A COMPARISON OF COMPOUND POISSON CLASS DISTRIBUTIONS

Dr. Deniz INAN
Oykum Esra ASKIN (PHD.CANDIDATE)
Dr. Ali Hakan BUYUKLU
In many applications the primary endpoint of interest is survival time.

- Medicine, Biology, Public health, Epidemiology, Engineering...

We may be interested in characterizing the distribution of survival time (such as death, going out remission...etc) for a given population;

- Comparing survival times among different groups
- Modelling relationship between survival time and observable covariates
PARAMETRIC SURVIVAL

- In parametric survival model is one in which survival time (the outcome) follow a known distribution;
  - Weibull
  - Exponential
  - Log-logistic
  - Lognormal
  - Generalized gamma
  - ...
- homogeneous population
If we study with heterogeneous population, can we rely on results when we assume the survival time follow pure classical distributions mentioned before?
In recent years, new classes of distributions have been proposed to deal with hardness of modelling heterogeneous data.

Some Examples

- Decreasing Failure Rate
  - exponential-geometric (Adamidis and Loukas, 1998)
  - exponential-Poisson (E-P)(Kus, 2007)
  - exponential-logarithmic (Tahmasbi and Rezaei, 2008)

- Failure Rate with decreasing, increasing and monotone decreasing
  - extended exponential-geometric (Adamidis et al. 2005)
  - weibull-geometric (Barreto-Souza et al., 2010)
  - weibull-logarithmic (Ciumara and Preda, 2009)
  - weibull-Poisson (W-P)(Lu and Shi, 2012)

SIMILAR MIXING PROCEDURE INTRODUCED BY ADAMIDIS AND LOUKAS
OUTLINE OF PRESENTATION

- Compound Poisson Class of Distributions
  - Exponential-zero truncated Poisson (E-P)
  - Weibull-zero truncated Poisson (W-P)
  - Rayleigh-zero truncated Poisson (RAY-P)
- Methodology
  - EM Algorithm
- Application
- Results
- Discussion
Think about a situation where failure (of a device for example) occurs due to the presence of an unknown number, $Z$, of same kind initial defects. Let us define $Z$ as a zero truncated Poisson distributed.

Then let $W$'s represent the failure times of a unit caused by initial defects and each defect can be detected only after causing failure, in which case it is repaired perfectly (Adamidis and Loukas, 1998).

According to $W$'s distributional assumptions ($W_1, W_2, ..., W_z$), we can model time to first failure $X = \text{Min}(W_1, W_2, ..., W_z)$.

In this study, we will take $W$'s as exponential, weibull and rayleigh distributed randoms.
E-P (kus, 2007)

Let $W_1, W_2, \ldots W_Z$ be iid random variables with the following pdf;

$$f(w, \beta) = \beta e^{-\beta w} \quad (1)$$

Also, $Z$ is a zero truncated Poisson variable with following pdf;

$$f(z, \lambda) = \frac{e^{-\lambda} \lambda^z}{\Gamma(z+1)(1-e^{-\lambda})} \quad (2)$$

Let us define $X = \text{min}(W_1, W_2, \ldots W_Z)$. Then, the marginal pdf of $X; \theta = (\lambda, \beta)$

$$f(x, \theta) = \frac{\lambda \beta}{(1-e^{-\lambda})} \exp(-\lambda - \beta x + \lambda \exp(-\beta x)) \quad (3)$$
**W-P (Lu and Shi, 2012)**

\( X = \min(W_1, W_2, \ldots, W_z) \quad \theta = (\lambda, \beta, \alpha) \)

\[
f(x, \theta) = \frac{\lambda \beta \alpha x^{\alpha - 1}}{(1 - e^{-\lambda})} \exp(-\lambda - \beta x^\alpha + \lambda \exp(-\beta x^\alpha))
\]  (4)

**RAY-P (Hemmati et al., 2011)**

\( X = \min(W_1, W_2, \ldots, W_z) \)

\[
f(x, \theta) = \frac{2 \lambda x \beta^2}{(1 - e^{-\lambda})} \exp(-\lambda - (\beta x)^2 + \lambda \exp(-(\beta x)^2))
\]  (5)
for EP and RayP $\lambda = 6; \beta = 2$
for WP $\lambda = 6; \beta = 2; \alpha = 3$
To summarize....

EP
\[ \text{exponential} + \text{zero truncated Poisson} \]

WP
\[ \text{weibull} + \text{zero truncated Poisson distributions,} \]

RayP
\[ \text{Rayleigh} + \text{zero truncated Poisson} \]

with the same mixing procedure
**Methodology**

- To find MLE’s of distribution parameters, Newton Raphson algorithm is one of the standard methods which is widely used. To employ the algorithm, second derivates of the log-likelihood are required.

- However EM algorithm is useful when maximizing observed log likelihood can be difficult then maximizing the complete data log likelihood.

- Recently, EM algorithm has been used by several authors such to find the ML estimations of compound distributions’ parameters.

- We will show the steps of EM algorithm for only WP distribution because of the limited time...
To find hypothetical complete data distribution, it is well known that the conditional density function can be defined as in equation (6). (Alkarni, and Oraby, 2012). Here, $\tau$ is the parameter vector of the weibull distribution.

$$f (x/z; \tau) = z f (x; \tau)\left[1 - F (x; \tau)\right]^{z-1}$$

$$= z\alpha\beta x^{\alpha-1} \exp(-\beta zx^\alpha)$$

Using (6), the hypothetical complete data distribution is given by (7). Here, $\theta$ is the parameter vector of weibull and zero truncated Poisson distributions;

$$f (x, z; \theta) = f (x/z; \tau) \ p (z; \lambda) = \frac{\alpha\beta zx^{\alpha-1} \exp(-\beta zx^\alpha) \lambda^z}{\Gamma(z+1)(\exp(\lambda)-1)}$$

$$x > 0, \ z = 1, 2, ..., \ \lambda, \beta > 0$$
E-step of EM cycle requires the computation of the conditional expectation of $Z$, which is given below:

$$E(Z \setminus X; \theta^{(k)})$$

Here, $\theta^{(k)} = (\lambda^{(k)}, \beta^{(k)}, \alpha^{(k)})$ is the current estimate of $\theta$. Conditional probability of $Z$ can be given as in equation (8).

$$P(z \setminus x; \theta) = \frac{f(x, z; \theta)}{f(x; \theta)} = \frac{\lambda^{z-1} \exp(-\beta z x^\alpha + \beta x^\alpha - \lambda \exp(-\beta x^\alpha))}{\Gamma(z)}$$ (8)

Using equation (8), we can find the conditional expectation of $Z$ for WP distribution as in equation (9).

$$E(z \setminus x; \theta) = \sum_{z=1}^{\infty} z P(z \setminus x; \theta) = 1 + \lambda e^{-\beta x^\alpha}$$ (9)
The EM cycle is completed with M-step. In this step, missing Z’s in complete data likelihood (given in equation (10)) are replaced by their conditional expectations.

\[
\text{C. D. Likelihood: } \prod_{i=1}^{n} \frac{\lambda^{z-1} \exp(-\beta z \alpha + \beta \alpha - \lambda \exp(-\beta \alpha))}{\Gamma(z)}
\]

(10)

Thus, an EM iteration, taking \( \theta^{(k)} \) into \( \theta^{(k+1)} \) is given by:

\[
\alpha^{(k+1)} = n / \left( \sum_{i=1}^{n} \beta^{k+1} \log(x_i) x_i^{\alpha^{(k+1)}} w_i^{k} - \log(x_i) \right)
\]

\[
\beta^{(k+1)} = n / \left( \sum_{i=1}^{n} w_i^{(k)} x_i^{\alpha^{(k+1)}} \right)
\]

\[
\lambda^{(k+1)} = n / \left[ \left( 1 - e^{-\lambda^{(k+1)}} \right) \sum_{i=1}^{n} w_i^{(k)} \right]
\]

\[
w_i^{k} = 1 + \lambda^{(k)} e^{-\beta^{(k)} x_i^{(k)}}
\]
THE FIRST DATA SET

The data concerns 46 observations reported on active repair times (hours) for an airborne communication transceiver.

Data set is used as a lifetime distribution by many authors.
**DATA CHARACTERISTICS**

<table>
<thead>
<tr>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Median</th>
<th>1st quartile</th>
<th>3rd quartile</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>24.5</td>
<td>3.607</td>
<td>1.75</td>
<td>0.800</td>
<td>4.375</td>
<td>2.794666</td>
<td>8.294985</td>
</tr>
</tbody>
</table>
The first data set

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameters</th>
<th>KS Test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP</td>
<td>$\theta: (\lambda = 3.41; \beta = 0.108)$</td>
<td>0.1051</td>
<td>0.6891</td>
</tr>
<tr>
<td>WP</td>
<td>$\theta: (\lambda = 3.52; \beta = 0.09; \alpha = 1.10)$</td>
<td>0.1111</td>
<td>0.6210</td>
</tr>
<tr>
<td>RP</td>
<td>$\theta: (\lambda = 5.92; \beta = 0.11)$</td>
<td>0.3498</td>
<td>$2.5 \times 10^{-5}$</td>
</tr>
</tbody>
</table>
GRAPHS OF PROBABILITY DENSITY FUNCTIONS
### Characteristics of EP distribution

<table>
<thead>
<tr>
<th>E(t)</th>
<th>1st quartile</th>
<th>3rd quartile</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.558</td>
<td>0.780</td>
<td>4.388</td>
<td>2.893</td>
<td>9.297</td>
</tr>
</tbody>
</table>

### Characteristics of WP distribution

<table>
<thead>
<tr>
<th>E(t)</th>
<th>1st quartile</th>
<th>3rd quartile</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.384</td>
<td>0.898</td>
<td>4.304</td>
<td>3.320</td>
<td>17.342</td>
</tr>
</tbody>
</table>

### Characteristics of data1

<table>
<thead>
<tr>
<th>Mean</th>
<th>1st quartile</th>
<th>3rd quartile</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.607</td>
<td>0.800</td>
<td>4.375</td>
<td>2.794666</td>
<td>8.294985</td>
</tr>
</tbody>
</table>
## Bootstrap Confidence Intervals

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Mean</th>
<th>Std.Err.</th>
<th>Bootstrap CI (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EP Distribution</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta = \lambda : \beta$</td>
<td>2.7569</td>
<td>1.8192</td>
<td>(0.0024, 6.949)</td>
</tr>
<tr>
<td></td>
<td>0.1589</td>
<td>0.09487</td>
<td>(0.054, 0.405)</td>
</tr>
<tr>
<td><strong>WP Distribution</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta = \lambda : \beta : \alpha$</td>
<td>3.1947</td>
<td>0.9761</td>
<td>(0.9114, 4.9331)</td>
</tr>
<tr>
<td></td>
<td>0.1069</td>
<td>0.0367</td>
<td>(0.0532, 0.2033)</td>
</tr>
<tr>
<td></td>
<td>1.1303</td>
<td>0.1124</td>
<td>(0.9444, 1.3835)</td>
</tr>
<tr>
<td>Distribution</td>
<td>boot. lambda</td>
<td>boot. beta</td>
<td>boot. alpha</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>EP</td>
<td><img src="image" alt="Histogram of boot. lambda" /></td>
<td><img src="image" alt="Histogram of boot. beta" /></td>
<td><img src="image" alt="Histogram of boot. alpha" /></td>
</tr>
<tr>
<td>WP</td>
<td><img src="image" alt="Histogram of boot. lambda" /></td>
<td><img src="image" alt="Histogram of boot. beta" /></td>
<td><img src="image" alt="Histogram of boot. alpha" /></td>
</tr>
</tbody>
</table>
DISCUSSION

EP or WP?
Ref.


THANK YOU