td>
<td>.04</td>
<td>.19</td>
</tr>
<tr>
<td>Polio1</td>
<td>.10</td>
<td>.30</td>
</tr>
<tr>
<td>Polio2</td>
<td>.09</td>
<td>.29</td>
</tr>
<tr>
<td>Deliver</td>
<td>.25</td>
<td>.43</td>
</tr>
<tr>
<td>DPT1</td>
<td>.05</td>
<td>.22</td>
</tr>
<tr>
<td>DPT2</td>
<td>.04</td>
<td>.19</td>
</tr>
<tr>
<td>Measles</td>
<td>.02</td>
<td>.15</td>
</tr>
<tr>
<td><strong>Maternal Items</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>.55</td>
<td>.50</td>
</tr>
<tr>
<td>Mage</td>
<td>26.65</td>
<td>6.80</td>
</tr>
<tr>
<td><strong>Household/Community Items</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caste1</td>
<td>.36</td>
<td>.48</td>
</tr>
<tr>
<td>Caste2</td>
<td>.50</td>
<td>.50</td>
</tr>
<tr>
<td>Religion</td>
<td>.15</td>
<td>3.6</td>
</tr>
<tr>
<td>Cem</td>
<td>.40</td>
<td>.49</td>
</tr>
<tr>
<td><strong>Dependent Variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cdeath</td>
<td>3.21</td>
<td>5.15</td>
</tr>
</tbody>
</table>
ances are non-constant under both the models. The implication is that GLM fits, for both the distributions, are inappropriate. This finding leads us to fit joint GLM for reducing the variance.

Results for fits for the LGM and the GM are displayed in Table 4. In this table we limit the model to those items that were statistically significant.

Table 3. Results for mean with constant variance models of infant survival time data from Log-Gaussian & Gamma fits (with all covariates)

<table>
<thead>
<tr>
<th>Covar.</th>
<th>Log Gaussian Model</th>
<th></th>
<th></th>
<th>Gamma Model</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>estimate</td>
<td>s.e</td>
<td>t</td>
<td>P-value</td>
<td>estimate</td>
<td>s.e</td>
</tr>
<tr>
<td>Mean model</td>
<td>Const.</td>
<td>-1.4819</td>
<td>0.8141</td>
<td>-1.8202</td>
<td>0.07</td>
<td>-1.3345</td>
</tr>
<tr>
<td></td>
<td>Mage</td>
<td>0.0284</td>
<td>0.0161</td>
<td>1.763</td>
<td>0.08</td>
<td>0.0403</td>
</tr>
<tr>
<td></td>
<td>Rural</td>
<td>0.3560</td>
<td>0.5699</td>
<td>0.6247</td>
<td>0.53</td>
<td>0.2823</td>
</tr>
<tr>
<td></td>
<td>Religion</td>
<td>-0.0869</td>
<td>0.3509</td>
<td>-0.2475</td>
<td>0.80</td>
<td>0.1074</td>
</tr>
<tr>
<td></td>
<td>Caste 1</td>
<td>0.0967</td>
<td>0.4014</td>
<td>0.2408</td>
<td>0.81</td>
<td>0.1543</td>
</tr>
<tr>
<td></td>
<td>Caste2</td>
<td>-0.0292</td>
<td>0.3814</td>
<td>-0.0767</td>
<td>0.93</td>
<td>0.0287</td>
</tr>
<tr>
<td></td>
<td>Cfem</td>
<td>-0.0650</td>
<td>0.2235</td>
<td>-0.2907</td>
<td>0.77</td>
<td>-0.0673</td>
</tr>
<tr>
<td></td>
<td>Tetanus1</td>
<td>0.1310</td>
<td>0.2339</td>
<td>0.5600</td>
<td>0.57</td>
<td>0.1926</td>
</tr>
<tr>
<td></td>
<td>Deliver</td>
<td>0.3276</td>
<td>0.2908</td>
<td>1.1267</td>
<td>0.26</td>
<td>0.6070</td>
</tr>
<tr>
<td></td>
<td>BCG 1</td>
<td>0.4075</td>
<td>1.2513</td>
<td>0.3257</td>
<td>0.74</td>
<td>1.0759</td>
</tr>
<tr>
<td></td>
<td>DPT 1</td>
<td>-0.0015</td>
<td>1.2023</td>
<td>-0.0013</td>
<td>0.99</td>
<td>-0.4947</td>
</tr>
<tr>
<td></td>
<td>POLIO 1</td>
<td>-0.2380</td>
<td>1.2629</td>
<td>-0.1885</td>
<td>0.85</td>
<td>-1.3946</td>
</tr>
<tr>
<td></td>
<td>DPT2</td>
<td>-0.4041</td>
<td>1.2401</td>
<td>-0.3259</td>
<td>0.74</td>
<td>-0.5802</td>
</tr>
<tr>
<td></td>
<td>POLIO 2</td>
<td>-0.2783</td>
<td>1.3566</td>
<td>-0.2051</td>
<td>0.84</td>
<td>0.6248</td>
</tr>
<tr>
<td></td>
<td>Measles</td>
<td>0.0659</td>
<td>1.3178</td>
<td>0.0500</td>
<td>0.96</td>
<td>-0.0123</td>
</tr>
<tr>
<td></td>
<td>Breast</td>
<td>0.9742</td>
<td>0.2406</td>
<td>4.0493</td>
<td>0.00</td>
<td>1.2864</td>
</tr>
<tr>
<td>Constant Dispersion</td>
<td>Const.</td>
<td>0.4024</td>
<td>0.1275</td>
<td>3.156</td>
<td>0.00</td>
<td>0.2451</td>
</tr>
<tr>
<td>AIC</td>
<td></td>
<td>499.100</td>
<td></td>
<td></td>
<td>532.128</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. (2).** Absolute residual plots with respect to fitted values for constant variance models of (a) Log-Gaussian & (b) Gamma for infant survival time in Table 3.

Table 4 shows that joint LGM fit is much better than joint GM fit based on AIC and standard errors (mean model parameters) as explained above. Indeed, while the GM fit remains unchanged (532.128 v. 531.897) the LGM fit is reduced by 3.75 percent (480.4 v. 499.1). Fig. 3(a) and Fig. 3(b) plot the absolute residuals with respect to fitted values and the normal probability plot for the mean of the joint GM
The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

The Open Demography Journal, 2011, Volume 4

Das et al.

Fig. 3. For final Gamma non-constant variance models (a) The Absolute residual plots with respect to fitted values & (b) The Normal probability plot for mean of infant survival time in Table 4.

Table 4. Results for mean & constant dispersion of infant survival time data from Log-Gaussian & Gamma fit

<table>
<thead>
<tr>
<th></th>
<th>Log Gaussian Model</th>
<th></th>
<th></th>
<th>Gamma Model</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Covar.</td>
<td>estimate</td>
<td>s.e</td>
<td>t</td>
<td>P-value</td>
<td>estimate</td>
</tr>
<tr>
<td>Mean model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Const.</td>
<td></td>
<td>-1.1103</td>
<td>0.3473</td>
<td>-3.197</td>
<td>0.00</td>
<td>-1.2571</td>
</tr>
<tr>
<td>Mage</td>
<td></td>
<td>0.0248</td>
<td>0.0132</td>
<td>1.878</td>
<td>0.06</td>
<td>0.0456</td>
</tr>
<tr>
<td>Deliver</td>
<td></td>
<td>0.5127</td>
<td>0.2368</td>
<td>2.165</td>
<td>0.03</td>
<td>0.9243</td>
</tr>
<tr>
<td>Breast</td>
<td></td>
<td>1.0252</td>
<td>0.1839</td>
<td>5.575</td>
<td>0.00</td>
<td>1.4411</td>
</tr>
<tr>
<td>Constant Dispersion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Const.</td>
<td></td>
<td>-1.6808</td>
<td>0.6027</td>
<td>-2.789</td>
<td>0.01</td>
<td>-0.9600</td>
</tr>
<tr>
<td>Mage</td>
<td></td>
<td>0.0352</td>
<td>0.0191</td>
<td>1.838</td>
<td>0.06</td>
<td>0.0245</td>
</tr>
<tr>
<td>Breast</td>
<td></td>
<td>1.2507</td>
<td>0.2828</td>
<td>4.423</td>
<td>0.00</td>
<td>0.5981</td>
</tr>
<tr>
<td>Deliver</td>
<td></td>
<td>0.9851</td>
<td>0.3228</td>
<td>3.052</td>
<td>0.00</td>
<td>0.5997</td>
</tr>
<tr>
<td>AIC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>480.400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. For final Gamma non-constant variance models (a) The Absolute residual plots with respect to fitted values & (b) The Normal probability plot for mean of infant survival time in Table 4.

Similarly, we plot Fig. 4(a) (absolute residual plot) and Fig. 4(b) (normal probability plot for the mean) for the joint LGM fit. Fig. 4(a) and Fig. 4(b) show almost the same features as Fig. 3(a) and Fig. 3(b). Thus, we conclude that both the fits are satisfactory. In addition, the absolute residual plot for joint LGM fit is much flatter than the joint GM fit. Thus, AIC, standard errors for mean model parameters, the number of significant effects, and graphical analysis suggest that the joint LGM fit is much better than the joint GM fit. The joint LGM fit shows that all the effects are significant (maximum at 6%), whereas in the joint GM model, the constant and Mother’s age fail to meet the 10 percent level of statistical significance (Table 4).

V. CONCLUSION

Infant mortality in India is an important social and medical problem. In attempting to better understand the issues surrounding this important problem, researchers have focused on factors related to pre-and postnatal factors. The majority of findings from this body of research suggest that medical care trumps economics and social status in reducing infant mortality.

Our research had two purposes. The first was to compare our results to those of previous research. A second purpose was to evaluate the statistical assumption made by previous research regarding the distribution of infant mortality data. Previous research estimated models assuming a Log-Gaussian distribution with constant variance, but infant mortality data appears to be Log-Gaussian or Gamma with non-constant variance. Our concern was that previous research, making the Gaussian assumption and then applying multivariate models, would draw important conclusions from erroneous assumptions. We therefore estimated and compared two models: a Log-Gaussian Model and a Gamma Model.

Our results, though not completely conclusive, are revealing—

- Our findings confirm previous research by noting that Proximate and House/Community factors tend to increase the life span of infants.
Contrary to what we had expected, the models were remarkably similar in their effects on infant mortality.

There were two exceptions. The intercept and Mage in the joint Log-Gaussian variance model were statistically significant, but not in the joint Gamma variance model. A second exception concerned model fit. The joint Log-Gaussian model exhibited a better fit to the data than the joint Gamma model. The AICs for the joint Log-Gaussian and joint models were: 480.400 vs. 531.897.

There are two conclusions that can be drawn from our research. First, in order to reduce infant mortality in India, policy and practice should continue to focus on pre- and post-natal care. Our findings and those of previous research have continually shown that such practices reduce infant mortality.

A second conclusion has to do with the use of statistical models. While further research is called for, we find that a joint Log-Gaussian model is much more effective than either traditional Log-Gaussian (with constant variance) or joint Gamma models because it better fits the data. In short, research should have greater faith in these results than those emanating from the joint Gamma and Log-Gaussian (with constant variance) models.

REFERENCES
