

Optimal PDF (Probability Density Function) Models for Stochastic Model Outcomes: Parametric Model Fitting on Tail Distributions

Y. Chueh⁽¹⁾

**Department of Mathematics, Central
Washington University, Ellensburg, WA**

chueh@cwu.edu

W.D.Curtis

**Department of Mathematics, Central
Washington University, Ellensburg, WA**

(1) corresponding author

Abstract

This paper documents a *Mathematica*-related development activity in the category of applied mathematics for the field of Actuarial Science. This is an on-going research project sponsored by a Central Washington University Seed Grant subsequent to the papers [ALC, YC] on efficient stochastic modeling techniques. Efficient stochastic modeling for the insurance business can be achieved by effective sampling of stochastic economic (risk) scenarios. The run-time can be significantly reduced and the model outcome distribution can be preserved on the critical tails. However, there is a lack of a well-known software application to help modelers analyze the model outcome distribution. This research will apply *Mathematica* as the main programming and computational resource to develop a user-friendly *Mathematica* application AMOOF (Actuarial Model Outcome Optimal Fit) that finds the best parametric probability density model for given stochastic model data. The AMOOF application will carry out a series of parametric probability model fitting processes and then create mathematical reports for actuaries. Unlike other commercial statistical data fitting packages, a comprehensive inventory of special and mixed distributions will be used to enhance fitting accuracy on the tails. The project will utilize *Mathematica*'s capabilities in calculus, graphing, and optimization and, when complete, will be valuable for actuarial modeling that is critical for a company's capital determination, reserve setting, risk analysis, and product pricing.

1. Introduction

We are developing a *Mathematica* application called the Actuarial Model Outcome Optimal Fit (AMOOF). This software is designed for actuarial model outcome analysis. Actuarial models apply stochastic risk scenarios to project profitability and financial strength for insurance business. The resulting stochastic modeling outcome distributions are critical to an insurance company's decision

find the optimal parametric probability models for the model outcome data. The parametric probability density models allow further analysis, given changes in business or contract assumptions. There will be numerical and graphical reports on the findings to assist model applications. The currently available statistical software can fit data to a variety of standard probability distributions, but not those inverse, transformed, or mixed distributions commonly used/required in the actuarial field. There is a lack of a software package to analyze actuarial model outcomes on the tails, even though the tail distributions are what really matter for the insurance business. We are developing an efficient and smart computer program (AMOOF) to explore all (or most) of the probability models used in actuarial science and determine the optimal one that fits the data well, especially at the tails. This is a challenging task not only because the optimization process is complex, but also because the long tails and extreme data values can cause problems in computation. The computer program will provide the user an actuarial report on how well the resulting model fits the input data and statistics for risk management and capital requirements. The program, once complete, can be used in future actuarial research and by practicing actuaries in insurance companies. Actuaries can use the program to conduct the analysis on a variety of insurance model outcomes that might have been too time prohibitive or onerous to do in the past.

2. Overview of the program

- A Mathematica program which finds the parametric distribution (PDF) that best fits the user's stochastic actuarial model outcomes
 - Can import data from the user's text files and export the results as graphic files to a report or presentation
 - Tests all the continuous distribution families for more than 22 different distribution types
 - Display results as presentation-quality graphs and a full statistical report
 - Automatic goodness-of-fit testing and CTE90 (CTE% conditional tail expectations/percentiles) calculation and comparisons
 - On-line Help and tutorial
-

3. Features of AMOOF

AMOOF will accept data in a variety of formats, including:

1. data points with equal probabilities
2. data points with unequal probabilities

Analysis of large data sets will be supported. Input data for AMOOF can be read in from a text file, entered by hand, or cut and pasted data from another Windows application.

Because the majority of insurance model outcomes are of continuous type and discrete outcomes such as claim frequency can be handled by a different approach, this project focuses on continuous probability distributions only.

AMOOF tests more than twenty-two continuous distributions to determine which distribution best fits the data. AMOOF may skip some distributions that are obviously not good fits by checking the skewness.

AMOOF determines the optimal parameters for selected distributions, performing three standard tests to determine the goodness-of-fit: chi-square, Anderson-Darling, and Kolmogorov-Smirnov. When AMOOF completes the process, the optimal distributions are listed in the order of goodness-of-fit.

The PDF Distributions available in AMOOF include the following continuous distributions selected from a well-known actuarial textbook [LM]:

1. Transformed Beta
2. Generalized Pareto
3. Burr
4. Inverse Burr
5. Pareto
6. Inverse Pareto
7. Loglogistic
8. Paralogistic
9. Inverse Paralogistic
10. Transformed Gamma
11. Inverse Transformed Gamma
12. Gamma
13. Inverse Gamma
14. Weibull
15. Inverse Weibull

- 17. Inverse Exponential
 - 18. Lognormal
 - 19. Inverse Gaussian
 - 20. Single-parameter Pareto
 - 21. Generalized beta
 - 22. Beta
 - 23. Mixed (1-22)
-

4. How AMOOF Works

AMOOF's goal is to find the distribution that best fits the input data. AMOOF does not produce an absolute answer; it identifies a density function that most likely produced the input data from a stochastic model.

For a given distribution candidate for the input data set, AMOOF looks for the parameters of the density function that optimize the value of the likelihood function of the data set, a measurement of the probability that the input data was produced by the given distribution. AMOOF allows results to be evaluated both quantitatively and qualitatively, by examining both the comparison graphs and statistics. Before maximizing the likelihood function of the input data, initial guesses of parameter values are produced using the method of moments that equals the parametric moments and sample moments.

The goodness-of-fit measures how well a parametric PDF fits the data set and is calculated for each PDF. All fitted PDF's are compared and the one with the lowest goodness-of-fit value is considered the best fit. AMOOF provides the information needed to decide which fit is the best, and whether that fit is good enough to use. All results, including graphs, statistics, and distribution functions, can easily be transferred to other programs for further analysis and presentation.

The AMOOF Report has a number of features to display results and perform advanced analyses as follows:

1. Graphics: High resolution graphics are used to present results. Comparison, difference, PP, and QQ graphs all lead to a powerful visual presentation of results.
2. Detailed Statistics: AMOOF can generate a full statistical report. Detailed statistics (mean, variance,

calculated and presented.

3. Confidence Levels: AMOOF generates Confidence Levels and Critical Values for each distribution (PDF) tested. These results are available in the statistics report. These statistics can be used to determine how good a fit AMOOF has found. If a particular distribution is rejected for all available confidence levels, this is indicated in the report.

4. The user can also have AMOOF compare the data set to a specific distribution by specifying the parameters.

5. Programming

The main challenge of this project is to write Mathematica code to create an effective user interface, import data sets, choose appropriate probability models, optimize the likelihood functions, conduct statistical tests, perform calculus operations, and create numerical and graphical reports for statistical or actuarial use. The authors wrote the following code example to illustrate the basic data fitting process. Two probability density functions/models were used to fit a given data set. Both Gamma and mixed Gamma density functions were fitted to a data set of one hundred values. To fit the best model, the likelihood functions were defined and maximized followed by the Kolmogorov-Smirnov tests to check the goodness of fit. The CTE code (not in the code example) involves computing improper integrals that can be done with the best accuracy only in Mathematica. Spreadsheet programs and numerical recipes can calculate CTE with less accuracy for some distribution models. Many of the tasks outlined in this paper can be carried out using a variety of statistical software and VBA programming in the Microsoft Excel. As part of the project the authors would like to compare Mathematica's Optimizer with Excel's add-ins Solver tool that has been extensively used in the financial industry.

6. Code Example

Below is a sample data set of 100 values.

10 78 20 762 23 058 24 365 27 216 28 505 29 560 31 7 32 138 32 166 33 610 37 077

38.599, 38.818, 38.964, 42.403, 42.959, 44.655, 46.935, 47.015, 47.078, 47.156, 47.787, 48.02, 49.448, 49.597, 49.612, 51.256, 51.632, 53.547, 54.179, 55.874, 55.91, 55.984, 56.526, 59.556, 61.071, 61.941, 61.951, 65.231, 65.354, 66.349, 72.411, 75.999, 76.355, 79.237, 79.891, 84.059, 101.718, 104.379, 158.74, 160.442, 162.186, 176.292, 190.798, 221.739, 230.269, 230.711, 231.018, 231.021, 235.92, 241.679, 243.443, 243.989, 244.53, 244.694, 257.412, 257.475, 261.791, 271.918, 277.635, 279.834, 281.53, 282.391, 285.835, 290.84, 296.917, 301.806, 309.584, 320.709, 323.789, 328.277, 332.701, 345.071, 346.511, 349.923, 358.246, 359.094, 369.45, 371.69, 375.842, 377.236, 386.892, 391.95, 399.141, 407.217, 409.285, 458.163, 471.869, 486.527

We will fit a Gamma distribution to this data, optimizing the distribution parameters α and θ using the method of maximum likelihood. Then we will fit a mixed density consisting of a convex combination of two Gamma densities. This combination is optimized using the method of maximum likelihood. We then apply the Kolmogorov-Smirnov test to see which distribution best fits the data.

The Gamma density, cumulative distribution, and data are as shown below.

```

In[5]:= gammaPDF[α_, θ_, x_] := 
$$\frac{\left(\frac{x}{\theta}\right)^\alpha e^{-\left(\frac{x}{\theta}\right)}}{x \text{Gamma}[\alpha]}$$


gammaCDF[α_, θ_, u_] := Integrate[
$$\frac{\left(\frac{x}{\theta}\right)^\alpha e^{-\left(\frac{x}{\theta}\right)}}{x \text{Gamma}[\alpha]}$$
, {x, 0, u}];

data = {19.78, 20.762, 23.058, 24.365, 27.216, 28.595, 29.569, 31.7,
32.438, 32.466, 33.649, 37.072, 38.599, 38.818, 38.964,
42.403, 42.959, 44.655, 46.935, 47.015, 47.078, 47.156,
47.787, 48.02, 49.448, 49.597, 49.612, 51.256, 51.632, 53.547,
54.179, 55.874, 55.91, 55.984, 56.526, 59.556, 61.071, 61.941,
61.951, 65.231, 65.354, 66.349, 72.411, 75.999, 76.355,
79.237, 79.891, 84.059, 101.718, 104.379, 158.74, 160.442,
162.186, 176.292, 190.798, 221.739, 230.269, 230.711, 231.018,
231.021, 235.92, 241.679, 243.443, 243.989, 244.53, 244.694,
257.412, 257.475, 261.791, 271.918, 277.635, 279.834, 281.53,
282.391, 285.835, 290.84, 296.917, 301.806, 309.584, 320.709,
323.789, 328.277, 332.701, 345.071, 346.511, 349.923, 358.246,
359.094, 369.45, 371.69, 375.842, 377.236, 386.892, 391.95,
399.141, 407.217, 409.285, 458.163, 471.869, 486.527};

dLen = Length[data];
dMax = Max[data];
dUpper = (Quotient[dMax, 10] + 1) 10;

```

Given a data array $\{x_1, x_2, \dots, x_N\}$ and a probability density function $f(x, \theta)$ depending on a parameter (or vector of parameters) θ , the likelihood function is $L(\theta) = \prod_{k=1}^N f(x_k, \theta)$. The log-likelihood function is $\ln(L(\theta)) = \sum_{k=1}^N \ln(f(x_k))$. The method of maximum likelihood chooses the parameters θ to maximize the log-likelihood function.

```

In[11]:=
logLikelihoodGamma[α_, θ_] :=
Plus@@ (Log[gammaPDF[α, θ, #]] & /@ data)

```

Next we form a mixed distribution by taking a convex combination of two Gamma distributions.

```

In[12]:= mixedGammaPDF[α1_, θ1_, α2_, θ2_, p_, x_] :=
  p gammaPDF[α1, θ1, x] + (1 - p) gammaPDF[α2, θ2, x]

mixedGammaCDF[α1_, θ1_, α2_, θ2_, p_, x_] :=
  p gammaCDF[α1, θ1, x] + (1 - p) gammaCDF[α2, θ2, x]

logLikelihoodMixedGamma[α1_, θ1_, α2_, θ2_, p_] :=
  Plus @@ (Log[mixedGammaPDF[α1, θ1, α2, θ2, p, #]] & /@ data)

```

Find optimal Maximum Likelihood parameter values for the Gamma distribution.

```

In[18]:= solGamma =
  NMaximize[{logLikelihoodGamma[α, θ], α > 0, θ > 0}, {α, θ},
  AccuracyGoal -> 5, PrecisionGoal -> 5, MaxIterations -> 500]

```

```

Out[18]= {-614.307, {α → 1.38601, θ → 127.287}}

```

Find optimal Maximum Likelihood parameter values for the mixed Gamma distribution.

```

In[19]:= solMixedGamma =
  NMaximize[{logLikelihoodMixedGamma[α1, θ1, α2, θ2, p], α1 > 0,
  α2 > 0, θ1 > 0, θ2 > 0, p > 0, p < 1}, {α1, θ1, α2, θ2, p},
  AccuracyGoal -> 5, PrecisionGoal -> 5, MaxIterations -> 500]

```

```

Out[19]= {-574.517, {p → 0.499756, α1 → 7.32034,
  α2 → 13.7522, θ1 → 7.02438, θ2 → 21.8901}}

```

Using the Kolmogorov-Smirnov test, we evaluate how well the data fit the distribution, first for the single Gamma distribution and then for the mixed distribution.

```

In[100]:=
  cdfListGamma = (gammaCDF[α, θ, #] & /@ data) /. solGamma[[2]];

  y = Sqrt[dLen] Max[Abs[cdfListGamma - Range[1./dLen, 1, 1./dLen]]];

  pValGamma = NestWhile[(2 (-1) ^ (j + 1) Exp[-2 j ^ 2 y ^ 2] + #) &,
  j = 1; 0, (Abs[(-1) ^ j ++ (#1 - #2)] > 10 ^ (-8)) &, 2]

```

```

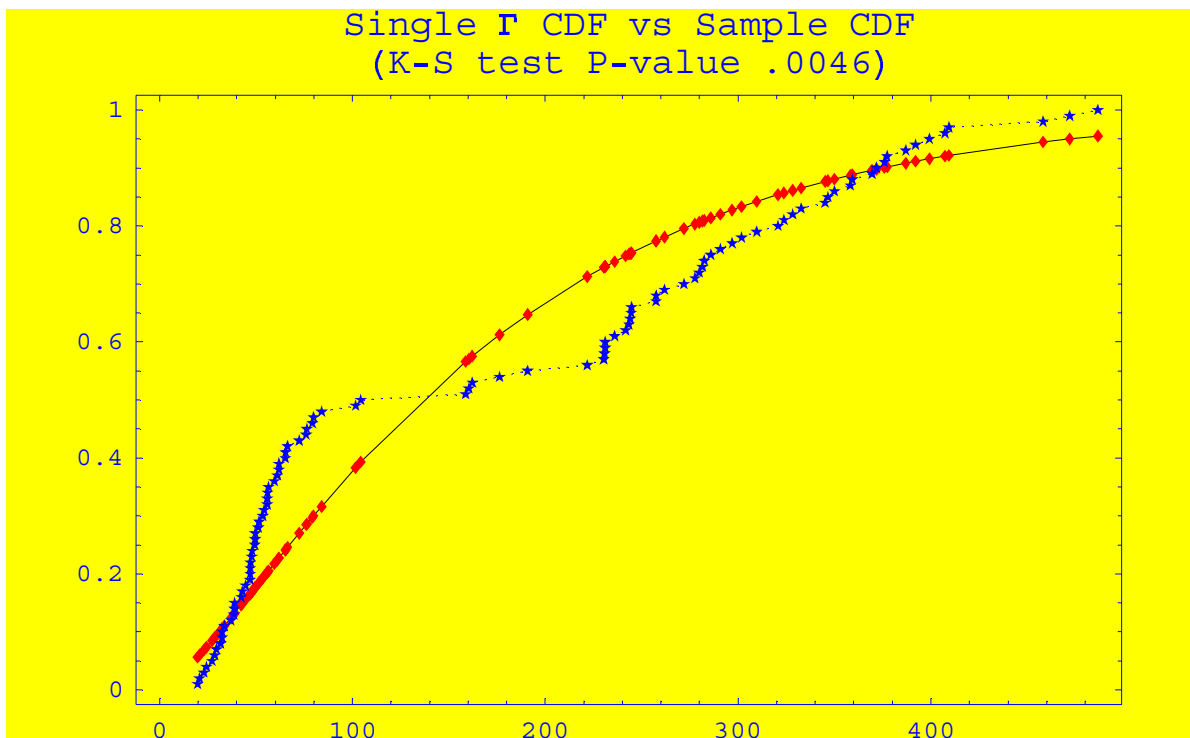
Out[102]=
  0.00458965

```

The very small P-value obtained is indicative of a poor fit, as is seen in the graph below.

```
In[103] :=
```

```
MultipleListPlot[Transpose[{data, cdflistGamma}],
  Transpose[{data, Range[1./dLen, 1, 1./dLen]}],
  PlotJoined → True, SymbolStyle → {Red, Blue},
  PlotLabel -> StyleForm["Single  $\Gamma$  CDF vs Sample CDF\n (K-S
    test P-value .0046)", FontSize → 14, FontWeight → Bold],
  DefaultColor → Blue, Background → Yellow, Frame → True];
```



```
In[90] := cdflistMixedGamma =
  (mixedGammaCDF[ $\alpha$ 1,  $\theta$ 1,  $\alpha$ 2,  $\theta$ 2, p, #] & /@ data) /.
  solMixedGamma[[2]];

y =  $\sqrt{dLen}$ 
  Max[Abs[cdflistMixedGamma - Range[1./dLen, 1, 1./dLen]]];

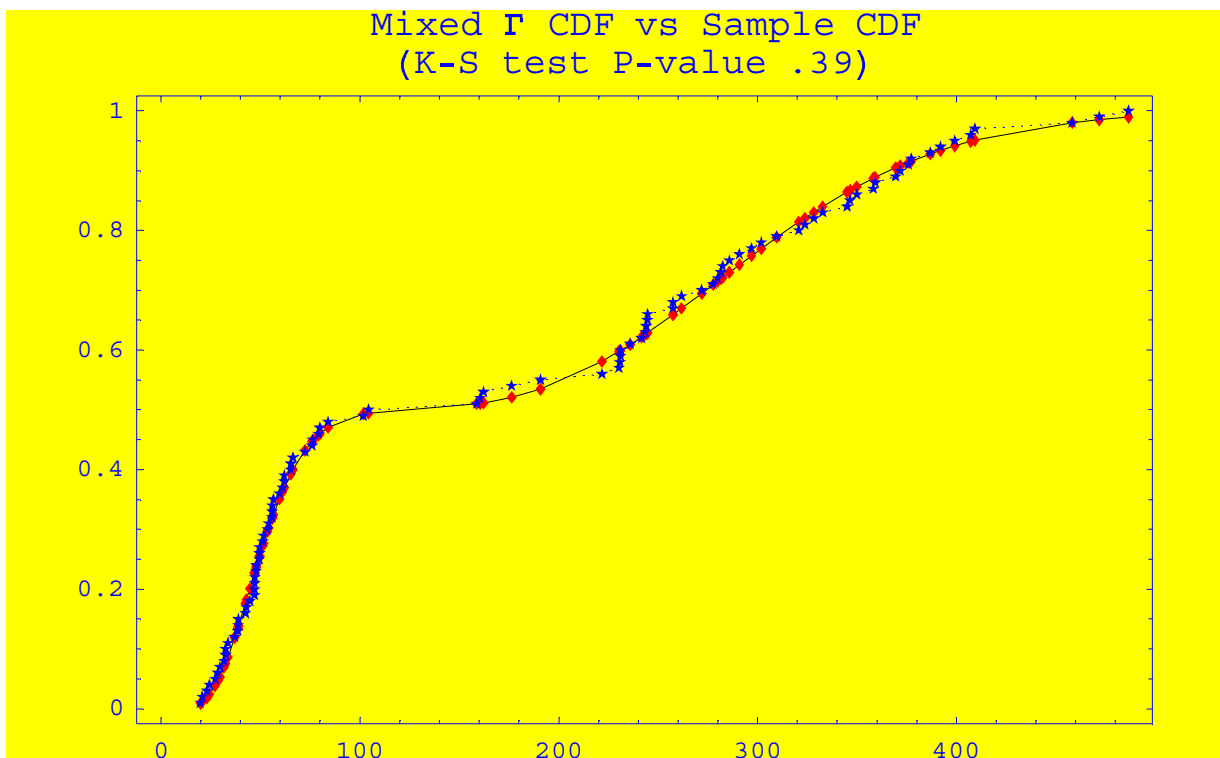
pValMixedGamma = NestWhile[(2 (-1) ^ (j + 1) Exp[-2 j ^ 2 y ^ 2] + #) &,
  j = 1; 0, (Abs[(-1) ^ j ++ (#1 - #2)] > 10 ^ (-8)) &, 2]
```

```
Out[95] = 0.999421
```

The high P-value obtained shows that the mixed distribution fits the data very well.

```
In[104] :=
```

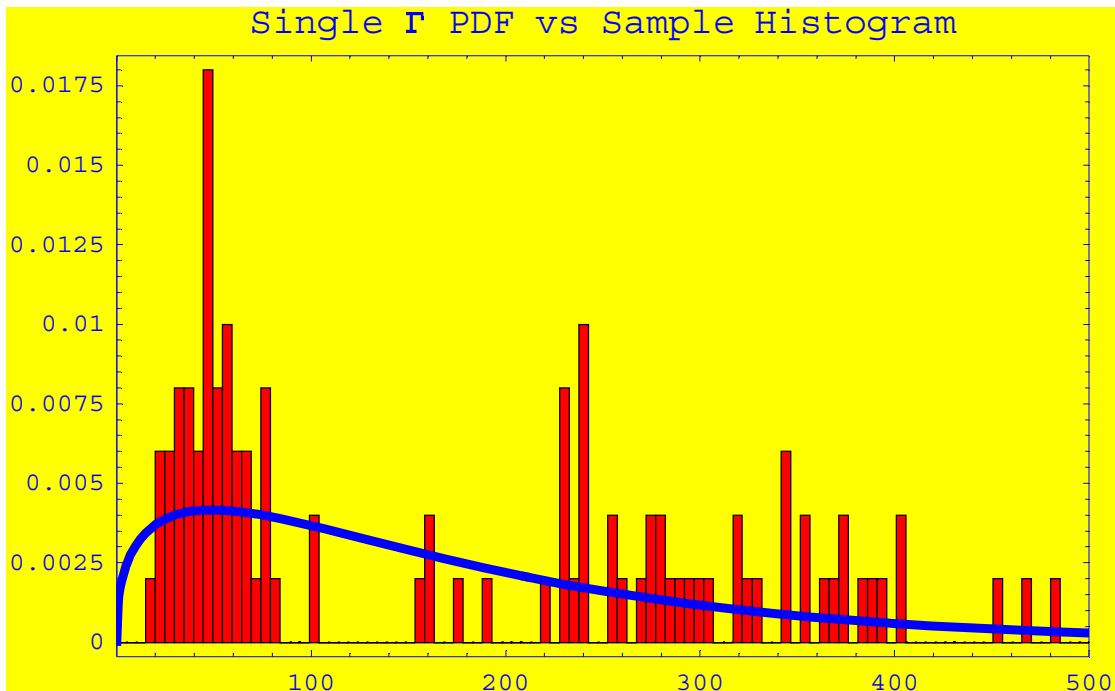
```
MultipleListPlot[Transpose[{data, cdflistMixedGamma}],  
  Transpose[{data, Range[1./dLen, 1, 1./dLen]}],  
  PlotJoined → True, SymbolStyle → {Red, Blue},  
  PlotLabel → StyleForm["Mixed  $\Gamma$  CDF vs Sample CDF\  
  (K-S test P-value .39)", FontSize → 14, FontWeight → Bold],  
  DefaultColor → Blue, Background → Yellow, Frame → True];
```

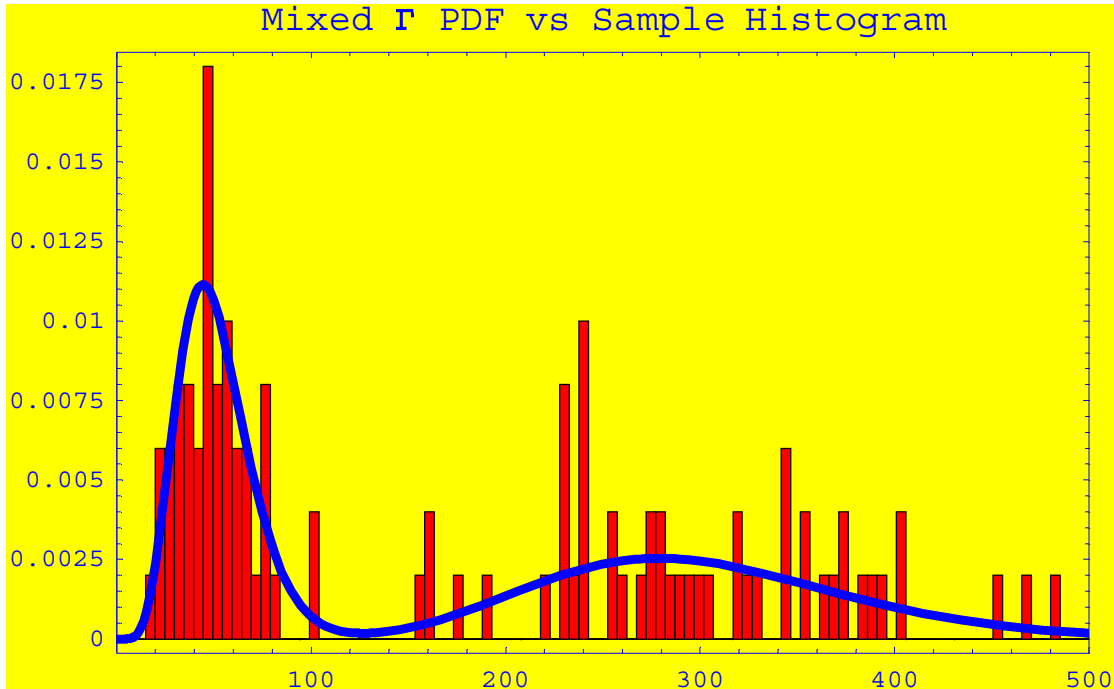


Next, we create a histogram of the data, and superimpose upon it the pdf for both the optimal Gamma density and the optimal mixed Gamma density.

In[131] :=

```
Show[Histogram[1 / dUpper Count[data, x_ /; # <= x && x < # + 5] & /@
  Range[0, dUpper, 5], FrequencyData -> True,
  HistogramRange -> {0, dUpper}, HistogramCategories -> Automatic,
  DisplayFunction -> Identity],
Plot[gammaPDF[α, θ, x] /. solGamma[[2]], {x, 0, dUpper},
  DisplayFunction -> Identity, PlotStyle -> {Thickness[.01]}],
PlotLabel -> StyleForm["Single Γ PDF vs Sample Histogram",
  FontSize -> 14, FontWeight -> Bold],
DisplayFunction -> $DisplayFunction, DefaultColor -> Blue,
Background -> Yellow, Frame -> True];
Show[Histogram[1 / dUpper Count[data, x_ /; # <= x && x < # + 5] & /@
  Range[0, dUpper, 5], FrequencyData -> True,
  HistogramRange -> {0, dUpper}, HistogramCategories -> Automatic,
  DisplayFunction -> Identity],
Plot[mixedGammaPDF[α1, θ1, α2, θ2, p, x] /. solMixedGamma[[2]],
  {x, 0, dUpper}, DisplayFunction -> Identity,
  PlotStyle -> {Thickness[.01]}],
DisplayFunction -> $DisplayFunction,
PlotLabel -> StyleForm["Mixed Γ PDF vs Sample Histogram",
  FontSize -> 14, FontWeight -> Bold],
DefaultColor -> Blue, Background -> Yellow, Frame -> True];
```





As is seen in this example, a single Gamma distribution does not fit the data well. However, an optimally chosen mixed distribution obtained as a convex combination of two Gamma distributions fits the data very well. Use of such mixed distributions to obtain better fits is major aspect of this project. *Mathematica's* powerful symbolic capabilities, optimization routines, data manipulation features, and flexible data visualization framework all combine to make it feasible to carry out calculations such as shown here.

Appendix A: Continuous Probability

Densities and Related Quantities

PDF : probability density function $f(x)$

Moments : the k -th moment is the expectation $E(X^k) = \int x^k f(x) dx$

Conditional Tail

$$\text{Expectations: } E(X | X < a) = \frac{\int_0^a x f(x) dx}{\int_0^a f(x) dx} \text{ and } E(X | X > b) = \frac{\int_b^\infty x f(x) dx}{\int_b^\infty f(x) dx}$$

1. Transformed Beta

$$f(x) = \frac{\Gamma(\alpha + \tau) \left(\gamma \left(\frac{x}{\theta}\right)^{\gamma\tau}\right)}{(\Gamma(\alpha) \Gamma(\tau)) \left(x \left(\left(\frac{x}{\theta}\right)^{\gamma} + 1\right)^{\alpha+\tau}\right)}$$

$$E(X^k) = \frac{\theta^k \Gamma(\tau + k/\gamma) \Gamma(\alpha - k/\gamma)}{\Gamma(\alpha) \Gamma(\tau)}, \quad -\tau\gamma < k < \alpha\gamma$$

2. Generalized Pareto

$$f(x) = \frac{\Gamma[\alpha + \tau]}{\Gamma[\alpha] \Gamma[\tau]} \frac{\theta^\alpha x^{\tau-1}}{(x + \theta)^{\alpha+\tau}}$$

$$E(X^k) = \frac{\theta^k \Gamma(\tau + k) \Gamma(\alpha - k)}{\Gamma(\alpha) \Gamma(\tau)}, \quad -\tau < k < \alpha$$

3. Burr

$$f(x) = \frac{\alpha \gamma \left(\frac{x}{\theta}\right)^\gamma}{x \left(\left(\frac{x}{\theta}\right)^\gamma + 1\right)^{\alpha+1}}$$

$$E(X^k) = \frac{\theta^k \Gamma(1 + k/\gamma) \Gamma(\alpha - k/\gamma)}{\Gamma(\alpha)}, \quad -\gamma < k < \alpha\gamma$$

4. Inverse Burr

$$f(x) = \frac{\tau \gamma \left(\frac{x}{\theta}\right)^{\gamma\tau}}{x \left(\left(\frac{x}{\theta}\right)^\gamma + 1\right)^{\tau+1}}$$

$$E(X^k) = \frac{\theta^k \Gamma(\tau + k/\gamma) \Gamma(1 - k/\gamma)}{\Gamma(\tau)}, \quad -\tau\gamma < k < \gamma$$

5. Pareto

$$f(x) = \frac{\alpha \theta^\alpha}{(x + \theta)^{\alpha+1}}$$

$$E(X^k) = \frac{\theta^k \Gamma(1 + k) \Gamma(\alpha - k)}{\Gamma(\alpha)}, \quad -1 < k < \alpha$$

6. Inverse Pareto

$$f(x) = \frac{\tau \theta x^{\tau-1}}{(x + \theta)^{\tau+1}}$$

$$E(X^k) = \frac{\theta^k \Gamma(\tau + k) \Gamma(1 - k)}{\Gamma(\tau)}, \quad -\tau < k < 1$$

7. Loglogistic

$$f(x) = \frac{\gamma \left(\frac{x}{\theta}\right)^\gamma}{x \left(\left(\frac{x}{\theta}\right)^\gamma + 1\right)^2}$$

$$E(X^k) = \theta^k \Gamma(1 + k/\gamma) \Gamma(1 - k/\gamma), \quad -\gamma < k < \gamma$$

8. Paralogistic

$$f(x) = \frac{\alpha^2 \left(\frac{x}{\theta}\right)^\alpha}{x \left(\left(\frac{x}{\theta}\right)^\alpha + 1\right)^{\alpha+1}}$$

$$E(X^k) = \frac{\theta^k \Gamma(1 + k/\alpha) \Gamma(\alpha - k/\alpha)}{\Gamma(\alpha)}, \quad -\alpha < k < \alpha^2$$

9. Inverse Paralogistic

$$f(x) = \frac{\tau^2 \left(\frac{x}{\theta}\right)^{\tau^2}}{x \left(\left(\frac{x}{\theta}\right)^\tau + 1\right)^{\tau+1}}$$

$$E(X^k) = \frac{\theta^k \Gamma(\tau + k/\tau) \Gamma(1 - k/\tau)}{\Gamma(\tau)}, \quad -\tau^2 < k < \tau$$

10. Transformed Gamma

$$f(x) = \frac{\tau \left(\frac{x}{\theta}\right)^{\alpha\tau} e^{-\left(\frac{x}{\theta}\right)^\tau}}{x \Gamma(\alpha)}$$

$$E(X^k) = \frac{\theta^k \Gamma(\alpha + k/\tau)}{\Gamma(\alpha)}, \quad -\alpha\tau < k$$

11. Inverse Transformed Gamma

$$f(x) = \frac{\tau \left(\frac{\theta}{x}\right)^{\alpha\tau} e^{-\left(\frac{\theta}{x}\right)^\tau}}{x \Gamma(\alpha)}$$

$$E(X^k) = \frac{\theta^k \Gamma(\alpha - k/\tau)}{\Gamma(\alpha)}, \quad \alpha\tau > k$$

$$f(x) = \frac{\left(\frac{x}{\theta}\right)^\alpha e^{-\frac{x}{\theta}}}{x \Gamma(\alpha)}$$

$$E(X^k) = \frac{\theta^k \Gamma(\alpha + k)}{\Gamma(\alpha)}, \alpha > -k$$

13. Inverse Gamma

$$f(x) = \frac{\left(\frac{\theta}{x}\right)^\alpha e^{-\frac{\theta}{x}}}{x \Gamma(\alpha)}$$

$$E(X^k) = \frac{\theta^k \Gamma(\alpha - k)}{\Gamma(\alpha)}, \alpha > k$$

14. Weibull

$$f(x) = \frac{\tau \left(\frac{x}{\theta}\right)^{\tau-1} e^{-\left(\frac{x}{\theta}\right)^\tau}}{x}$$

$$E(X^k) = \theta^k \Gamma(1 + k/\tau), k > -\tau$$

15. Inverse Weibull

$$f(x) = \frac{\tau \left(\frac{\theta}{x}\right)^{\tau-1} e^{-\left(\frac{\theta}{x}\right)^\tau}}{x}$$

$$E(X^k) = \theta^k \Gamma(1 - k/\tau), k < \tau$$

16. Exponential

$$f(x) = \frac{e^{-\frac{x}{\theta}}}{\theta}$$

$$E(X^k) = \theta^k \Gamma(1 + k), k > -1$$

17. Inverse Exponential

$$f(x) = \frac{\theta e^{-\frac{\theta}{x}}}{x^2}$$

$$E(X^k) = \theta^k \Gamma(1 - k), k < 1$$

18. Lognormal

$$f(x) = \frac{e^{-\frac{(\log(x)-\mu)^2}{2\sigma^2}}}{x\sigma\sqrt{2\pi}}$$

$$E(X^k) = \exp(k\mu + k^2\sigma^2/2)$$

19. Inverse Gaussian

$$f(x) = \sqrt{\frac{\theta}{2\pi x^3}} e^{-\frac{\theta(x-\mu)^2}{2x\mu^2}}$$

$$E(X) = \mu, \text{Var}(X) = \frac{\mu^3}{\theta}$$

20. Single - parameter Pareto

$$f(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}$$

$$E(X^k) = \frac{\alpha\theta^k}{\alpha - k}, k < \alpha$$

21. Generalized beta

$$f(x) = \frac{\Gamma(a+b) \left(\frac{x}{\theta}\right)^{\tau a} \left(1 - \left(\frac{x}{\theta}\right)^\tau\right)^{b-1} \tau}{(\Gamma(a)\Gamma(b)) x}$$

$$E(X^k) = \frac{\theta^k \Gamma(a+b) \Gamma(a+k/\tau)}{\Gamma(a)\Gamma(a+b+k/\tau)}, k > -a\tau$$

22. Beta

$$f(x) = \frac{\Gamma(a+b) \left(\frac{x}{\theta}\right)^a \left(1 - \frac{x}{\theta}\right)^{b-1}}{(\Gamma(a)\Gamma(b)) x}$$

$$E(X^k) = \frac{\theta^k \Gamma(a+b) \Gamma(a+k)}{\Gamma(a)\Gamma(a+b+k)}, k > -a$$

23. Mixed (1 - 22)

$$f(x) = p f_1(x) + (1-p) f_2(x), 0 < p < 1$$

Null

References

- [LM] Klugman, Stuart A., Panjer, H. H., and Willmot, G. E., Loss Models, Wiley-Interscience (1998).
- [ALC] Longley-Cook, A. G., Efficient Stochastic Modeling Using Representative Scenarios: Application to Equity Risks. The Stochastic Modeling Symposium (2003).
- [YC] Chueh, Y., Efficient Stochastic Modeling for Large and Consolidated Insurance Business: Interest Rate Sampling Algorithms. North American Actuarial Journal, Volume 6, no. 3 (2002).