EXPLORING THE MODIFIED GAMMA FRAILTY DISTRIBUTION: AN OPTIMAL DESIGN APPROACH USING PYTHON

Abdulazeez, Sikiru Adeyinka

Department of Mathematical Sciences, Kaduna State University, Kaduna, Kaduna State, Nigeria

*Corresponding authors’ email: ysabdul94@gmail.com

KEYWORDS: Frailty Models, Censorship, Survival Analysis, Python, Frailty Models

INTRODUCTION

The design of optimal experiments is crucial in the analysis of survival data. When studying time-to-event data or survival data, frailty models, including the Gamma frailty model, are essential for capturing unobserved heterogeneity among subjects. Abdulazeez, (2020). Hazard models have become widespread in their use for the analysis of duration time data in many scientific disciplines, including biology and medicine (Cox, 1972; Kalbfeisch & Prentice, 1980), sociology (Petersen, 1998, Vermunt, 1996), marketing research (Vilcassim & Jain, 1991; Wedel et al., 1995). (Getachew & Bekele 2016) and economics (Kiefer, 1988; Lancaster, 1990). The Cox proportional hazards model (Cox, 1972) is commonly used in the analysis of survival time data. An often unstated assumption of the proportional hazards model is that all individuals will experience the event of interest. However, in some situation a fraction of individuals is not expected to experience the event of interest; that is, these individuals are not at risk. (Anthony et al. 2019). The terminology to describe the never-at-risk group varies from field to field, but includes ‘long-term survivors’ or ‘cured’ in epidemiology, ‘non-susceptibles’ in toxicology, ‘stayers’ in finite Markov transition models of occupational mobility, the ‘non-fecundable’ in fertility models, and ‘non-recidivists’ among convicted criminals. In epidemiology and medicine, researchers may be interested in analyzing the occurrence of a disease. Many individuals may never experience that disease; therefore, there exists a fraction in the population that is protected. Cure models are survival models which allow for the presence of cure in the study population. These models extend the understanding of time-to-event data by allowing for the formulation of more accurate and informative conclusions than previously made. These conclusions would otherwise be unobtainable from an analysis that fails to account for the cured fraction in the population. If a cured component is not present, the analysis reduces to standard approaches of survival analysis. In cure models, the population is divided into two subpopulations so that an individual either cured with probability 1 − φ, or has a proper survival function S(t), with probability φ. Here, proper survival function means \( \lim_{t \to \infty} S(t) = 0 \). Individuals regarded as cured will never experience the event of interest and their survival time will be defined as infinity. Therefore, the hazard and survival functions of cured individuals are set to zero and one, respectively, for all finite values of t. Longini and Halloran (1996) have proposed frailty cure models that extend standard frailty models. The frailty random variable in the former has point mass at zero with probability 1 − φ while heterogeneity among those experiencing the event of interest is modelled via a continuous distribution with probability φ. Price and Manatunga (2001) gave an excellent introduction to this area and applied leukaemia remission data to different cure, frailty and frailty cure models. They found that frailty models are useful in modeling data with a cured fraction and that the gamma frailty cure model provides a better fit to their remission data compared to the standard cure model.
In the next section we describe the existing models and a proposed model, then provide an application of the models to an existing data on occupational exposure tagged – IRANIAN data.

MATERIALS AND METHODS

Cox PH models

The notation used for Cox PH models (Cox, 1972), Lee & Song (2001) with one more subscript to capture multiple events is generalized. Let \( T_{ik} \) be the total time of the \( k^{th} \) event for the \( i^{th} \) subject, \( C_{ik} \) be the censoring time of the \( k^{th} \) event for the \( i^{th} \) subject. Let \( U_{ik} \) be the observation time, that is, \( U_{ik} = \min(T_{ik}, C_{ik}) \).

and

\[
Z_i = \frac{\lambda_0}{\lambda_1} Y_0 + Y_1 \sim \Gamma(k_0 + k_1, \lambda_1)
\]

\[
Z_i = \frac{\lambda_0}{\lambda_2} Y_0 + Y_2 \sim \Gamma(k_0 + k_2, \lambda_2)
\]

\[
E(Z_1) = E(Z_2) = 1,
\]

\[
V(Z_1) = \frac{1}{\lambda_1^2},
\]

\[
V(Z_2) = \frac{1}{\lambda_2^2}.
\]

The following relation holds

\[
E(Y_0^2) = V(Y_0) + (E(Y_0))^2
\]

\[
= \frac{k_0 + (k_0^2)}{\lambda_0^2},
\]

\[
= \frac{k_0^2 + k_0}{\lambda_0^2} + 1.
\]

Consequently, because of relation

\[
k_0 + k_1 = k_1 = \lambda_1
\]

\[
= \frac{1}{\sigma_i^2} (i, j = 1,2; i \neq j).
\]

Correlated Gamma Frailty (CGF) Model


Let \( k_0, k_1, k_2 \) be some real positive values. Set \( \lambda_1 = k_0 + k_1, \lambda_2 = k_0 + k_2 \).

Let \( Y_0, Y_1, Y_2 \) be independently gamma distributed random variables with

\[
Y_0 \sim \Gamma(k_0, \lambda_0), Y_1 \sim \Gamma(k_1, \lambda_1), Y_2 \sim \Gamma(k_2, \lambda_2)
\]

Consequently,

\[
S(t_1, t_2) = E[S(t_1, t_2, |Z_1, Z_2)]
\]

\[
= E\left[\sum_i \min(t_1, Z_i, t_2, Z_i)\right]
\]

\[
= E\left[e^{\sum_i \Lambda_i(t_1) e^{-\sum_i \Lambda_i(t_2)}} \right]
\]

\[
\begin{align*}
S(t_1, t_2) &= E\left\{ e^{-\sum_i \lambda_i(t_1)} e^{-\sum_i \lambda_i(t_2)} \right\} \\
&= E\left\{ e^{-\sum_i \lambda_i(t_1)} e^{-\sum_i \lambda_i(t_2)} \right\} \\
&= \left( 1 + \frac{\lambda_0}{\lambda_1} \Lambda_1(t_1) + \frac{\lambda_0}{\lambda_2} \Lambda_2(t_2) \right)^{-k_0} \left( 1 + \frac{\lambda_1}{\lambda_1} \Lambda_1(t_1) + \frac{\lambda_2}{\lambda_2} \Lambda_2(t_2) \right)^{-k_1} \\
&= \left( 1 + \frac{\lambda_0}{\lambda_1} \Lambda_1(t_1) + \frac{\lambda_0}{\lambda_2} \Lambda_2(t_2) \right)^{-k_0} \left( 1 + \frac{\lambda_1}{\lambda_1} \Lambda_1(t_1) + \frac{\lambda_2}{\lambda_2} \Lambda_2(t_2) \right)^{-k_1}
\end{align*}
\]
which results in the representation of the Correlated Gamma Frailty model given as

\[
S(t_1, t_2) = \frac{\gamma(t_1)^{\rho_1} \gamma(t_2)^{\rho_2}}{\gamma(t_1 + t_2)^{\rho_1 + \rho_2}}
\]

(17)

**The Proposed Model – Modified Gamma Frailty (MGF) Model**

In order to include heterogeneity in the model, we assume a correlated gamma frailty model. Let \( z_i (j = 1, 2) \) be the frailties, and \( x_i (j = 1, 2) \) vectors of observable covariates of the two individuals of a twin pair. Assume that their individual hazards are represented by the proportional hazards model

\[
\lambda(t) = \alpha(t) \exp(\beta X j)(j = 1, 2)
\]

(18)

with a baseline hazard function \( \alpha(t) \) describing the risk of respiratory infection as a function of age and \( \beta \) denotes the vector of regression parameters. Let the lifetimes of the two twin partners be conditionally independent given their frailties \( z_1 \) and \( z_2 \). Because frailties \( X j (j = 1, 2) \) are usually unobservable, their correlation coefficient used cannot be estimated directly from the empirical data. So a bivariate lifetime model which allows indirect calculation of the parameters is needed. The unconditional bivariate survival function of the correlated gamma frailty model with observed covariates is given by:

\[
S(t_1, t_2) = S(t_1 \mid X_1) \cdot S(t_2 \mid X_2) = \gamma(t_1)^{\rho_1} \gamma(t_2)^{\rho_2} \gamma(t_1 + t_2)^{\rho_1 + \rho_2}
\]

(19)

Where \( S(t \mid X) \) denotes the marginal univariate survival function, assumed to be equal for both partners in a twin pair. Using a parametric approach we fit a model to the data, such that

\[
S(t | X_{ijk}) = \left[1 + \left(1 + \sigma_i^2 \beta \gamma \left(e^{bt} - 1\right) \right)^{\sigma_i^2} \right] \gamma \left( e^{bt} \right)
\]

(20)

Where \( \sigma_i^2, \sigma_j^2, \beta, \text{and} \rho \) are parameters to be estimated. The lifetimes are assumed to be independently censored from the right by independent and identically distributed pairs of non-negative random variables, which are independent of the lifetimes. Thus, observe

\[
(T_{1i}, T_{2i}, \Delta_{1i}, \Delta_{2i}, X_{1i}, X_{2i})
\]

(21)

with \( \Delta_{1i}(i = 1, 2, \ldots; m; j = 1, 2) \) as a binary variable with values 1 (event) and 0 (no event). Let the lifetimes follow a distribution (dependent on covariates \( X_{1i}, X_{2i} \)) given by the bivariate survival function

\[
S(t_1, t_2 | X_{1i}, X_{2i}) = P(T_{1i} > t_1, T_{2i} > t_2 | X_{1i}, X_{2i})
\]

(22)

Starting from this model, we are able to derive the likelihood function given by

\[
L(t_1, t_2, \delta_1, \delta_2, X_{1i}, X_{2i}) = \delta_1 \delta_2 S(t_1 | t_2 | X_{1i}, X_{2i}) - \delta_1 (1 - \delta_2) S(t_1, t_2 | X_{1i}, X_{2i}) - (1 - \delta_1) \delta_2 S(t_1 | t_2 | X_{1i}, X_{2i}) + (1 - \delta_1)(1 - \delta_2) S(t_1, t_2 | X_{1i}, X_{2i})
\]

(23)

Partial derivatives of the marginal survival functions are given by

\[
S_i(t_1, t_2) = \frac{\partial S(t_1, t_2)}{\partial t_j}
\]

(24)

and

\[
S_i(t_1, t_2) = \frac{\partial S(t_1, t_2)}{\partial t_i}
\]

(25)

The model is called the **Modified Gamma Frailty (MGF) Model.**

**Numerical Illustration**

An application of the models to an existing data on occupational exposure tagged - IRANIAN data is demonstrated here. Relationships between occupational exposures and morbidity, morbidity and job category were analyzed using proportional hazard analysis, allowing for exposure status (never exposed, ever smoked and ever exposed) until the time of carrying out the study. The survival-analysis was performed using the Python programming. Below is Python code framework to estimate parameters in the context of a gamma frailty model using the survival package:

```python
pip install lifelines
import pandas as pd
import numpy as np
from lifelines import CoxPHFitter

# Simulate some survival data with gamma frailty (simplified for illustration)
np.random.seed(123)
n = 200
clusters = np.repeat(np.arange(n // 2, 2))
frailty_effect = np.repeat(np.random.gamma(1, 1, n // 2, 2))
baseline_hazard = 0.05
time = np.log(np.random.uniform(0, 1, n)) / (baseline_hazard * frailty_effect)
censoring_time = np.random.exponential(1.0/0.3, n)
event = (time <= censoring_time).astype(int)
obs_time = np.minimum(time, censoring_time)
data = pd.DataFrame({'id': clusters, 'obs_time': obs_time, 'event': event})

cph = CoxPHFitter().fit(data, 'obs_time', 'event', cluster_col='id')
print(cph.print_summary())
```

The code above uses the `lifelines` package in Python to perform Cox regression analysis with frailty effects.
EXPLORING THE MODIFIED GAMMA...  

Abdulazeez  

import statsmodels.api as sm  
from lifelines.datasets import load_dd  

# Load a sample dataset from lifelines for demonstration  
data = load_dd()  

# Here, the dataset has the columns 'duration' and 'observed'.  
# 'duration' is the observed duration and 'observed' is a binary column indicating if the event was observed or not.  

# To fit a PH model with gamma shared frailty:  
  fml = "duration ~ age + education + np.log(1 + income) + np.log(1 + income)**2"  
  phf = sm.PHReg(data["duration"], data[fml], status=data["observed"], ties="efron", strata=data["strata"])
  result = phf.fit()  
  print(result.summary())

# The output will provide parameter estimates, p-values

The discrete algorithm was used, since the time-scale (person-years) was discrete. All exposures were first analyzed separately, allowing for age and smoking habits. Two-sided p-values < 0.05 were considered as statistically significant. The relationship between occupational exposures and morbidity was also analyzed simultaneously. Using the stepwise option of Python programming, and allowing for age and smoking habits, specific exposures were included and excluded until the following conditions were met: the significance of the residual Chi-squared was less than 0.25, and the significance of the relative risks was less than 0.10. Using the standard error of the regression coefficient, the 95% confidence intervals were estimated. The Python programming was also applied in analyzing the Correlated Gamma Frailty Model and the Modified Gamma Frailty Model. Hazard function and survival functions for the exposure data for large and small samples were estimated.

RESULTS AND DISCUSSION

Table 1 shows the results of analysis of the Iranian data and the goodness of fit table. The exponentiated coefficients in the third column of each table of the output shown are interpretable as multiplicative effects on the hazard. In tables 1, for example, holding the other covariates constant, one additional year of age increases the yearly hazard of exposure of worker by a factor of $e^{\beta} = 1.053376$ on average – that is, by 5.3 percent. Similarly, each Forced Ventilatory Function (FVC) factor increases the hazard by a factor of 1.059079 or 5.9 percent.

The fifth column is the result of the test of significance of $\beta$ using the Wald Statistic which is the ratio of the coefficients to the standard error of $\beta$. The obtained value is compared with the Z value and a decision is made. In table 2, holding the other covariates constant, an additional year of age increases the yearly hazard of exposure of worker by a factor of $e^{\beta} = 1.034585$ on average – that is, by 3.5 percent. Similarly, each FVC factor increases the hazard by a factor of 1.001301 or 0.1 percent.

In table 3, holding the other covariates constant, an additional year of age increases the yearly hazard of exposure of worker by a factor of $e^{\beta} = 1.053481$ on average – that is, by 5.3 percent. Similarly, each FVC factor increases the hazard by a factor of 1.062155 or 6.2 percent.

The exposure status (never exposed, exposed and ever smoked), Job category and pack years smoked is considered to be insignificant for the Iranian data using the Cox Model. The CGF captures the exposure status and Job category to be insignificant for the Iranian data while the proposed MGF considers all the variables to be significant for the Iranian data.

Table 1: Regression Coefficients in the Cox Model for the Iranian Study

<table>
<thead>
<tr>
<th>Covariate</th>
<th>coeff((\beta))</th>
<th>Exp(coeff((\beta)))</th>
<th>Std error coeff((\beta))</th>
<th>Z</th>
<th>P</th>
<th>95% C.I for coeff((\beta))</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>0.0520</td>
<td>1.053376</td>
<td>0.0227</td>
<td>2.290749</td>
<td>0.0010</td>
<td>1.0212 - 1.0828</td>
</tr>
<tr>
<td>BMI</td>
<td>0.0555</td>
<td>1.057069</td>
<td>0.0119</td>
<td>4.663866</td>
<td>0.0010</td>
<td>1.0321 - 1.0788</td>
</tr>
<tr>
<td>EXPOSURE STATUS</td>
<td>-0.1498</td>
<td>0.86088</td>
<td>0.2122</td>
<td>-0.70594</td>
<td>0.2145*</td>
<td>0.1531 - 1.9124</td>
</tr>
<tr>
<td>JOB CATEGORY</td>
<td>0.4337</td>
<td>1.542956</td>
<td>0.3819</td>
<td>1.135638</td>
<td>0.1041*</td>
<td>0.2009 - 1.6029</td>
</tr>
<tr>
<td>SYST B P</td>
<td>0.0915</td>
<td>1.099581</td>
<td>0.0286</td>
<td>3.199301</td>
<td>0.0029</td>
<td>1.0899 - 1.1103</td>
</tr>
<tr>
<td>PACK YRS SMOKED</td>
<td>-0.2038</td>
<td>0.815625</td>
<td>0.1914</td>
<td>-1.06479</td>
<td>0.9856*</td>
<td>0.2133 - 1.9339</td>
</tr>
<tr>
<td>FVC</td>
<td>0.0574</td>
<td>1.059079</td>
<td>0.0221</td>
<td>2.597285</td>
<td>0.0085</td>
<td>1.0444 - 1.0679</td>
</tr>
<tr>
<td>FEV1</td>
<td>0.0849</td>
<td>1.088608</td>
<td>0.1956</td>
<td>0.434049</td>
<td>0.0007</td>
<td>1.0415 - 1.0940</td>
</tr>
</tbody>
</table>

* Not significant.

Table 2: Regression Coefficients in the Correlated Gamma Frailty Model for the Iranian Study

<table>
<thead>
<tr>
<th>Covariate</th>
<th>coeff((\beta))</th>
<th>Exp(coeff((\beta)))</th>
<th>Std error coeff((\beta))</th>
<th>Z</th>
<th>P</th>
<th>95% C.I for coeff((\beta))</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>0.0340</td>
<td>1.034585</td>
<td>0.0154</td>
<td>2.207792</td>
<td>0.0011</td>
<td>1.0112 - 1.0542</td>
</tr>
<tr>
<td>BMI</td>
<td>0.0437</td>
<td>1.044669</td>
<td>0.0231</td>
<td>1.891775</td>
<td>0.0053</td>
<td>1.0221 - 1.0658</td>
</tr>
<tr>
<td>EXPOSURE STATUS</td>
<td>-0.0249</td>
<td>0.975407</td>
<td>0.0713</td>
<td>-0.34923</td>
<td>0.6714*</td>
<td>0.0313 - 1.0224</td>
</tr>
<tr>
<td>JOB CATEGORY</td>
<td>0.0023</td>
<td>1.002303</td>
<td>0.0044</td>
<td>0.522727</td>
<td>0.1304*</td>
<td>0.0019 - 1.0329</td>
</tr>
<tr>
<td>SYST B P</td>
<td>0.0021</td>
<td>1.002102</td>
<td>0.0009</td>
<td>2.333333</td>
<td>0.0039</td>
<td>1.0001 - 1.0030</td>
</tr>
<tr>
<td>PACK YRS SMOKED</td>
<td>-0.2138</td>
<td>0.80751</td>
<td>0.0371</td>
<td>-5.7628</td>
<td>0.0035</td>
<td>0.033 - 0.9339</td>
</tr>
</tbody>
</table>
EXPLORING THE MODIFIED GAMMA... Abdulazeez FJS

FVC  0.0013  1.001301  0.0006  2.16667  0.0018  1.0003 - 1.0489
FEV\textsubscript{1}  0.0362  1.036863  0.0146  2.479452  0.0029  1.0235 - 1.0440

* Not significant.

Table 3: Regression Coefficients in the Modified Gamma Frailty Model for the Iranian Study

<table>
<thead>
<tr>
<th>Covariate</th>
<th>coeff(β)</th>
<th>Exp(coeff(β))</th>
<th>Std error coeff(β)</th>
<th>Z</th>
<th>P</th>
<th>HR† (CI§ 95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>0.0521</td>
<td>1.053481</td>
<td>0.0167</td>
<td>3.11976</td>
<td>0.0011</td>
<td>1.0215 - 1.0848</td>
</tr>
<tr>
<td>BMI</td>
<td>0.0547</td>
<td>1.056224</td>
<td>0.0111</td>
<td>4.927928</td>
<td>0.0001</td>
<td>1.0324 - 1.0798</td>
</tr>
<tr>
<td>EXPOSURE STATUS</td>
<td>-0.1049</td>
<td>0.900415</td>
<td>0.2013</td>
<td>-0.52111</td>
<td>0.0045</td>
<td>0.0312 - 0.9814</td>
</tr>
<tr>
<td>JOB CATEGORY</td>
<td>0.3987</td>
<td>1.489887</td>
<td>0.4009</td>
<td>0.994512</td>
<td>0.0041</td>
<td>1.0009 - 1.6029</td>
</tr>
<tr>
<td>SYST B P</td>
<td>0.0891</td>
<td>1.09319</td>
<td>0.0271</td>
<td>3.287823</td>
<td>0.0021</td>
<td>1.0021 - 1.5003</td>
</tr>
<tr>
<td>PACK YRS SMOKED</td>
<td>-0.2038</td>
<td>0.815625</td>
<td>0.1717</td>
<td>-1.18695</td>
<td>0.0056</td>
<td>0.363 - 0.9939</td>
</tr>
<tr>
<td>FVC</td>
<td>0.0603</td>
<td>1.062155</td>
<td>0.0211</td>
<td>2.85782</td>
<td>0.0015</td>
<td>1.0044 - 1.2479</td>
</tr>
<tr>
<td>FEV\textsubscript{1}</td>
<td>0.0762</td>
<td>1.0797178</td>
<td>0.0377</td>
<td>2.02122</td>
<td>0.0029</td>
<td>1.0135 - 1.8340</td>
</tr>
</tbody>
</table>

* Not significant.

Table 4: Prognostic Factors of Occupational Exposure using Cox and frailty Models for Iranian study

<table>
<thead>
<tr>
<th>Prognostic factors</th>
<th>Cox regression</th>
<th>Correlated Gamma</th>
<th>Modified Gamma Frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (CI 95%)</td>
<td>HR† (CI§ 95%)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.0530 (1.0212 - 1.0828)</td>
<td>1.0346 (1.0112 - 1.0542)</td>
<td>1.053481 (1.0215 - 1.0848)</td>
</tr>
<tr>
<td>BMI</td>
<td>1.0571 (1.0321 - 1.0788)</td>
<td>1.0447 (1.0221 - 1.0658)</td>
<td>1.056224 (1.0324 - 1.0798)</td>
</tr>
<tr>
<td>EXPOSURE STATUS</td>
<td>0.8609 (0.8153 - 0.9124)</td>
<td>0.9754 (0.9013 - 0.9254)</td>
<td>0.900415 (0.8512 - 0.9814)</td>
</tr>
<tr>
<td>JOB CATEGORY</td>
<td>1.5430 (1.0209 - 1.6029)</td>
<td>1.0023 (0.9919 - 1.0129)</td>
<td>1.489891 (1.0009 - 1.6029)</td>
</tr>
<tr>
<td>SYST B P</td>
<td>1.0998 (1.0899 - 1.1103)</td>
<td>1.0021 (1.0001 - 1.0030)</td>
<td>1.09321 (1.0021 - 1.5003)</td>
</tr>
<tr>
<td>PACK YRS SMOKED</td>
<td>0.8156 (0.8133 - 0.9339)</td>
<td>0.8075 (0.7533 - 0.9339)</td>
<td>0.8156 (0.633 - 0.9939)</td>
</tr>
<tr>
<td>FVC</td>
<td>1.0591 (1.0444 - 1.0679)</td>
<td>1.0013 (1.0003 - 1.0489)</td>
<td>1.062155 (1.0044 - 1.2479)</td>
</tr>
<tr>
<td>FEV\textsubscript{1}</td>
<td>1.0886 (1.0415 - 1.0940)</td>
<td>1.0369 (1.0235 - 1.0440)</td>
<td>1.079176 (1.0135 - 1.8340)</td>
</tr>
<tr>
<td>AIC#</td>
<td>1.157</td>
<td>751</td>
<td>704</td>
</tr>
</tbody>
</table>

† Hazard Ratio § Confidence interval * Not significant # Akaike Information Criterion

Figure 1: Survival Function at mean of covariates - Iranian Study

Figure 2: Hazard Function at mean of covariates - Iranian Study

CONCLUSION

Interestingly, parameter estimates are quite different depending on distribution of the base-line hazard function. The newly introduced Modified Gamma frailty model offers a very elegant approach to integrate the concept of clusters into frailty modelling. The survival function is explicitly available and of easy form which allows traditional maximum likelihood parameter estimation. This is the most important advantage of the suggested model compared to the model introduced by Moger & Aalen (2005). Our simulation study revealed insights into the properties of the estimator under the modified gamma frailty model.

The present work contributes to three aspects of Frailty models with censored data. First, It presents several important extensions of the existing models. Secondly, It develops a general asymptotic theory for the Frailty models. Thirdly, It provides simple and efficient numerical method to implement the corresponding inference procedures. It is hoped that this work will facilitate further development and applications of Frailty models. It has been demonstrated that the MGF is a very general and powerful approach to the analysis of Frailty models with censored data. This approach can be used to study many other problems. Of great interest would be a non-parametric version...
of the correlated compound Poisson frailty model, where the baseline hazard functions are not specified. A part of future research is envisaged in this direction. Another aspect that will be of interest for further research is the problem of identifiability. The identifiability problem is growing with increased censoring, but is reduced by the parametric modelling of the baseline hazard. This study furnishes a structured approach for optimal experiment design using the modified gamma frailty distribution, supported by a demonstrative Python-based simulation.

REFERENCES


©2023 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via [https://creativecommons.org/licenses/by/4.0](https://creativecommons.org/licenses/by/4.0) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.