Moment-based approximation with finite mixed Erlang distributions

Hélène Cossette

with D. Landriault, E. Marceau, K. Moutanabbir

49th Actuarial Research Conference (ARC) 2014
University of California Santa Barbara

Université Laval

July 15th 2014

Acknowledgements

- For this research, the authors would like to gratefully acknowledge the financial support received from The Actuarial Foundation ("TAF"), The Casualty Actuarial Society ("CAS") and The Society of Actuaries ("SOA").
- This project was funded by Individual Grants Competition under the name of "Mixed Erlang Moment-Based Approximation: Applications in Actuarial Science and Risk Management".

0. Outline

- Introduction
- Definitions and Notations
- Mixed Erlang distribution
- Moment-based approximation methods
- Approximation method based on finite mixed Erlang distributions: 2 contexts
- Numerical examples

1. Introduction

- We consider a positive continuous rv S for which we know the first m moments.
- Applications in actuarial science and quantitative risk management:
  - S = aggregate claim amount for a portfolio of insurance risks
  - S = aggregate claim amount for a line of business
  - S = aggregate losses for a portfolio of investment risks (e.g. credit risks)
- Main objective: evaluate cdf of S i.e. F_S
- Impossible or very difficult to find F_S analytically
- Possible to use aggregation methods:
  - Methods based on recursive numerical methods
  - Methods based on MC simulation
- May be very time-consuming to find numerically F_S
- A **moment-based approximation** can be used

2. Definitions and notations

- S : rv with cdf F_S
- j^{th} raw moments : $\mu_j(S) = E\left[S^j\right]$, $j \in \mathbb{N}^+$
- Risk measure VaR
  - $VaR_\kappa(S) = F_S^{-1}(\kappa)$, for $\kappa \in (0,1)$, where $F_S^{-1}(u) = \inf\{x \in \mathbb{R} : F_S(x) \geq u\}$
- Risk measure TVaR
  - $TVaR_\kappa(S) = \frac{1}{1-\kappa}\int_\kappa^1 VaR_u(S)\,\mathrm{d}u$ for $\kappa \in (0,1)$
  - $TVaR_\kappa(S) = \frac{E\left[S \times 1_{\{S > VaR_\kappa(S)\}}\right] + VaR_\kappa(S)(F_S(VaR_\kappa(S)) - \kappa)}{1-\kappa}$
  - If the rv S is continuous, $TVaR_\kappa(S) = \frac{E\left[S \times 1_{\{S > VaR_\kappa(S)\}}\right]}{1-\kappa}$
- Stop-loss premium :
  $\pi_S(b) = E\left[\max(S-b;0)\right] = E\left[S \times 1_{\{S>b\}}\right] - bF_S(b)$

3. Mixed Erlang distribution

- Attractive features of this class of distributions: studied by e.g. Willmot & Woo (2007), Lee & Lin (2010), and Willmot & Lin (2010)
  - Illustrate the versatility of this distribution to model claim amounts
  - Illustrate the faisability to obtain closed-form expressions for various quantities of interest in risk theory.
  - Provide several non trivial examples of distributions which belong to the class of mixed Erlang distributions
  - Provide a detailed procedure to express e.g. mixtures of exponentials, generalized Erlang distributions in terms of mixed Erlang distributions
- Closed under various operations such as convolutions, Esscher transformations (risk aggregation and ruin problems).
- Risk measures VaR, TVaR and stop-loss premium are easily obtained.

3. Mixed Erlang distribution

- Tijms (1994):
  - class of mixed Erlang distributions is dense in the set of all continuous and positive distributions
  - any nonnegative continuous distribution may be approximated by an Erlang mixture to any given accuracy
  - **Theorem.** Let F be the cdf of a positive rv. For any given $h > 0$, define the cdf F_h by

$$F_h(x) = \sum_{j=1}^{\infty} \left(F(jh) - F((j-1)h)\right) H\left(x; j, \frac{1}{h}\right), \qquad x \geq 0,$$

    where $H(x; n, \beta)$ is the Erlang cdf. Then, for any continuity point x of F,

$$\lim_{h \to 0} F_h(x) \to F(x).$$

- **Moment-based approximation** based on class of mixed Erlang dist.

3. Mixed Erlang distribution

- W : mixed Erlang rv with common rate parameter β
- Cdf of W : $F_W(y) = \sum_{k=1}^{l} \zeta_k H(x; k, \beta)$, l finite or infinite
- $H(x; k, \beta) = 1 - e^{-\beta x} \sum_{i=0}^{k-1} \frac{(\beta x)^i}{i!}$: cdf of Erlang rv of order k
- ζ_k : non-negative mass probability associated to the kth Erlang distribution in the mixture
- $\sum_{k=1}^{\infty} \zeta_k = 1$

- Representation of the mixed Erlang distribution as a compound distribution with discrete primary distribution $\{\zeta_k\}_{k=1}^{l}$ and secondary exponential distribution with rate parameter β :

$$W = \sum_{k=1}^{M} C_k,$$

  - $C_k \sim Exp(\beta)$ $(k = 1, 2, ...)$
  - M = discrete r.v. with pmf $f_M(k) = \Pr(M = k) = \zeta_k$, $k \in \mathbb{N}^+$.

- j^{th} raw moment: $\mu_j(W) = \sum_{k=1}^{\infty} \zeta_k \frac{\prod_{i=0}^{j-1}(k+i)}{\beta^j}$
- $CV(W) > \frac{1}{\sqrt{l}}$

3. Mixed Erlang distribution

- No general closed-form expression for *Value-at-risk* but easily obtained with simple numerical optimization method

$$VaR_\kappa(W) = F_W^{-1}(u) = \inf\{x \in \mathbb{R} : F_W(x) \geq u\}, \; \kappa \in (0,1)$$

- Explicit expression for *Tail Value-at-risk*:

$$\begin{aligned} TVaR_\kappa(W) &= \frac{1}{1-\kappa}\int_\kappa^1 VaR_u(W)\,\mathrm{d}u, \; \kappa \in (0,1) \\ &= \frac{1}{1-\kappa}\sum_{k=1}^{\infty} \zeta_k \frac{k}{\beta}\overline{H}(VaR_\kappa(W); k+1, \beta). \end{aligned}$$

- Explicit expression for *Stop-loss premium*:

$$\begin{aligned} \pi_W(b) &= E[\max(W-b; 0)] \\ &= \frac{1}{1-\kappa}\sum_{k=1}^{\infty} \zeta_k \left(\frac{k}{\beta}\overline{H}(b; k+1, \beta) - b\overline{H}(b; k, \beta)\right). \end{aligned}$$

4. Moment-based approximation methods

- **One approach** : approximate the unknown distribution by a mixture of known distributions.
- Several approximation methods motivated by Tijms' theorem were proposed over the years
- Examples of such methods:
  - Whitt (1982) :
    - 3 moments and 2 moments
    - $CV(S) > 1$: mixtures of 2 exponential distributions
      $F_W(x) = p_1\left(1 - e^{-\beta_1 x}\right) + p_2\left(1 - e^{-\beta_2 x}\right)$
    - $CV(S) < 1$: generalized Erlang distribution
      $F_W(x) = H(x; \beta_1, ..., \beta_r) = \sum_{i=1}^{r}\left(\prod_{l=1, l\neq i}^{r} \frac{\beta_l}{\beta_l - \beta_i}\right)\left(1 - e^{-\beta_i x}\right)$

- Continued...
  - Altiok (1985) :
    - 3 moments and 2 moments
    - $CV(S) > 1$: mixture of generalized Erlang distribution and exponential distribution $F_W(x) = pH(x; \beta_1, \beta_2) + (1-p)\left(1 - e^{-\beta_1 x}\right)$
    - $CV(S) < 1$: generalized Erlang distribution $F_W(x) = H(x; \beta_1, ..., \beta_r) = \sum_{i=1}^{r} \left(\prod_{l=1, l \neq i}^{r} \frac{\beta_l}{\beta_l - \beta_i} \right) \left(1 - e^{-\beta_i x}\right)$
  - Johnson & Taaffe (1989) :
    - 3 moments: mixture of two Erlang distributions of common order and different scale factors
    - generalize the approximation of Whitt (1982) and Altiok (1985) (for $CV(S) > 1$)
    - more than 3 moments...

4. Moment-based approximation methods

- Substantial body of literature on **3-moment** based approximations within phase-type class of distributions
- Matching first 3 moments: effective to provide a reasonable approximation (see Osogami and Harchol-Balter (2006)) but does not always suffice.
- Development of more flexible moment-based approximation methods:
  - Johnson and Taaffe (1989)
  - Horvath and Telek (2007)
  - Our proposed method.

5. Approx. method based on finite mixed Erlang distribution

- S : rv with m known moments $\mu_1(S)$, ..., $\mu_m(S)$
- Idea : map (approximate) F_S to a subclass of distributions which belongs to the class of mixed Erlang distributions
- Subclass = class of **finite** mixed Erlang distributions with
  - $F_W(y) = \sum_{k=1}^{l} \zeta_k H(x; k, \beta)$; $l < \infty$
  - $\mu_j(W) = E\left[W^j\right] = \sum_{k=1}^{l} \zeta_k \frac{\prod_{i=0}^{j-1}(k+i)}{\beta^j}$ $(j = 1, 2, ..., m)$

5. Approx. method based on finite mixed Erlang distribution

- Consider a set of first m moments $(\mu_1, ..., \mu_m) = (\mu_1(W),\mu_2(W),...,\mu_m(W))$ and $A_l = \{1, 2, ..., l\}$
- $\mathcal{ME}(\mu_1, ..., \mu_m, A_l)$: set of all **finite** mixtures of Erlang with cdf $F(y) = \sum_{k=1}^{l} \zeta_k H(x; k, \beta)$ and first m moments $(\mu_1,\mu_2,...,\mu_m)$.
- Identification of all solutions to the problem:

$$\mu_j(S) = \sum_{k=1}^{l} \zeta_k \frac{\prod_{i=0}^{j-1}(k+i)}{\beta^j}, j = 1, ..., m.$$

- Constraints: $\beta, \{\zeta_k\}_{k=1}^{l}$ are non-negative and $\sum_{k=1}^{l} \zeta_k = 1$.

5. Approx. method based on finite mixed Erlang distribution

- $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_l)$:
  - (restricted) subset of $\mathcal{ME}(\mu_1, ..., \mu_m, A_l)$ such that **at most** m of the mixing weights $\{\zeta_k\}_{k=1}^{l}$ are non-zero.
  - Propose to use it as our class of approximations.
  - $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_{l_1}) \subseteq \mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_{l_2})$ for $l_1 \leq l_2$.
  - Members are identified by rewriting moment expressions in matrix form

5. Approx. method based on finite mixed Erlang distribution

- Obtain **all** sets of m atoms $\{i_k\}_{k=1}^m$ $(i_1 < i_2 < ... < i_m < l)$ in $A_l = \{1, 2, ..., l\}$
- For a given set of atoms $\{i_k\}_{k=1}^m$, $\mu_j(S) = \sum_{k=1}^{l} \zeta_k \frac{\prod_{i=0}^{j-1}(k+i)}{\beta^j}$ $j = 1, ..., m$ can be written as:

$$\mathbf{A}_{m,m}^T \zeta_m = \mathbf{M}\beta$$

- $\boldsymbol{\zeta}_m^T = (\zeta_{i_1}, \zeta_{i_2}, ..., \zeta_{i_m})$, $\mathbf{M} = diag(\mu_1, \mu_2, ..., \mu_m)$, $\boldsymbol{\beta}^T = (\beta, \beta^2, ..., \beta^m)$
- $\mathbf{A}_{m_1,m_2} = \begin{pmatrix} i_1 & i_2 & \cdots & i_{m_1} \\ i_1(i_1+1) & i_2(i_2+1) & \cdots & i_{m_1}(i_{m_1}+1) \\ \vdots & \vdots & \ddots & \vdots \\ \prod_{i=0}^{m_2-1}(i_1+i) & \prod_{i=0}^{m_2-1}(i_2+i) & \cdots & \prod_{i=0}^{m_2-1}(i_{m_1}+i) \end{pmatrix}$
- m_1 : number of Erlang terms and m_2 : number of moments

5. Approx. method based on finite mixed Erlang distribution

- $\zeta_m = \left[\mathbf{A}_{m,m}^{-1}\mathbf{M}\right]\boldsymbol{\beta}$ under the constraint that $\mathbf{e}^T\zeta_m = 1$, with $\mathbf{e}$ a column vector of 1s.
- $\mathbf{e}^T \left[\mathbf{A}_{m,m}^{-1}\mathbf{M}\right]\boldsymbol{\beta}$: polynomial of degree m in β.
- Look for positive solutions in β to $\mathbf{e}^T \left[\mathbf{A}_{m,m}^{-1}\mathbf{M}\right]\boldsymbol{\beta} = \mathbf{1}$.
- Complete mixed Erlang representations via identification of mixing weights through $\zeta_m = \left[\mathbf{A}_{m,m}^{-1}\mathbf{M}\right]\boldsymbol{\beta}$.
- Repeat procedure for all possible sets of atoms.

5. Approx. method based on finite mixed Erlang distribution

- Criteria of quality among all legitimate candidates in $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_l)$: Kolmogorov-Smirnov (KS) distance
- KS distance for two rv's S and W (with respective cdf F_S and F_W):

$$d_{KS}(S, W) = \sup_{x \geq 0} |F_S(x) - F_W(x)|.$$

- Denote by $F_{W_{m,l}}$ this approximation:

$$d_{KS}(S, W_{m,l}) = \inf_{F_W \in \mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_l)} \sup_{x \geq 0} |F_S(x) - F_W(x)|,$$

where $W_{m,l}$ is a rv with cdf $F_{W_{m,l}}$.

6. Numerical examples

- **Example #1: Weibull rv S**
- $F_S(x) = 1 - \exp\left\{-(x/\beta)^{\tau}\right\}$ for $x, \tau, \beta > 0$.
- Parameters: $\tau = 1.5$ and $\beta = \Gamma(5/3)$
- $CV = 0.6790$.
- Consider class of mixed Erlang distributions $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_{20})$ with $A_{20} = \{1, 2, ..., 20\}$

- Cardinalities of $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_{20})$:

m	Cardinality $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_{20})$
3	298
4	577
5	1010

- Resulting mixed Erlang approximations for $m = 3, 4$:

$$\begin{aligned} F_{W_{3,20}}(x) &= 0.0564H(x; 1, 3.0114) + 0.4097H(x; 2, 3.0114) \\ &+0.5339H(x; 4, 3.0114), \end{aligned}$$

$$\begin{aligned} F_{W_{4,20}}(x) &= 0.0355H(x; 1, 3.9083) + 0.2777H(x; 2, 3.9083) \\ &+0.4966H(x; 4, 3.9083) + 0.1901H(x; 7, 3.9083), \end{aligned}$$

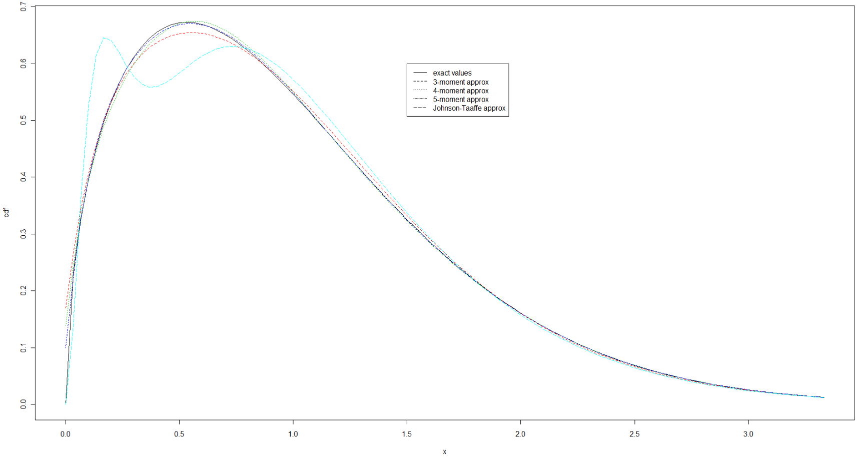
- Kolmogorov-Smirnov distances:

m	$d_{KS}\left(S, W_{m,20}\right)$
3	0.0042
4	0.0020
5	0.0007

- Quality of the approximation improves from 3 to 5 moments.

6. Numerical examples

- Comparison of pdfs of $W_{3,20}$, $W_{4,20}$, $W_{5,20}$ and S. The 3-moment approximation of Johnson and Taffee (1989) is also provided.



- Examine the tail fit: VaR and TVaR for the exact and approximated distributions

κ	$VaR_\kappa(W_{3,20})$	$VaR_\kappa(W_{4,20})$	$VaR_\kappa(W_{5,20})$	$VaR_\kappa(S)$
0.9	1.9224	1.9334	1.9324	1.9316
0.99	3.0670	3.0647	3.0651	3.0662
0.999	4.0823	4.0116	4.0146	4.0178
0.9999	5.0375	4.8706	4.8734	4.8674

κ	$TVaR_\kappa(W_{3,20})$	$TVaR_\kappa(W_{4,20})$	$TVaR_\kappa(W_{5,20})$	$TVaR_\kappa(S)$
0.9	2.4287	2.4370	2.4361	2.4354
0.99	3.5112	3.4807	3.4823	3.4844
0.999	4.4988	4.3871	4.3901	4.3897
0.9999	5.4378	5.2227	5.2244	5.2083

- **Example #2: Lognormal rv** S
- $S = \exp(\nu + \sigma Z)$ where Z is a standard normal rv
- Consider Example 5.4 of Dufresne (2007) where $\nu = 0$ and $\sigma^2 = 0.25$.
- $CV = 0.5329$.
- Lognormal has a heavier tail than mixed Erlang: no guarantee that our mixed Erlang approximation would perform well, especially for tail risk measures.

6. Numerical examples

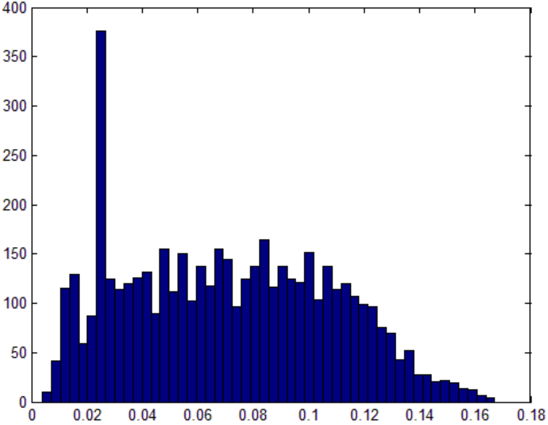
- Consider class of mixed Erlang distributions $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_{50})$.
- Kolmogorov-Smirnov distances:

m	$d_{KS}(S, W_{m,50})$
3	0.0040
4	0.0018
5	0.0025

- KS distance increases from the 4-moment to the 5-moment approximation.
- **Remark:** $F_{W_{5,50}}$ uses Erlang-50 cdf, where 50 is the upper boundary point of A_{50} : believe that a mixed Erlang approximation with a KS distance lower than 0.0018 could be found by increasing the value l in set A_l.

6. Numerical examples

- Histograms of the KS distance for all the mixed Erlang distributions in $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_{50})$, $m = 3, 4, 5$
- KS distances (x-axis) vs counts (y-axis)



6. Numerical examples

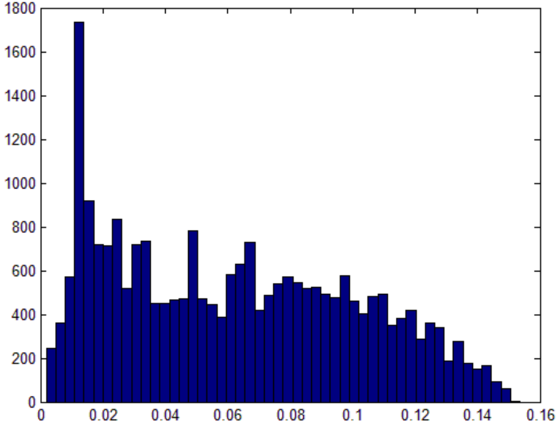


Figure: 4-moment approximations

6. Numerical examples

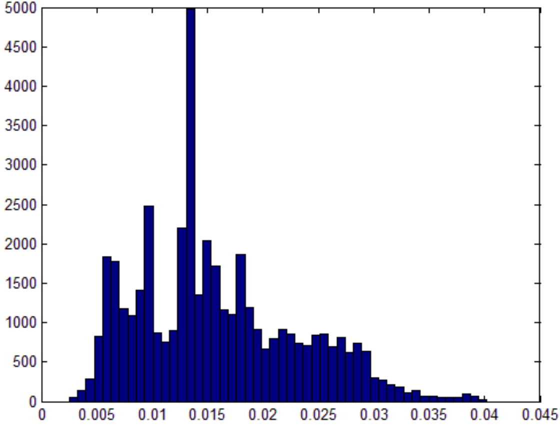
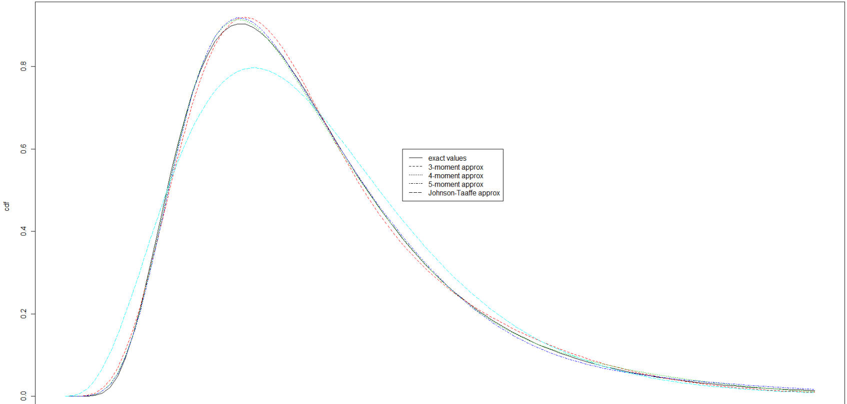


Figure: 5-moment approximations

- Overall quality of the approximations (judged by values and dispersion of KS distances) increases with number of moments matched.
- Comparison of the pdfs of $W_{3,50}$, $W_{4,50}$, $W_{5,50}$, and S. The 3-moment approximation of Johnson and Taaffe (1989) is also plotted.



6. Numerical examples

- All three mixed Erlang approximations provide an overall good fit to the exact distribution.
- To further examine the tail fit, specific values of VaR and TVaR for the exact and approximated distributions are provided below:

κ	$VaR_\kappa(W_{3,50})$	$VaR_\kappa(W_{4,50})$	$VaR_\kappa(W_{5,50})$	$VaR_\kappa(S)$
0.9	1.9129	1.9056	1.8891	1.8980
0.99	3.1223	3.0991	3.2220	3.2001
0.999	4.9237	5.0623	4.3746	4.6885
0.9999	6.0352	6.2642	6.8895	6.4206

κ	$TVaR_\kappa(W_{3,50})$	$TVaR_\kappa(W_{4,50})$	$TVaR_\kappa(W_{5,50})$	$TVaR_\kappa(S)$
0.9	2.4540	2.4550	2.4697	2.4616
0.99	3.9008	3.8625	3.7700	3.8413
0.999	5.4245	5.5431	5.5204	5.4341
0.9999	6.4189	6.6093	7.4023	7.2879

6. Numerical examples

- *VaR* and *TVaR* values of our mixed Erlang approximations compare reasonably well to their lognormal counterparts.
- Improvement is not monotone with the number of moments matched: well known that increasing the number of moments does not necessarily lead to a higher quality approximation in moment-matching techniques.

- **Example #3: Real data**
- Normalized damage amounts from 30 most damaging hurricanes in United States from 1925 to 1995 (provided by Pielke and Landsea (1998) and analyzed by Brazauskas et al. (2009)).
- Purpose of this example: not to carry an exhaustive statistical analysis of this dataset, but provide a simple fit with a finite mixed Erlang distribution
- First 4 empirical moments:

j	1	2	3	4
μ_j	11.7499	317.5154	15604.47	986686.4

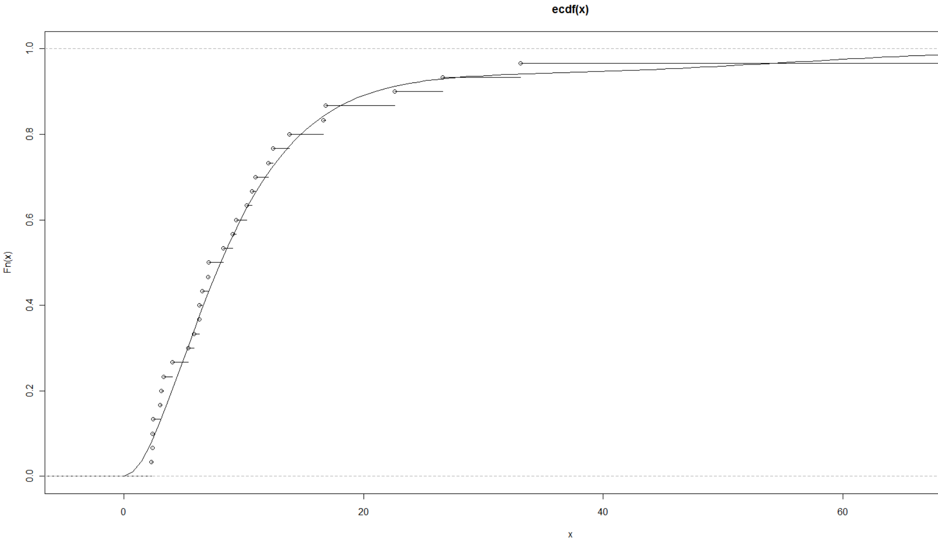
- $CV = 1.1401$.

- Perform the approximation with $\mathcal{ME}^{res}(\mu_1, ..., \mu_m, A_{30})$ for $m = 3$ and 4.
- Kolmogorov-Smirnov distances (with empirical distribution) :

m	$d_{KS}(S, W_{m,30})$
3	0.0773
4	0.0769

- Critical value of the KS hypothesis test at a significance level of 1%: $1.63/\sqrt{30} = 0.2976$
- Do not reject both distributions as a plausible model for the dataset.

6. Numerical examples



7. Moment-based approx. with known rate parameter

- Slightly different context.
- Distribution function F_S is known to be of mixed Erlang form with known $\beta > 0$ and $(\mu_1, \mu_2, ..., \mu_m)$.
- Distribution itself unknown or difficult to evaluate.
- Restrict to sets of **finite** mixture of Erlang distributions.
- Bounds on risk measures can be established.
- Connection with extremal points of a discrete moment-matching problem.

7. Moment-based approx. with known rate parameter

- $F_S \in \mathcal{ME}(\mu_1, ..., \mu_m, \beta)$: set of all mixed Erlang dist. for $l = \infty$, rate parameter β and first m moments $(\mu_1, ..., \mu_m)$.
- $\mathcal{ME}(\mu_1, ..., \mu_m, A_l, \beta)$: subset of $\mathcal{ME}(\mu_1, ..., \mu_m, \beta)$ for a given $l \in \mathbb{N}^+$.
- $\mathcal{ME}^{ext}(\mu_1, ..., \mu_m, A_l, \beta)$: subset of $\mathcal{ME}(\mu_1, ..., \mu_m, A_l, \beta)$ such that **at most** $(m+1)$ of mixing weights $\{\zeta_k\}_{k=1}^{l}$ are non-zero.
- Consider two approaches to derive bounds on $E[\phi(S)]$ for ϕ a given function (such that expectation exists):
  - Based on discrete s-convex extremal distributions
  - Based on moment bounds on discrete expected stop-loss transforms

8. Discrete s-convex extremal distributions

- $D(\alpha_1, ..., \alpha_m, A_l)$: all discrete dist. with support A_l with first m moments $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_m)$.
- $D^{ext}(\alpha_1, ..., \alpha_m, A_l)$: all discrete dist. with support A_l with at most $(m+1)$ non-zero mass points with first m moments are $\boldsymbol{\alpha}$.
- For a given $\beta > 0$: one-to-one correspondence between discrete classes and mixed-Erlang classes
- Each dist. in $D(\alpha_1, ..., \alpha_m, A_l)$ (and $D^{ext}(\alpha_1, ..., \alpha_m, A_l)$) corresponds to a mixed Erlang dist. in $\mathcal{ME}(\mu_1, ..., \mu_m, A_l, \beta)$ (and $\mathcal{ME}^{ext}(\mu_1, ..., \mu_m, A_l, \beta)$) (see De Vylder 1996)
- Allows to use theory on sets of discrete distributions e.g. in Prékopa (1990), Denuit, Lefèvre and Mesfioui (1999), Courtois et. al (2006).

- **Definition s-convex:** Let C be a subinterval of $\mathbb{R}$ or a subset of $\mathbb{N}$ and ϕ a function on C. For two rv's X and Y defined on C, X is said to be smaller than Y in the s-convex sense, namely $X \preceq_{s-cx}^{C} Y$, if $E[\phi(X)] \leq E[\phi(Y)]$ for all s-convex functions ϕ.
- Examples of s-convex functions: $\phi(x) = x^{s+j}$ and $\phi(x) = \exp(cx)$ for $c \geq 0$.
- $K_{s,\min}$ and $K_{s,\max}$: s-extremum rv's on $D(\alpha_1, ..., \alpha_m, A_l)$

$$E[\phi(K_{s,\min})] \leq E[\phi(K)] \leq E[\phi(K_{s,\max})]$$

for any s-convex function ϕ and any $K \in D(\alpha_1, ..., \alpha_m, A_l)$.

- General distribution forms of $K_{s,\min}$ and $K_{s,\max}$ are given in Prékopa (1990) and Courtois et al. (2006)
- $W_K = \sum_{j=1}^{K} C_j$ be a mixed Erlang rv.
- Denuit, Lefèvre and Utev (1999) state that the s-convex order is stable under compounding.
- **Lemma:** If $K \preceq_{s-cx}^{A_l} K'$', then $W_K \preceq_{s-cx}^{\mathbb{R}^+} W_{K'}$.
- Can apply this Lemma to $W_{K_{s-\min}}$ and $W_{K_{s-\max}}$:

$$W_{K_{s-\min}} \preceq_{s-cx}^{\mathbb{R}^+} W_K \preceq_{s-cx}^{\mathbb{R}^+} W_{K_{s-\max}}$$

- Allows to find general distribution forms of $F_{W_{K_{s-\min}}}$ and $F_{W_{K_{s-\max}}}$
- For s-convex functions $\phi(x) = x^{s+j}$ and $\phi(x) = \exp(cx)$,can obtain bounds:

$$\begin{aligned} E\left[W_{K_{s-\min}}^{s+j}\right] &\leq E\left[W_K\right] \leq E\left[W_{K_{s-\max}}^{s+j}\right] \\ E\left[\exp(cW_{K_{s-\min}})\right] &\leq E\left[\exp(cW_K)\right] \leq E\left[\exp(cW_{K_{s-\max}})\right] \end{aligned}$$

9. Moment bounds on discrete expected stop-loss transforms

- Extrema with respect to s-convex order allows to derive bounds on $E\left[\phi(S)\right]$ for all s-convex functions ϕ.
- Approach not appropriate to derive bounds for TVaR and stop-loss premium when $m \geq 2$.
- Use an approach (based on increasing convex order) inspired from Courtois and Denuit (2009) and Hürlimann (2002).
- **Main idea:**
  - consider $D(\alpha_1, ..., \alpha_m, A_l)$ for $m \in \{2, 3, ...\}$
  - find lower and upper bounds for $E\left[(K-k)_+\right]$ on $D(\alpha_1, ..., \alpha_m, A_l)$ for all $k \in A_l$
  - from lower (upper) bound, derive corresponding rv K_{m-low} (K_{m-up})
  - $E\left[(K_{m-low}-k)_+\right] \leq E\left[(K-k)_+\right] \leq E\left[(K_{m-up}-k)_+\right]$ on $D(\alpha_1, ..., \alpha_m, A_l)$ for all $k \in A_l$
  - implies under the increasing convex order: $K_{m-low} \preceq_{icx} K \preceq_{icx} K_{m-up}$

9. Moment bounds on discrete expected stop-loss transforms

- Increasing convex order is stable under compounding:

$$W_{K_{m-low}} \preceq_{icx} W_K \preceq_{icx} W_{K_{m-up}}$$

- From Denuit et al. (2005):

$$TVaR(W_{K_{m-low}}) \leq TVaR(W_K) \leq TVaR(W_{K_{m-up}})$$

10. Example - Portfolio of dependent risks

- Portfolio of n dependent risks (common mixture model of Cossette and al. (2002))
- $S = X_1 + ... + X_n$: aggregate claim amount with $X_i = B_i I_i$.
- Conditional on a common mixture rv Θ with pmf p_Θ, $\{I_i\}_{i=1}^n$ are assumed to form a sequence of independent Bernoulli rv's with

$$\Pr\left(I_i = 1 \left| \Theta = \theta \right.\right) = 1 - r_i{}^\theta \text{ for } r_i \in (0, 1).$$

- B_i $(i = 1, ..., n)$ are assumed to form a sequence of iid rv's, independent of $\{I_i\}_{i=1}^{20}$ and Θ.
- B_i $(i = 1, ..., n)$: exponentially distributed with mean 1
- Distribution of S : two-point mixture of a degenerate rv at 0 and a mixed Erlang with $l = n$ and $\beta = 1$.

- Parameters:
  - $n = 20$ risks
  - Θ has a logarithmic distribution with pmf $p_{\Theta}(j) = (0.5)^{j} / (j \ln 2)$ for $j = 1, 2, ...$
  - constants r_i are set such that the (unconditional) mean of I_i is $q_i = 1 - E\left[(r_i)^{\Theta}\right]$ with $q_1 = ... = q_{10} = 0.1$ and $q_{11} = ... = q_{20} = 0.02$. It
- Perform moment-based approximation on rv $Y = (S \,|S > 0)$ rather than S
- j-th moment of Y : $\mu'_j \equiv E\left[Y^j\right] = \frac{E\left[S^j\right]}{1 - F_S(0)}$
- $CV(Y) = 0.9603$.
- Methods of Whitt (1982) and Altiok (1985) not applicable here: constraints on CV and third moment ($\mu_3 \mu_1 \geq 1.5 {\mu_2}^2$) not satisfied.
- Method of Johnson and Taaffe (1989): $r = 2$, $\beta_1 = 0.7627$, $\beta_2 = 2.8939$ and $p = 0.5742$.

- **First approach: discrete s-convex extremal distributions**
- Find cdfs $F_{W_{K_{s-\min}}}$ and $F_{W_{K_{s-\max}}}$ for $m = 4, 5$ $(s = m + 1)$
- Consider two distributional characteristics of S :
  - higher-order moments $E\left[S^j\right]$ for $j = 4, 5, 6$
  - exponential premium principle $\varphi_\eta\left(S\right) = \frac{1}{\eta}\ln E\left[e^{\eta S}\right]$ for $\eta > 0$.
- Distributions $F_{W_{K_{m+1-\min}}}$ and $F_{W_{K_{m+1-\max}}}$ provide bounds to these risk measures associated to the rv S

10. Example - Portfolio of dependent risks

- Bounds on $E\left[S^j\right]$ and $\varphi_\eta(S) = \frac{1}{\eta}\ln E\left[e^{\eta S}\right]$:

j	$E\left[W^j_{K_{5-\min}}\right]$	$E\left[W^j_{K_{6-\min}}\right]$	$E\left[S^j\right]$	$E\left[W^j_{K_{6-\max}}\right]$	$E\left[W^j_{K_{5-\max}}\right]$
4	138.7579	138.7579	138.7579	138.7579	138.7579
5	1125.9592	1129.1880	1129.1880	1129.1880	1149.9348
6	10748.5738	10873.8020	10881.2732	10922.7337	11993.6176

θ	$\varphi_\eta\left(W_{K_{5-\min}}\right)$	$\varphi_\eta\left(W_{K_{6-\min}}\right)$	$\varphi_\eta(S)$	$\varphi_\eta\left(W_{K_{6-\max}}\right)$	$\varphi_\eta\left(W_{K_{5-\max}}\right)$
0.2	1.5545	1.5546	1.5546	1.5548	1.5564
0.1	1.3536	1.3536	1.3536	1.3536	1.3536
0.01	1.2137	1.2137	1.2137	1.2137	1.2137

- Bounds get sharper as the number of moments involved increases.

10. Example - Portfolio of dependent risks

- **Second approach: moment bounds with discrete expected stop-loss transforms**
- Values of TVaR for $W_{K_{m-low}}$ and $W_{K_{m-up}}$ $(m = 4, 5)$:

	Exact	J&T	$TVaR_\kappa(...)$ for $m = 3$	
κ	$TVaR_\kappa(S)$	$TVaR_\kappa(W)$	W_{3-low}	W_{3-up}
0.9	5.0696	5.1389	4.798911	5.333275
0.95	6.2214	6.2563	5.771565	6.615174
0.99	8.8460	8.7491	7.911982	9.675684
0.995	9.9589	9.7892	8.799191	11.116631
0.999	12.5066	12.156	10.805712	15.181871

10. Example - Portfolio of dependent risks

	Exact	$TVaR_\kappa(...)$ for $m=4$		$TVaR_\kappa(...)$ for $m=5$	
κ	$TVaR_\kappa(S)$	$W_{K_{4-low}}$	$W_{K_{4-up}}$	$W_{K_{5-low}}$	$W_{K_{5-up}}$
0.9	5.0696	4.9222	5.2062	4.9800	5.1490
0.95	6.2214	5.9708	6.4548	6.0594	6.3642
0.99	8.8460	8.2899	9.3301	8.4655	9.1767
0.995	9.9589	9.2500	10.5629	9.4679	10.3775
0.999	12.5066	11.4122	13.4854	11.7323	13.1382

- Inequality verified:

$$TVaR(W_{K_{m-low}}) \leq TVaR(W_K) \leq TVaR(W_{K_{m-up}})$$

- Interval estimate of $TVaR_\kappa(S)$ shrinks as number of moments matched increases.