Journal of Statistical Software

MMMMMM YYYY, Volume VV, Issue II. http://www.jstatsoft.org/

frailtyr: A General Semiparametric Shared Frailty Model

John V. Monaco
Pace University

Malka Gorfine
Tel Aviv University

Li Hsu
Fred Hutchinson
Cancer Research Center

Abstract

TODO

Keywords: survival analysis, shared frailty model, R.

1. Introduction

Based on Gorfine, Zucker, and Hsu (2006). Background, motivation, survival packages.

Table 1: R functions for fitting survival models. P=parametric, NP=nonparametric, PPL=partial penalized likelihood, MML=maximum marginal likelihood, DLs=downloads.

Package::function	λ_0	Estimation	Gen?	Frailty distr	Weekly DLs
survival::coxph	NP	PPL	No	Γ, LN, LT	2003.1
coxme::coxme	NP	PPL	No	LN	167.4
frailtypack::frailtyPenal	P or NP	PPL	No	Γ, LN,	80.3
phmm::phmm	NP	EM	No	LN,	46
parfm::parfm	P	MML	No	Γ, PS, IG	45.9
survBayes::survBayes	NP	Bayes	No	Γ, LN	35.7

Additional: survMisc, survsim, MST.

2. Data generation

Data can be generated under a wide variety of conditions to simulate survival times. The genfrail function in frailtyr is used to generate survival times from a shared frailty model.

That is, we have survival function

$$S(t|\mathbf{Z}_{ij}, \omega_i) = \exp\left\{-\Lambda_0(t)\,\omega_i \exp(\beta'\mathbf{Z}_{ij})\right\} \tag{1}$$

where t is time, ω_i is the frailty value for cluster i, $\Lambda_0(t)$ is the unspecified cumulative baseline hazard function, β is the coefficient vector, and $\mathbf{Z}_{ij}$ is the vector of covariates for member j in cluster i. We let $\omega \sim f(\theta)$, where θ is a parameter vector for frailty distribution f, and $Z \sim U(0,1)$ or $Z \sim N(0,1)$. Covariates are draw from either a standard normal or uniform distribution, as specified by the covariates parameter.

2.1. Baseline hazard

There are three ways the baseline hazard can be specified to generate survival data: as the inverse cumulative baseline hazard Λ_0^{-1}, the cumulative baseline hazard Λ_0, or the baseline hazard λ_0. Due to Bender Bender, Augustin, and Blettner (2005), if the cumulative baseline hazard function can be directly inverted, then the failure times can solved directly by

$$T_{ij}^0 = \Lambda_0^{-1}\left[-\ln(u)\exp\left(-\beta'\mathbf{X}_{ij}\right)\right] \tag{2}$$

where $u \sim U(0,1)$ and T_{ij}^0 is the failure time for member j in cluster i. Consequently, if Λ_0^{-1} is provided, then survival times are determined by Equation (2).

When Λ_0 cannot be inverted, we must solve the following equation for t to determine failure time T_{ij}^0,

$$S(t|\mathbf{Z}_{ij}, \omega_i) - u = 0$$

using a univariate root-finding algorithm as described in Crowther and Lambert (2013). Alternatively, we may take the logarithm and solve

$$-\Lambda_0(t)\,\omega_i \exp(\beta'\mathbf{Z}_{ij}) - \ln u = 0 \tag{3}$$

for greater numerical stability in the root-finding algorithm. genfrail takes this approach when Λ_0 is provided and uses the R function uniroot to solve Equation (3), which is based on Brent's algorithm Brent (2013). The resulting failure times take on values $[0, \infty)$ for $u \in (0, 1]$.

If neither Λ_0^{-1} or Λ_0 are provided to genfrail, then the baseline hazard function λ_0 must be passed. In this case, we have

$$\Lambda_0(t) = \int_0^t \lambda_0(s)\,ds$$

which must be evaluated numerically. Using the numerical integrations, Equation (3) can be solved for t. This approach is the most computationally expensive since it requires numerical integration to be performed at each iteration in the root-finding algorithm, but allows for generating data from arbitrary baseline hazard functions. Section 2.6 demonstrates generating data using each method.

2.2. Shared frailty

Shared frailty values are generated for each cluster according to the frailty distribution specified. Gamma variates are generating using the R stats function rgamma, which relies on two different techniques depending on whether $\theta > 1$. For other distributions with Laplace

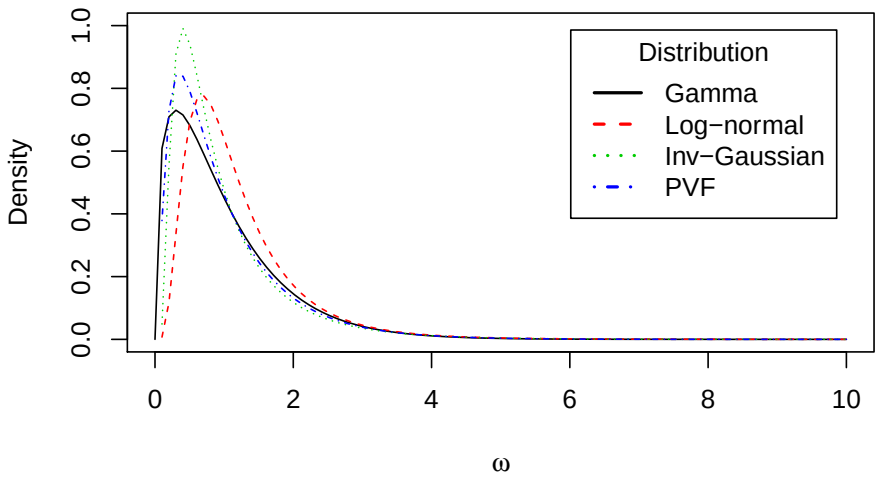

Figure 1: Frailty distribution densities with $\mathsf{VAR}[\omega] = 0.7$. Gamma, Inv-Gaussian, and PVF have $\mathsf{E}[\omega] = 1$ and log-normal has $\mathsf{E}[\omega] \approx 1.168$.

transforms, frailty values are generated using a modified Newton-Raphson algorithm for numerical transform inversionRidout (2009). This technique is used for the PVF and positive stable distributions. Note that while frailtyr can generate survival data with a positive stable frailty distribution, it cannot estimate parameters for this model due to the positive stable having an infinite mean. Frailty values from a log-normal distribution are generated in the usual way, and inverse Gaussian variates are generated using a transformation method in the statmod package Smyth, Hu, Dunn, Phipson, and Chen (2015). See Appendix A for a full list of frailty distributions and parametrization.

For the gamma, log-normal, and inverse Gaussian, there is a positive relationship between the distribution parameter θ and the strength of dependence between cluster members. As θ increases, inter-cluster differences decrease and intra-cluster differences increase. The opposite is true for the PVF, and as α increases, the dependence between cluster members decreases. All the supported frailty distributions have degenerate cases, which are defined in Appendix A.

Figure 1 shows the densities for the supported distributions, all with 0.7 variance. The gamma and inverse Gaussian have variance $\mathsf{VAR}[\omega] = \theta = 0.7$, while the PVF has $\mathsf{VAR}[\omega] = 1 - \alpha = 0.7$ for $\alpha = 0.3$. For the log-normal $\mathsf{VAR}[\omega] = e^{\theta}\left(e^{\theta} - 1\right) = 0.7$ must be solved for θ to get $\theta \approx 0.3884$.

2.3. Cluster sizes

Depending on how a cluster is defined, the cluster sizes in a shared frailty dataset will vary. For example, in the Diabetic Retinopathy Study (DRS), two failure times are observed for each subject corresponding to the left and right eye Group *et al.* (1976); Huster, Brookmeyer, and

Self (1989). Observations are clustered by subject, so every cluster has exactly two members. If instead the observations were clustered by geographical location, the cluster sizes would vary and perhaps follow a discrete power law. genfrail is able to generate data with fixed or varying cluster sizes.

Cluster sizes can be generated from a k-truncated Poisson Geyer. (2014), where the expected cluster size is given by

$$\mathsf{E}\left[K\right] = \frac{\lambda - e^{-\lambda}\sum_{j=1}^{l}\frac{\lambda^j}{(j-1)!}}{1 - e^{-\lambda}\sum_{j=0}^{l}\frac{\lambda^j}{j!}}$$

The typical case is with $l = 0$ for a zero-truncated Poisson. If we also have $\lambda = 2$, then $\mathsf{E}[K] \approx 2.313$. The parameters for the k-truncated Poisson are c(l, lambda).

A discrete zeta distribution can also be used to generate cluster sizes. Accurately fitting and generating from a discrete power-law distribution is generally difficult Clauset, Shalizi, and Newman (2009), and genfrail uses a truncated Zeta distribution to avoid some of the pitfalls. The density is given by

$$f\left(K; s, u, l\right) = \frac{\left(K-l\right)^{-s}/\zeta\left(s\right)}{\sum_{j=1}^{u-l} j^{-s}/\zeta\left(s\right)} \quad s > 1$$

where $\zeta\left(s\right)$ is the Riemann zeta function, s is the scaling exponent, l is the noninclusive lower bound, and u is the inclusive upper bound. The resulting distribution has support $K \in \{l+1, \ldots, u\}$. With $s \gg 1$, large enough u, the distribution behaves similar to the zeta distribution, where $\mathsf{E}\left[K\right] = \frac{1}{\zeta(s)}\sum_{j=1}^{\infty}\frac{1}{j^{s-1}}$.

Finally, a discrete uniform distribution can be specified by passing K = "duniform" to genfrail. The respective parameters are c(l, u), where l is the noninclusive lower bound and u is the inclusive upper bound. Similar to the truncated zeta, $K \in \{l+1, \ldots, u\}$ and cluster sizes will be selected from this set. Since the lower bound is noninclusive, $\mathsf{E}\left[K\right] = \frac{1+l+u}{2}$.

The cluster sizes may also be specified by the user. In this case, a vector of the desired cluster sizes should be passed as parameter K. In this case, N is ignored with a warning.

2.4. Censoring

The observed follow-up times and failure indicators are determined by right-censoring the failure times. Let the observed follow-up time for member j in cluster i be given by

$$T_{ij} = \min\left(T_{ij}^0, C_{ij}\right)$$

where C_{ij} is the censoring time. The failure indicator is given by

$$\delta_{ij} = I\left(T_{ij}^0 \leq C_{ij}\right)$$

The censoring distribution is specified through the parameters censoring.distr and censoring.param for the distribution name and parameter vector, respectfully. A normal distribution is used by default.

Sometimes, a particular censoring rate is desired. Typically, the censoring distribution parameters are varied to obtain a desired censoring rate. genfrail can avoid this effort on behalf of the user by letting the desired censoring rate be specified instead. In this case, the appropriate parameters for the censoring distribution are determined to achieve the desired censoring rate, given the generated failure times.

Let $\dot{F}(t)$ be the empirical cumulative distribution of failure times. With a normal censoring distribution, the expected censoring rate is

$$C(\mu) = \int_0^\infty \mathcal{N}\left(t; \mu, \sigma^2\right) F(t)\, dt$$

where $C(\mu)$ is the empirical censoring rate as a function of μ. To obtain a particular censoring rate r, we must solve the equation

$$r - C(\mu) = 0$$

for μ. The censoring variance σ^2 must be specified for the problem to be identifiable. Solving for μ will give the censoring distribution with expected censoring rate r for the data generated. Thus, this method works with any empirical distribution of failure times. genfrail uses this method to achieve a desired censoring rate with either normal or log-normal censoring distributions.

2.5. Rounding

The observed times may optionally be rounded to the nearest multiple of B by

$$\dot{T}_{ij} = B\left\lfloor \frac{T_{ij}}{B} + 0.5 \right\rfloor$$

This has the effect of simulating multiple observations (and thus, multiple failures depending on the failure indicator) at a given time step when there are many subjects and a course-grained measurement of time. If $B = 1$, the observed times are simply rounded to the nearest integer.

2.6. Examples

Generating survival data from a model with known parameters.

$$S(t|\mathbf{Z}, \omega) = \exp\left\{-\omega \exp(\beta'\mathbf{Z})(0.01t)^{4.6}\right\}$$

$$\lambda_0(t) = \{\tau (Ct)^\tau\} t^{-1} \tag{4}$$

Oscillating BH:

$$\lambda_0(t) = A^{sin(f\pi t)} \{\tau (Ct)^\tau\} t^{-1} \tag{5}$$

The baseline hazard can be verified by passing a zero covariate coefficient and no censoring. In this case, the empirical failure rate should reflect the baseline hazard.

3. Model estimation

$$\ell = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \delta_{ij} \log e^{\underline{\beta}^T \underline{Z}_{ij}} + \sum_{i=1}^{n} \log \mathcal{L}^{(m_i)}(H_{i\cdot}(\tau))$$

$$\begin{aligned}\underline{U}_{\underline{\beta}} = \frac{\partial}{\partial \underline{\beta}}\ell &= \sum_{i=1}^{n}\left[\sum_{j=1}^{m_i}\delta_{ij}\underline{Z}_{ij} + \frac{\frac{\partial}{\partial \underline{\beta}}H_{i.}(\tau)\,\frac{\partial}{\partial H_{i.}(\tau)}\mathcal{L}^{(m_i)}(H_{i.}(\tau))}{\mathcal{L}^{(m_i)}(H_{i.}(\tau))}\right] \\ &= \sum_{i=1}^{n}\left[\sum_{j=1}^{m_i}\delta_{ij}\underline{Z}_{ij} + \sum_{j=1}^{m_i}H_{ij}(T_{ij})\,\underline{Z}_{ij}\frac{\mathcal{L}^{(m_i+1)}(H_{i.}(\tau))}{\mathcal{L}^{(m_i)}(H_{i.}(\tau))}\right]\end{aligned}$$

$$U_\theta = \frac{\partial}{\partial\theta}\ell = \sum_{i=1}^{n}\frac{\frac{\partial}{\partial\theta}\mathcal{L}^{(m_i)}(H_{i.}(\tau))}{\mathcal{L}^{(m_i)}(H_{i.}(\tau))}$$

Maximize likelihood vs score equations
Likelihood

$$L = \prod_{i=1}^{n}\prod_{j=1}^{m_i}\left\{\lambda_0(T_{ij})\exp\left(\beta^T Z_{ij}\right)\right\}^{\delta_{ij}}\prod_{i=1}^{n}\int\omega^{N_{i.}(\tau)+1}\exp\{-\omega H_{i.}(\tau)\}f(\omega)\,d\omega \tag{6}$$

Loglikelihood

$$l = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\delta_{ij}\log\left\{\lambda_0(T_{ij})\exp\left(\beta^T Z_{ij}\right)\right\} + \sum_{i=1}^{n}\log\left[\int\omega^{N_{i.}(\tau)+1}\exp\{-\omega H_{i.}(\tau)\}f(\omega)\,d\omega\right] \tag{7}$$

Coefficient score

$$U_r = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\delta_{ij}Z_{ijr} - \frac{1}{n}\sum_{i=1}^{n}\frac{\left\{\sum_{j=1}^{m_i}H_{ij}(T_{ij})\,Z_{ijr}\right\}\int\omega^{N_{i.}(\tau)+1}\exp\{-\omega H_{i.}(\tau)\}f(\omega)\,d\omega}{\int\omega^{N_{i.}(\tau)}\exp\{-\omega H_{i.}(\tau)\}f(\omega)\,d\omega}$$

Coefficient score, using psi?

$$U_r = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\delta_{ij}Z_{ijr} - \frac{1}{n}\sum_{i=1}^{n}\left\{\sum_{j=1}^{m_i}H_{ij}(T_{ij})\,Z_{ijr}\right\}\psi_i(\gamma,\Lambda,\tau)$$

Frailty parameter score

$$U_{p+1} = \frac{1}{n}\sum_{i=1}^{n}\frac{\int\omega^{N_{i.}(\tau)}\exp\{-\omega H_{i.}(\tau)\}f'(\omega)\,d\omega}{\int\omega^{N_{i.}(\tau)}\exp\{-\omega H_{i.}(\tau)\}f(\omega)\,d\omega}$$

Baseline hazard estimation

$$\psi_i(\gamma,\Lambda,t) = \phi_{2i}(\gamma,\Lambda,t)/\phi_{1i}(\gamma,\Lambda,t)$$

$$\phi_{ki}(\gamma,\Lambda_0,t) = \int\omega^{N_{i.}(\tau)+(k-1)}\exp\{-\omega H_{i.}(\tau)\}f(\omega)\,d\omega \quad 1\le k\le 4$$

Likelihood and score functions are function of

$$H_{i\cdot}(\tau) = \sum_{i=1}^{m_i} \Lambda_0(T_{ij}) \exp\left(\beta^T Z_{ij}\right)$$

$$H_{i\cdot}(\tau_{k-1}) = \sum_{i=1}^{m_i} \Lambda_0\{\min(T_{ij}, \tau_{k-1})\} \exp\left(\beta^T Z_{ij}\right)$$

TODO:

LT to solve integrals

Performance remarks

fitfrail arguments, return value

Analytic and empirical asymptotic results.

Correlation with existing packages for gamma, log-normal frailty.

4. Example

Example generating data and fitting model, displaying and interpreting results

5. Case study

TODO: Application to DRS data, maybe HD failure?

6. Discussion

TODO

References

Bender R, Augustin T, Blettner M (2005). "Generating survival times to simulate Cox proportional hazards models." *Statistics in medicine*, **24**(11), 1713–1723.

Brent RP (2013). *Algorithms for minimization without derivatives*. Courier Corporation.

Clauset A, Shalizi CR, Newman ME (2009). "Power-law distributions in empirical data." *SIAM review*, **51**(4), 661–703.

Crowther MJ, Lambert PC (2013). "Simulating biologically plausible complex survival data." *Statistics in medicine*, **32**(23), 4118–4134.

Geyer CJ (2014). *aster: Aster Models*. R package version 0.8-30, URL http://CRAN.R-project.org/package=aster.

Gorfine M, Zucker DM, Hsu L (2006). "Prospective survival analysis with a general semiparametric shared frailty model: A pseudo full likelihood approach." *Biometrika*, **93**(3), 735–741.

Group DRSR, *et al.* (1976). "Preliminary report on effects of photocoagulation therapy." *American journal of ophthalmology*, **81**(4), 383–396.

Hanagal DD (2009). "Modeling heterogeneity for bivariate survival data by power variance function distribution." *Journal of Reliability and Statistical Studies (ISSN: 0974-8024)*, **2**(1), 14–27.

Hougaard P, Hougaard P (2000). *Analysis of multivariate survival data*, volume 564. Springer New York.

Huster WJ, Brookmeyer R, Self SG (1989). "Modelling paired survival data with covariates." *Biometrics*, pp. 145–156.

Ridout MS (2009). "Generating random numbers from a distribution specified by its Laplace transform." *Statistics and Computing*, **19**(4), 439–450.

Smyth G, Hu Y, Dunn P, Phipson B, Chen Y (2015). *statmod: Statistical Modeling*. R package version 1.4.21, URL http://CRAN.R-project.org/package=statmod.

A. Frailty distributions

All frailty distributions have support $\omega \in (0, \infty)$. It is usually necessary to place constraints on the parameters to avoid identifiability problems.

TODO: For each frailty distribution: density, Laplace transform, method to generate random samples

A.1. Gamma

Frailty values from a gamma distribution are denoted by $\omega \sim \Gamma(\frac{1}{\theta}, \frac{1}{\theta})$, where $E[\omega] = 1$ and $Var[\omega] = \theta$. The frailtyr package uses a one-parameter gamma distribution with shape and scale both $\frac{1}{\theta}$ and density given by

$$\Gamma(\omega; \theta) = \frac{\omega^{\frac{1}{\theta}-1} \exp\left(\frac{-\omega}{\theta}\right)}{\theta^{\frac{1}{\theta}} \Gamma(\frac{1}{\theta})} \tag{8}$$

The special case when $\theta = 0$ is taken to be a degenerate distribution where $\omega = 1$, i.e. there is no hidden frailty in the hazard function. The partial derivative, with respect to the single parameter θ, has closed form solution

$$\frac{\partial \Gamma(\omega; \theta)}{\partial \theta} = \frac{\left(\frac{\omega}{\theta}\right)^{\frac{1}{\theta}-1} \exp\left(\frac{-\omega}{\theta}\right) \left\{ \ln\left(\frac{\theta}{\omega}\right) + \psi^{(0)}\left(\frac{1}{\theta}\right) + \omega - 1 \right\}}{\theta^3 \Gamma\left(\frac{1}{\theta}\right)} \tag{9}$$

LT gamma

$$L^{(p)}(s) = (-1)^p \theta^{-\frac{1}{\theta}} \left(\theta^{-1} + s\right)^{-\left(\frac{1}{\theta}+p\right)} \Gamma\left(\theta^{-1} + p\right) / \Gamma\left(\theta^{-1}\right)$$

A.2. Log-normal

Frailty values from a log-normal distribution are denoted $\omega \sim LN(\theta)$, where the log-mean and log-variance of ω are 0 and θ, respectively. This distribution has $E[\omega] = \exp\frac{\theta}{2}$ and $Var[\omega] = \exp 2\theta - \exp\theta$, with density given by

$$LN(\omega;\theta) = \frac{1}{\omega\sqrt{\theta 2\pi}} \exp\left\{\frac{-(\ln\omega)^2}{2\theta}\right\} \tag{10}$$

Similar to the gamma distribution, the special case of $\theta = 0$ indicates $\omega = 1$. The partial derivative with respect to θ is given by

$$\frac{\partial LN(\omega;\theta)}{\partial\theta} = \frac{\ln^2(\omega)\exp\left(\frac{-\ln^2\omega}{2\theta}\right)}{2\sqrt{2\pi}\theta^{5/2}\omega} - \frac{\exp\left(\frac{-\ln^2\omega}{2\theta}\right)}{2\sqrt{2\pi}\theta^{3/2}\omega} \tag{11}$$

A.3. Inverse Gaussian

Frailty values from an inverse Gaussian distribution are denoted by $\omega \sim IG(\theta)$, where $\frac{1}{\theta}$ is the scale and the mean is fixed at 1. The density is given by

$$IG(\omega;\theta) = \left(2\pi\theta\omega^3\right)^{-1/2} \exp\left\{\frac{-(\omega-1)^2}{2\theta\omega}\right\} \tag{12}$$

where $\theta > 0$. The IG has $E[\omega] = 1$ and $Var[\omega] = \theta$. Similar to the gamma and log-normal, ω is taken to be 1 when $\theta = 0$. The derivative wrt. θ is given by

$$\frac{\partial IG(\omega;\theta)}{\partial\theta} = \frac{(\omega-1)^2\exp\left\{-\frac{(\omega-1)^2}{2\theta\omega}\right\}}{2\sqrt{2\pi}\theta^2\omega\sqrt{\theta\omega^3}} - \frac{\omega^3\exp\left\{-\frac{(\omega-1)^2}{2\theta\omega}\right\}}{2\sqrt{2\pi}\left(\theta\omega^3\right)^{3/2}}$$

A.4. Positive stable

The one-parameter positive stable distribution is given by $\omega \sim PS(\alpha)$, where $\beta = 1$, $\mu = 0$, $0 < \alpha < 1$, and $\delta = \alpha$ as defined in Hougaard Hougaard and Hougaard (2000). The density is given by

$$PS(\omega;\alpha) = -\frac{1}{\pi\omega}\sum_{k=1}^{\infty}\frac{\Gamma(k\alpha+1)}{k!}\left(-\omega^{-\alpha}\right)^k \sin(\alpha k\pi) \tag{13}$$

When $\alpha = 1$, a degenerate distribution at 1 is obtained and there is no shared frailty.

A.5. Power variance function

The power variance function (PVF) distribution is denoted $PVF(\alpha,\delta,\theta)$ with density

$$PVF\left(\omega;\alpha,\delta,\theta\right)=\exp\left(-\theta\omega+\frac{\delta^{\alpha}}{\alpha}\right)\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\Gamma\left(k\alpha+1\right)}{k!}\left(-\frac{1}{\omega}\right)^{\alpha k+1}\sin\left(\alpha k\pi\right)$$

where $0<\alpha\leq 1,\theta\geq 0,\delta>0$. To avoid identifiability problems, we let $\delta=\theta=1$ as in Hanagal (2009), and get the one-parameter PVF density

$$PVF\left(\omega;\alpha\right)=\exp\left(-\omega+\alpha^{-1}\right)\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\Gamma\left(k\alpha+1\right)}{k!}\left(-\frac{1}{\omega}\right)^{\alpha k+1}\sin\left(\alpha k\pi\right) \tag{14}$$

When $\alpha=1$, the degenerate distribution with $\omega=1$ is obtained. The PVF has $E[\omega]=1$ and $Var[\omega]=(1-\alpha)$.

Laplace transform

$$L\left(s\right)=\exp\left[-\left\{\left(1+s\right)^{\alpha}-1\right\}/\alpha\right] \tag{15}$$

LT derivs

$$L^{(m)}\left(s\right)=\left(-1\right)^{m}L\left(s\right)\sum_{j=1}^{m}c_{m,j}\left(\alpha\right)\left(1+s\right)^{j\alpha-m} \tag{16}$$

LT deriv wrt alpha

$$\frac{\partial L\left(s\right)}{\partial\alpha}=\exp\left\{\frac{1-\left(s+1\right)^{\alpha}}{\alpha}\right\}\left\{-\frac{1-\left(s+1\right)^{\alpha}}{\alpha^{2}}-\frac{\left(s+1\right)^{\alpha}\log\left(s+1\right)}{\alpha}\right\} \tag{17}$$

LT higher deriv wrt alpha

$$\begin{aligned}c_{m,m}\left(\alpha\right) &= 0\\ c_{m,1}\left(\alpha\right) &= \frac{\Gamma\left(m-\alpha\right)}{\Gamma\left(1-\alpha\right)}\\ c_{m,j}\left(\alpha\right) &= c_{m-1,j-1}\left(\alpha\right)+c_{m-1,j}\left(\alpha\right)\left\{\left(m-1\right)-j\alpha\right\}\end{aligned} \tag{18}$$

$$\begin{aligned}\frac{\partial}{\partial\alpha}\mathcal{L}^{(m)}\left(s\right) &= \frac{\partial}{\partial\alpha}\left[\left(-1\right)^{m}\mathcal{L}\left(s\right)\sum_{j=1}^{m}c_{m,j}\left(\alpha\right)\left(1+s\right)^{j\alpha-m}\right]\\ &= \left(-1\right)^{m}\left\{\frac{\partial}{\partial\alpha}\mathcal{L}\left(s\right)\right\}\sum_{j=1}^{m}c_{m,j}\left(\alpha\right)\left(1+s\right)^{j\alpha-m}\\ &\quad+\left(-1\right)^{m}\mathcal{L}\left(s\right)\sum_{j=1}^{m}\left\{\frac{\partial}{\partial\alpha}c_{m,j}\left(\alpha\right)\left(1+s\right)^{j\alpha-m}\right.\\ &\quad\left.+c_{m,j}\left(\alpha\right)j\left(1+s\right)^{j\alpha-m}\ln\left(1+s\right)\right\}\end{aligned} \tag{19}$$

$$
\begin{aligned}
\frac{\partial}{\partial \alpha} c_{m,m}(\alpha) &= 0 \\
\frac{\partial}{\partial \alpha} c_{m,1}(\alpha) &= \frac{\Gamma(m-\alpha)\left\{\psi^{(0)}(1-\alpha) - \psi^{(0)}(m-\alpha)\right\}}{\Gamma(1-\alpha)} \\
\frac{\partial}{\partial \alpha} c_{m,j}(\alpha) &= \frac{\partial}{\partial \alpha} c_{m-1,j-1}(\alpha) + \frac{\partial}{\partial \alpha} c_{m-1,j}(\alpha)\{(m-1) - j\alpha\} - j c_{m-1,j}(\alpha) \qquad (20)
\end{aligned}
$$

B. Simulation results

TODO

Affiliation:

John V. Monaco
Seidenberg School of Computer Science and Information Systems
Pace University
861 Bedford Road
Pleasantville, NY 10570
E-mail: jmonaco@pace.edu
URL: http://vmonaco.com

Malka Gorfine
TODO: department?
Tel Aviv University
Address line 1
Address line 2
E-mail: TODO
URL: TODO

Li Hsu TODO: department?
Fred Hutchinson Cancer Research Center
Address line 1
Address line 2
E-mail: TODO
URL: TODO

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume VV, Issue II *Submitted:* yyyy-mm-dd
MMMMMM YYYY *Accepted:* yyyy-mm-dd