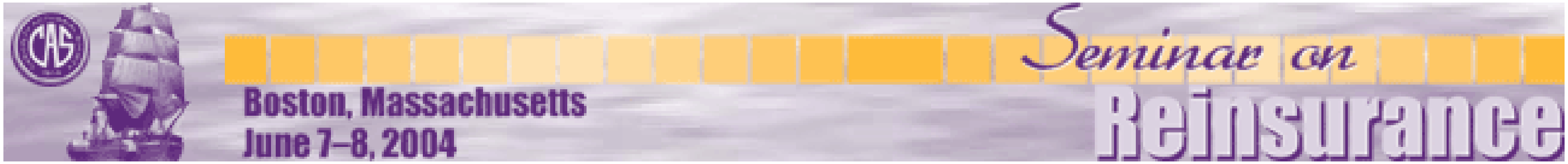


June 7-8, 2004

Uncertainty in Actuarial Modeling
An Applied Approach

Steve White
2004 CARe Meeting - Boston

Section 1

Uncertainty

What is Modeling Uncertainty

In **Loss Models from data to decisions** we have the following two definitions

A ***mathematical model*** is an abstract and simplified representation of a given phenomenon that can be expresses in mathematical terms

A ***stochastic model*** is a mathematical model for a phenomenon displaying statistical regularity that can accurately describe the probabilities of outcomes

What is Modeling Uncertainty

Some Types of Risk

- Model Selection Risk
  - Is our abstract simplification reasonably predictive?
  - Did we choose the wrong model?
- Parameter Risk
  - Even if we are happy with the model, are we using the right parameters in the model
  - If we think that the parameters are reasonable on average, do we think that they are an exact value or could they have a range of values
- Process Risk
  - The nature of risk is that the results are random even if we have the “right” model and the “right” parameters

Some Areas where you can include Uncertainty in Modeling

As Actuaries it is a Standard of Practice that a Loss Reserve estimate should be a range of values (uncertainty) rather than a single point estimate

Yet in our other work we often resort back to point estimates

Curve Fitting

- use of the MLE estimates (the most likely or modal value)

Experience Rating

- May report a few values, paid vs incurred, BF vs CL

Exposure Rating

- Usually report a single value

With more sensible estimate of uncertainty

- Easier to do a minimum variance credibility weighting of estimates
- Less likely to fall into the trap of understating the volatility in Aggregate Loss/ DFA models

Some Areas where you can include Uncertainty in Modeling

- Curve Fitting
  - Parameter Uncertainty
  - Parameter Correlation
  - Trend
  - Development
- Experience Rating
  - Loss Trend
  - Loss Development
  - Limits Profile
  - Exposure Trend
  - Premium Trend
  - Rate Adequacy
- Exposure Rating
  - Loss Trend
  - ALAE Treatment
  - Limits Profile
  - Rate Adequacy (Loss Ratio)

Section 2

Tools for Modeling Uncertainty

Common Tools for Modeling Uncertainty

- Simulation
- Mixing Distributions
  - Theoretical Mixing
  - Numerical Mixing

Common Tools for Modeling Uncertainty
Simulation-Notation

$f(x;\varphi)$ probability density function of x with parameter(s) φ

$F(X;\varphi)=\int_0^X f(x)dx$ cumulative distribution function of x with parameter(s) φ

$p=F(X;\varphi)$ probability, p, that x ≤ X

$X=F^{-1}(p;\varphi)$ Inverse of the F(X) which finds the X value associated with p
This is the function used for simulation.
Some F(X)'s are directly invertible. Others require other Methods to simulate.

Simple Example – Exponential Distribution

$$f(x;\theta)=\frac{1}{\theta}e^{-x/\theta}$$

$$F(X;\theta)=1-e^{-x/\theta}$$

$$F^{-1}(p,\theta)=-\theta\, ln(1-p)$$

Common Tools for Modeling Uncertainty

Modeling Uncertainty with Simulation

Let's assume that we are happy with the exponential distribution as our choice for the severity distribution.

But we are not sure about the value for θ, the mean of the exponential distribution.

You can assume that θ also has a distribution function G(θ) and a corresponding $G^{-1}(p_\theta)$.

The process for each simulated year will be

1) draw a random p_θ from a Uniform(0,1)
2) $\Theta=G^{-1}(p_\theta)$.
3) draw a random p_{Exp} from a Uniform(0,1)
4) $X=F^{-1}(p_{Exp};\theta)$

This process can be described as mixing

Common Tools for Modeling Uncertainty

Mixing Distributions

$$h(y;\phi,\theta) = \int_{\psi} f(y;\phi,\psi) g(\psi;\theta) d\psi$$

$$h(y;\phi,\theta) = f(y;\phi,c \mid c = \psi) g(\psi;\theta)$$

$$h(y;\phi,\theta) = f(y;\phi,\psi) \hat{\psi} \, g(\psi;\theta)$$

$f(y;\phi,\psi)$ structural loss distribution with independent parameter(s) ϕ and dependent parameter(s) ψ

$g(\psi;\theta)$ mixing distribution on parameter(s) ψ with it's own parameter(s) $\boldsymbol{\theta}$

$h(y;\phi,\theta)$ Resulting (mixed) distribution with parameters ϕ and $\boldsymbol{\theta}$

The basic idea is that you assume the losses are from a given distribution f(y) of a known form.

f(y) has parameter(s) ϕ are fixed and parameter(s) ψ which are not fixed values.

ψ has it's own distribution g(ψ) with parameter(s) $\boldsymbol{\theta}$.

When you mix (combine) these two distributions, the distribution h(y) will depend on the form of structural and mixing distributions and on the parmater(s) ϕ and $\boldsymbol{\theta}$.

h(y)may or may not have a recognizable form, it can be very useful when h(y) has a known form

Common Tools for Modeling Uncertainty

Some Theoretical Mixing Distribution Combinations

Probably the best known Mixing Distribution Combination

$$NB(n;\theta,\beta) = P(n;\lambda)\ \hat{\lambda}\ Ga(\lambda;\theta,\beta)$$

Other Mixing Distribution Combinations

$$TB(y;\tau,\theta,\beta,\alpha) = TG(y;\tau,\psi,\beta)\ \hat{\psi}\ InvTG(\psi;\tau,\theta,\alpha)_{1,2}$$

$$GP(y;\theta,\beta,\alpha) = GA(y;\psi,\beta)\ \hat{\psi}\ InvGA(\psi;\theta,\alpha);\tau = 1_2$$

$$Burr(y;\theta,\beta,\alpha) = W(y;\psi,\beta)\ \hat{\psi}\ InvTG(\psi;\beta,\theta,\alpha);\tau = \beta_2$$

$$InvBurr(y;\theta,\beta,\alpha) = TG(y;\alpha,\psi,\beta)\ \hat{\psi}\ InvW(\psi;\theta,\alpha);\tau = \alpha_2$$

$$BP(y;\theta,\alpha) = Exp(y;\psi)\ \hat{\psi}\ InvGA(\psi;\theta,\alpha)_2$$

$$LL(y;\theta,\alpha) = W(y;\alpha,\psi)\ \hat{\psi}\ InvW(\psi;\theta,\alpha);\tau = \alpha = \beta_2$$

$$LT(y;\mu,\sigma,q) = LN(y;\mu,\psi)\ \hat{\psi}\ InvGA(\psi;\sigma^2 q,q)_2$$

$$LN(y;\mu,\sqrt{\sigma_s^2+\sigma_m^2}) = LN(y;\psi,\sigma_s)\ \hat{\psi}\ LN(\psi;\mu,\sigma_m)_3$$

1 - Venter, Gary "Transformed Beta and Gamma Distributions and Aggregate Lossed". Proceedings of the CAS (1984), 156-193

2 – McDonald, James B and Butler, Richard J " Some Generalized Mixture Distributions with an Application to Unemployment Duration" The Review and Economics and Statistics, Vol LXIX-2 (May 1987), 232-240

3 – Foundations of Casualty Actuarial Science, 3rd edition, 490

NB-Negative Binomial, P – Poisson, Ga – Gamma, TB – Transformed Beta, TG – Transformed Gamma, ITG – Inverse Transformed Gamma
GB – Generalized Pareto, IGa – Inverse Gamma, Burr – Burr, W – Weibull, IBurr – Inverse Burr, IW – Inverse Weibull, BP – Ballasted Pareto
Exp – Exponential, LL – Log Logistic, LT – Log T, LN – Log Normal (See Appendix for further definition of the distributions)

Mixing Distributions

Theoretical Mixing Distribution Combinations Parameterization – pdf's

Resulting Distribution	Structural Distribution	Mixing Distribution
$NB(n;\theta,\alpha)=\binom{n+\alpha-1}{n}\left(\frac{\theta}{1+\theta}\right)^{n}\left(\frac{1}{1+\theta}\right)^{\alpha}$	$P(n;\lambda)=\frac{\lambda^{n}}{n!}e^{-\lambda}$	$Ga(\lambda;\theta,\alpha)=\frac{1}{\theta}\frac{(\lambda/\theta)^{\alpha-1}}{\Gamma(\alpha)}e^{-\lambda/\theta}$
$TB(x;\tau,\theta,\beta,\alpha)=\frac{\tau}{\theta}\frac{(x/\theta)^{\beta-1}}{\left(1+(x/\theta)^{\tau}\right)^{(\alpha+\beta)/\tau}}\frac{\Gamma((\alpha+\beta)/\tau)}{\Gamma(\alpha/\tau)\Gamma(\beta/\tau)}$	$TG(x;\tau,\psi,\beta)=\frac{\tau}{\psi}\frac{(x/\psi)^{\beta-1}}{\Gamma(\beta/\tau)}e^{-(x/\psi)^{\tau}}$	$InvTG(\psi;\tau,\theta,\alpha)=\frac{\tau}{\theta}\frac{(\theta/\psi)^{\alpha+1}}{\Gamma(\alpha/\tau)}e^{-(\theta/\psi)^{\tau}}$
$GP(x;\theta,\beta,\alpha)=\frac{1}{\theta}\frac{(x/\theta)^{\beta-1}}{(1+(x/\theta))^{(\alpha+\beta)}}\frac{\Gamma((\alpha+\beta))}{\Gamma(\alpha)\Gamma(\beta)}$	$GA(x;\psi,\beta)=\frac{1}{\psi}\frac{(x/\psi)^{\beta-1}}{\Gamma(\beta)}e^{-(x/\psi)}$	$InvGA(\psi;\theta,\alpha)=\frac{1}{\theta}\frac{(\theta/\psi)^{\alpha+1}}{\Gamma(\alpha)}e^{-\theta/\psi}$
$Burr(x;\theta,\beta,\alpha)=\frac{\alpha}{\theta}\frac{(x/\theta)^{\beta-1}}{\left(1+(x/\theta)^{\beta}\right)^{(\alpha+\beta)/\beta}}$	$W(x;\psi,\beta)=\frac{\beta}{\psi}(x/\psi)^{\beta-1}e^{-(x/\psi)^{\beta}}$	$InvTG(\psi;\beta,\theta,\alpha)=\frac{\beta}{\theta}\frac{(\theta/\psi)^{\alpha+1}}{\Gamma(\alpha/\beta)}e^{-(\theta/\psi)^{\beta}}$
$InvBurr(x;\theta,\beta,\alpha)=\frac{\beta}{\theta}\frac{(x/\theta)^{\beta-1}}{\left(1+(x/\theta)^{\alpha}\right)^{(\alpha+\beta)/\alpha}}$	$TG(x;\alpha,\psi,\beta)=\frac{\alpha}{\psi}\frac{(x/\psi)^{\beta-1}}{\Gamma(\beta/\alpha)}e^{-(x/\psi)^{\alpha}}$	$InvW(\psi;\theta,\alpha)=\frac{\alpha}{\theta}(\theta/\psi)^{\alpha+1}e^{-(\theta/\psi)^{\alpha}}$
$BP(x;\theta,\alpha)=\frac{\alpha}{\theta}\frac{1}{(1+(x/\theta))^{(\alpha+1)}}$	$Exp(x;\psi)=\frac{1}{\psi}e^{-(x/\psi)}$	$InvGA(\psi;\theta,\alpha)=\frac{1}{\theta}\frac{(\theta/\psi)^{\alpha+1}}{\Gamma(\alpha)}e^{-\theta/\psi}$
$LL(x;\theta,\alpha)=\frac{\alpha}{\theta}\frac{(x/\theta)^{\alpha-1}}{\left(1+(x/\theta)^{\alpha}\right)^{2}}$	$W(x;\psi,\beta)=\frac{\beta}{\psi}(x/\psi)^{\beta-1}e^{-(x/\psi)^{\beta}}$	$InvW(\psi;\theta,\alpha)=\frac{\alpha}{\theta}(\theta/\psi)^{\alpha+1}e^{-(\theta/\psi)^{\alpha}}$
$LT(x;\mu,\sigma,\nu)=\frac{1}{x\sqrt{2\nu}\sigma\left(1+\left(\frac{\ln(x)-\mu}{\sqrt{2\nu}\sigma}\right)^{2}\right)}\frac{\Gamma((\nu+1/2))}{\Gamma(\nu)\Gamma(1/2)}$	$LN(x;\mu,\psi)=\frac{1}{x\psi\sqrt{2}}e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\psi}\right)^{2}}$	$InvGA(\psi;\sigma^{2}\nu,\nu)=\frac{1}{\sigma^{2}\nu}\frac{(\sigma^{2}\nu/\psi)^{\nu+1}}{\Gamma(\nu)}e^{-(\sigma^{2}\nu/\psi)}$

Common Tools for Modeling Uncertainty
Numerical Mixing Methods

The mixing distributions don't help the choice of f() and g() do not yield a recognizable h()

$$h(y;\phi,\theta) = f(y;\phi,\psi)\ \hat{\psi}\ \ g(\psi;\theta)$$

We have already discussed a simulation approach to mixing

$$H^{-1}(u_f;\phi,\theta) \cong F^{-1}(u_f;\phi,G^{-1}(u_g;\theta))$$

A Numerical Integration approach to mixing can give very nice results
Without being as computer time intensive as simulation

$$h(y;\phi,\theta) = \int_{\psi} f(y;\phi,\psi)g(\psi;\theta)d\psi$$

If ψ is a vector, then multivariate integration is required

Gaussian Integration seems to work very nicely. The number of points required will depend of the shape of the mixing distribution. Multivariate mixing distributions tend to require more points. A seven point Gaussian Integration has given good convergence on a two parameter mixing distribution

Common Tools for Modeling Uncertainty

Numerical Integration "Quadrature"

Common numerical integration methods covered in the Numerical Analysis portion of the Actuarial Exams

- Simpson's Rule
- Trapezoidal Rule
- Romberg Rule

These methods will often require many terms to converge and do not work well over an indefinite range $(0,\infty)$ or $(-\infty,\infty)$ and therefore are of limited use for modeling parameter uncertainty

- Gaussian Quadrature can be defined to work well over an indefinite interval using relatively few points

Common Tools for Modeling Uncertainty

Numerical Integration - Gaussian Integration Values for Normal Distribution

$$h(x) = \int f(x;\psi)g(\psi)dx$$ The Mixing Equation

$$h(x) = \sum_i w_i f(x; z_i)$$ If the Mixing distribution is Standard Normal

$$h(x) = \sum_i w_i f(x; G^{-1}(p_i))$$ If the Mixing distribution is other than Normal

3 Point			5 Point			7 Point		
w	z	p	w	z	p	w	z	p
0.166667	-1.732051	0.041632	0.011257	-2.856970	0.002139	0.000548	-3.750440	0.000088
0.666667	0.000000	0.500000	0.222076	-1.355626	0.087609	0.030757	-2.366759	0.008972
0.166667	1.732051	0.958368	0.533333	0.000000	0.500000	0.240123	-1.154405	0.124167
			0.222076	1.355626	0.912391	0.457143	0.000000	0.500000
			0.011257	2.856970	0.997861	0.240123	1.154405	0.875833
						0.030757	2.366759	0.991028
						0.000548	3.750440	0.999912

For background or more points see any of the following:

Abramowitz, Milton and Stegun, Irene – Handbook of Mathematical Tables

Press, William and Flannery, Brian - Numerical Recipes in C (available in other languages)

Burden, Richard and Fiares, Douglas – Numerical Analysis

Common Tools for Modeling Uncertainty

A Simple Example – Ballasted Pareto as a Mixture

Ballasted Pareto as a function of an Exponential distribution mixed with an Inverse Gamma distribution

Choosing Parameters for the distributions

The Exponential has an assumed mean, ***μ***

- Select parameters for the Inverse Gamma Distribution
- You can choose and Inverse Gamma with a mean, μ, and an assumed CV
  - $\alpha = 2+1/CV^2$, $\theta=\mu/(\alpha-1)$
  - Inverse Gamma parameters define the Ballasted Pareto Parameters
  - Intuitive approach based on CV
  - 2nd moment exists for Ballasted Pareto
  - For a thicker tailed Ballasted Pareto, you can select an $\alpha < 2$
  - If you select an $\alpha<1$ an unconstrained distribution will have undefined (infinite) mean and the results can be very unstable

Common Tools for Modeling Uncertainty
A Simple Example – Ballasted Pareto

Theoretical Mixing – we can directly model the Ballasted Pareto

$$BP(y;\theta,\alpha) = Exp(y;\mu)\,\hat{\mu}\, InvGA(\mu;\theta,\alpha)$$

Simulation – First Simulate the mean of the exponential, then simulate from the exponential distribution

$$y = BP^{-1}(u_{BP};\theta,\alpha) = Exp^{-1}\left(u_{BP}; InvGA^{-1}\left(u_{IG};\theta,\alpha\right)\right)$$

Numerical Integration – Evaluate the Exponential at a few points and then weigh them together to estimate the Ballasted Pareto

$$BP(y;\theta,\alpha) = \int_{\mu} Exp(y;\mu)\, InvGA(\mu;\theta,\alpha)$$

$$BP(y;\theta,\alpha) = \sum_{i} w_i Exp(y;\mu_i); \quad \mu_i = InvGA^{-1}(p_i;\theta,\alpha)$$

Common Tools for Modeling Uncertainty

A Simple Example - Comparing Methods

Using all three methods compare the results for the following

$$BP(y;\theta,\alpha) = Exp(y;\mu)\,\hat{\mu}\,InvGA(\mu;\theta,\alpha)$$

$$Mean_{IGa} = 10{,}000;\ \ CV_{IGa} = 2;$$

$$\alpha_{IGa} = 2 + (1/2)^2 = 2.25;\ \ \theta_{BP} = Mean_{IGa} \times (\alpha_{IGa} - 1) = 12{,}500$$

Compare the Following Statistics

- E(x)
- Var(x)
- F(x); x=1K, 10K, 100K, 1M
- LAS(x); x=1K, 10K, 100K, 1M

Using the following methods

- •Theoretical Mixing
- •Simulation - 1K and 10k draws from the InvGamma, each with 20K Exponential draws
- •Gaussian Integration (3 Point, 5 Point and 7 Point)

Common Tools for Modeling Uncertainty

A Simple Example - Comparing Methods/Summarized Results

	Exponential	Ballasted Pareto	Gaussian Quadrature 3 Pt	Gaussian Quadrature 5 Pt	Gaussian Quadrature 7 Pt	Sim(1K) Mean	Sim(10K) Mean
E(X)=	10,000	10,000	9,685	9,973	9,998	9,742	9,803
StdDev(X)=	10,000	30,000	16,151	21,683	24,833	20,529	21,515
F(1,000)=	0.09516	0.15900	0.15896	0.15900	0.15900	0.16020	0.15972
F(10,000)=	0.63212	0.73354	0.73582	0.73364	0.73354	0.73533	0.73549
F(100,000)=	0.99995	0.99287	0.99416	0.99344	0.99266	0.99339	0.99331
F(1,000,000)=	1.00000	0.99995	1.00000	1.00000	0.99992	0.99999	0.99997
LAS(1,000)=	951.63	917.19	917.20	917.19	917.19	916.55	916.79
LAS(10,000)=	6,321.21	5,203.67	5,197.65	5,203.41	5,203.66	5,187.61	5,188.13
LAS(100,000)=	9,999.55	9,358.50	9,511.15	9,328.19	9,355.96	9,209.60	9,263.91
LAS(1,000,000)=	10,000.00	9,958.85	9,685.34	9,972.34	9,956.41	9,738.12	9,794.51
		Pct. Error	Pct. Error	Pct. Error	Pct. Error	Pct. Error	Pct. Error
E(X)=		0.00%	-3.15%	-0.27%	-0.02%	-2.58%	-1.97%
StdDev(X)=		0.00%	-46.16%	-27.72%	-17.22%	-31.57%	-28.28%
F(1,000)=		0.00%	-0.03%	0.00%	0.00%	0.75%	0.45%
F(10,000)=		0.00%	0.31%	0.01%	0.00%	0.24%	0.27%
F(100,000)=		0.00%	0.13%	0.06%	-0.02%	0.05%	0.04%
F(1,000,000)=		0.00%	0.01%	0.00%	0.00%	0.00%	0.00%
LAS(1,000)=		0.00%	0.00%	0.00%	0.00%	-0.07%	-0.04%
LAS(10,000)=		0.00%	-0.12%	0.00%	0.00%	-0.31%	-0.30%
LAS(100,000)=		0.00%	1.63%	-0.32%	-0.03%	-1.59%	-1.01%
LAS(1,000,000)=		0.00%	-2.75%	0.14%	-0.02%	-2.22%	-1.65%

7 point integration is a good approximation for all of the statistics except for the standard deviation (but better than simulation in these cases). This may be causes because based on the α used, 2.25, the variance is close to being undefined which is anytime $\alpha \leq 2$.

Common Tools for Modeling Uncertainty

A Simple Example - Comparing Methods/Summarized Results

z_i	3.7504	2.3668	1.1544	0.0000	-1.1544	-2.3668	-3.7504			
w_i	0.00055	0.03076	0.24012	0.45714	0.24012	0.03076	0.00055			
p_i	0.9999	0.9910	0.8758	0.5000	0.1242	0.0090	0.0001			
α_i										2.2500
θ_i	519,029	63,068	16,652	6,487	3,147	1,730	1,003		10,000	12,500
	Exp_1	Exp_2	Exp_3	Exp_4	Exp_5	Exp_5	Exp_7	Wgtd	Exp	BP
Limit	LAS	LAS	LAS	LAS	LAS	LAS	LAS	LAS	LAS	LAS
10,000	9,904.28	9,247.50	7,517.97	5,098.54	3,015.95	1,724.79	1,003.07	5,203.65	6,321.20	5203.66342
100,000	90,956.61	50,150.19	16,610.53	6,487.18	3,147.16	1,730.13	1,003.12	9,355.95	9,999.53	9,358.49
	CDF	CDF	CDF	CDF	CDF	CDF	CDF	CDF	CDF	CDF
10,000	0.01908	0.14663	0.45149	0.78594	0.95831	0.99691	0.99995	0.73354	0.63212	0.73354
100,000	0.17524	0.79517	0.99753	1.00000	1.00000	1.00000	1.00000	0.99266	0.99995	0.99287

z_i & w_i – Gaussian Integration constants
p_i – $F(z_i)$, where $F(z)$ is the standard normal CDF
θ_i – Mean of the ith Exponential = $IG^{-1}(p_i;\theta,\alpha)$, where IG^{-1} is the Inverse CDF of the Inverse Gamma distribution.

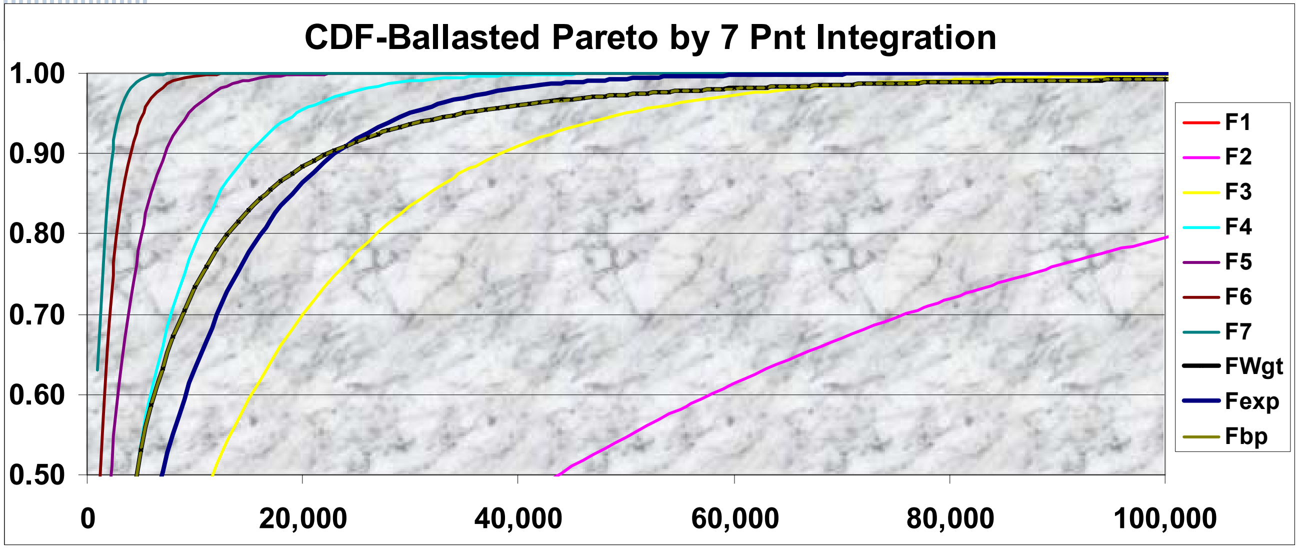
$$CDF_{\exp}(x;\theta_i)=\left(1-e^{-x/\theta_i}\right) \qquad CDF_{BP}(x)=1-\left(1+x/\theta\right)^{-\alpha} \qquad WgtCDF_{Exp}(x)=\sum_i w_i\left(1-e^{-x/\theta_i}\right)$$

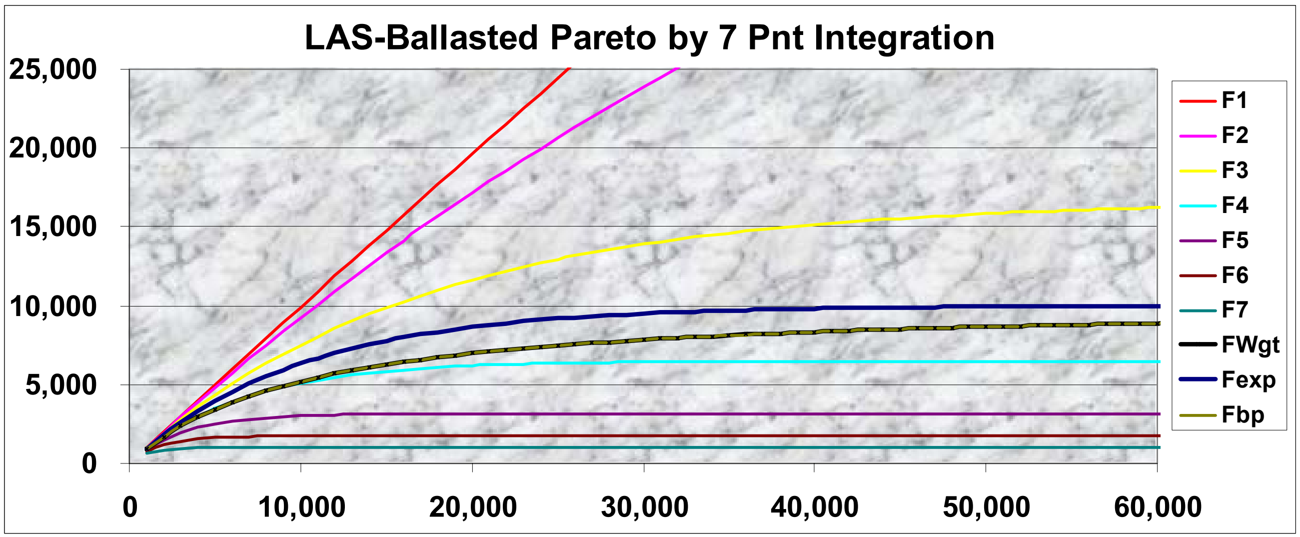
$$LAS_{\exp}(x)=\theta_i\left(1-e^{-x/\theta_i}\right) \qquad LAS_{BP}(x)=\frac{\theta}{1-\alpha}\left(\left(\frac{x+\theta}{\theta}\right)^{1-\alpha}-1\right) \qquad WgtLAS_{Exp}(x)=\sum_i w_i\theta_i\left(1-e^{-x/\theta_i}\right)$$

Simple weighting works for E(X), LAS(X), CDF(X)
To estimate Variance, you must estimate wgtd $E(X^2)$ and wgtd E(X) then
calculate wgtd $Var(X)=\text{wgtd } E(X^2)-[\text{wgtd } E(X)]^2$

Common Tools for Modeling Uncertainty

A Simple Example - Comparing Methods/Summarized Results





Section 3

Fitting Size of Loss Distributions

Fitting Size of Loss Distributions
Considerations

Considerations when curve fitting

- Fewer Parameters are better unless significant improvement is gained by the additional parameters (Parsimony)
- It is not good enough to only look at the most likely parameter values (the MLE predictors)
- Testing the significance of additional parameters
  - Likelihood Ratio Test
  - T-Test
- Many common distributions have correlated parameters. Correlation of these parameters adds to the complexity of modeling the loss amounts. Ignoring the correlation is wrong.
- MLE parameter estimates are asymptotically normally distributed

Fitting Size of Loss Distributions
Sources of Uncertainty

Some Sources of Uncertainty when Curve Fitting

- Parameter Uncertainty
- Parameter Correlation

Other Factors not directly addressed here but could by adding an additional dimension to the uncertainty

- Severity Trend
- Limits Profile
- Severity Loss Development

Fitting Size of Loss Distributions
Maximum Likelihood

The process of fitting a Size of Loss distribution using maximum likelihood is to find the set of Parameters, ψ, that maximizes the likelihood function L.

For complete individual data the Likelihood function is defined as follows where f(x;ψ) is the probability density function

$$L = \prod_i f(x_i;\psi)$$

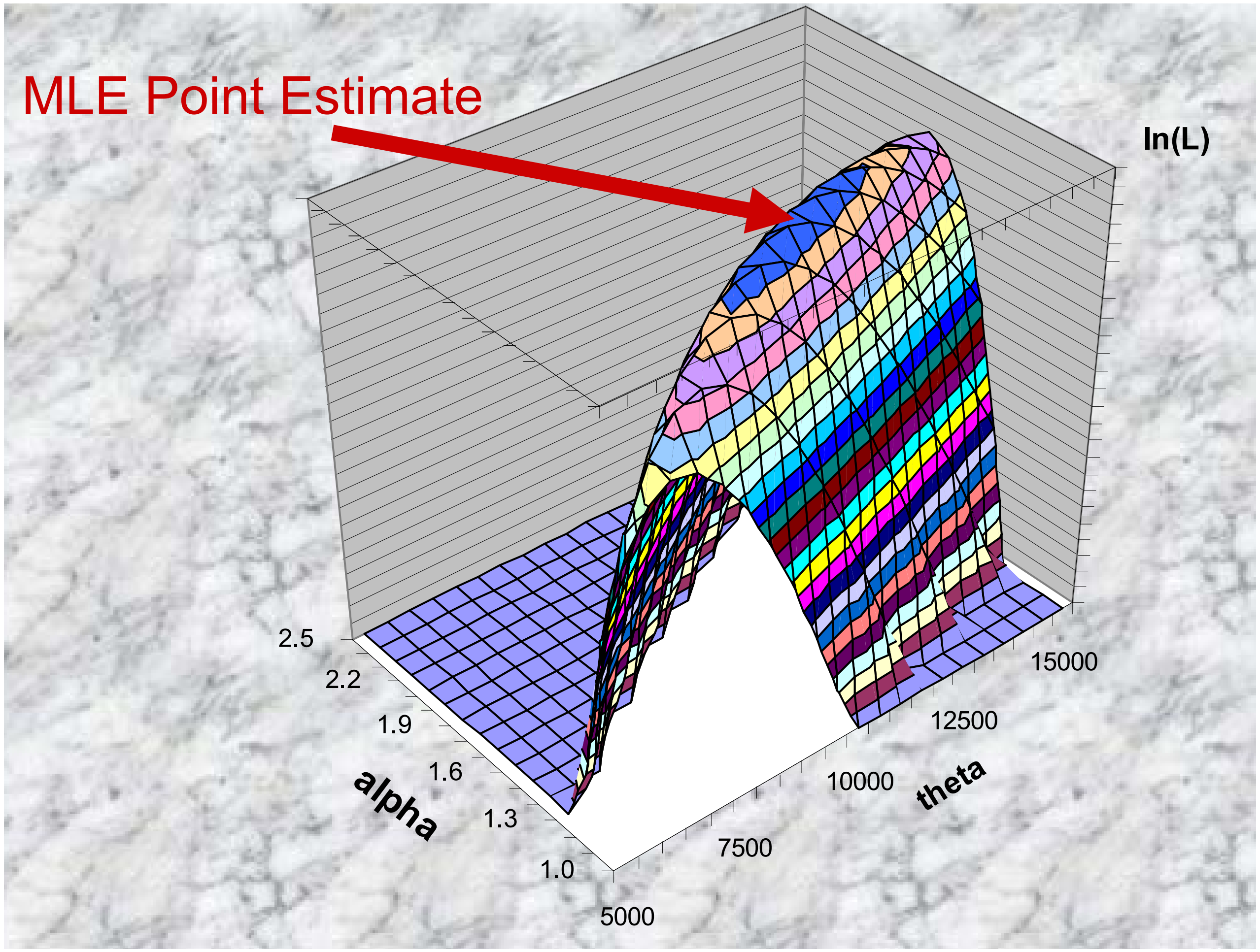
Generally it is easier to work with the log of the likelihood function, ℓ. It is equivalent to fit the Likelihood function, L, or the log-likelihood function, ℓ.

$$\lambda = \ln\left(\prod_i f(x_i;\psi)\right) = \sum_i \ln\left(f(x_i;\psi)\right)$$

Once you have found the parameters, $\tilde{\psi}$ that maximizes the likelihood function. You need to estimate the covariance matrix, **V**, of the parameters in order to model the uncertainty of those parameters.

Fitting Size of Loss Distributions
Likelihood Contour for a Ballasted Pareto

Ballasted Pareto Log-Likelihood Contour



When fitting a Size of Loss Distribution. Often users only use the point estimate and do not take into account that the parameters actually have a distribution.
The maximum likelihood estimates asymptotically follow a normal distribution.

As you can see from this likelihood plot, not only do you need estimates of the standard deviations for each of the parameters, you need estimates of the correlation between each of the parameters.

Fitting Size of Loss Distributions

Estimating the Co-Variance Matrix

The covariance matrix, V, for the parameter(s) can be estimated from the information matrix A, which is the expected value of the second derivative matrix for the log-likelihood function, ℓ.

$$a_{i,j} = -n \times E\left[\frac{\partial^2 \ln f(x;\tilde{\psi})}{\partial \psi_i \partial \psi_j}\right]$$

Approximations to the information Matrix

$$\hat{a}_{i,j} = -\sum_{k=1}^{n} \frac{\partial^2 \ln f\left(x_k;\tilde{\psi}\right)}{\partial \psi_i \partial \psi_j}$$

You can also approximate the information Matrix using Numerical Differentiation
The covariance matrix is calculated by taking the Matrix inverse of the information matrix.
The diagonal elements on the covariance matrix are the variance of the individual parameter. The off-diagonal elements define the covariances between parameters

$$V = [A]^{-1}; \quad \sigma_i = \sqrt{v_{i,i}}; \quad \rho_{i,j} = \frac{v_{i,j}}{\sigma_i \sigma_j}, i \neq j$$

Fitting Size of Loss Distributions

Estimating Parameter Uncertainty-A Ballasted Pareto example

$$a_{\theta,\theta} = \frac{\alpha}{\theta^2(2+\alpha)} \qquad a_{\theta,\alpha} = -\frac{1}{\theta(1+\alpha)} \qquad a_{\alpha,\alpha} = \psi'(\alpha) - \psi'(1+\alpha)$$

Kleiber, Christian and Kotz, Samuel - Statistical Size Distributions in Economics and Actuarial Sciences
The also have the information matrices for the Transformed Beta(GB2), Transformed Gamma(GG), Log-Normal and related distributions.

$$\psi'(\alpha) = \frac{\partial^2 \ln(\Gamma(\alpha))}{\partial \alpha^2}$$

$\psi'(\alpha)$ is the trigamma function, the second derivative of the natural log of the gamma function, $\Gamma(\alpha)$. See Abramowitz and Stegun, Handbook of Mathematical Tables

The Ballasted Pareto Information Matrix

$$A = n \times \begin{bmatrix} \dfrac{\alpha}{\theta^2(2+\alpha)} & -\dfrac{1}{\theta(1+\alpha)} \\ -\dfrac{1}{\theta(1+\alpha)} & \psi'(\alpha) - \psi'(1+\alpha) \end{bmatrix}$$

The next step is to use the M.L.E. estimates in the Information Matrix

Fitting Size of Loss Distributions

Estimating the Uncertainty-Ballasted Pareto example

Using the MLE estimates below

$$\tilde{\theta} = 10{,}000; \quad \tilde{\alpha} = 1.4057$$

Gives the following values in the information matrix

$$A = n \times \begin{bmatrix} \frac{1.4057}{10{,}000^2(3.4057)} & -\frac{1}{10{,}000(2.4057)} \\ -\frac{1}{10{,}000(2.4057)} & \psi'(1.4057) - \psi'(2.4057) \end{bmatrix}$$

You can then use a function like excel's MINVERSE to invert the matrix

$$V = \left[n \times \begin{bmatrix} \frac{1.4057}{10{,}000^2(3.4057)} & -\frac{1}{10{,}000(2.4057)} \\ -\frac{1}{10{,}000(2.4057)} & 0.986859 - 0.480785 \end{bmatrix} \right]^{-1} = \frac{1}{n} \begin{bmatrix} 1{,}402{,}166{,}287 & 115{,}173.7 \\ 115{,}173.7 & 11.43641 \end{bmatrix}$$

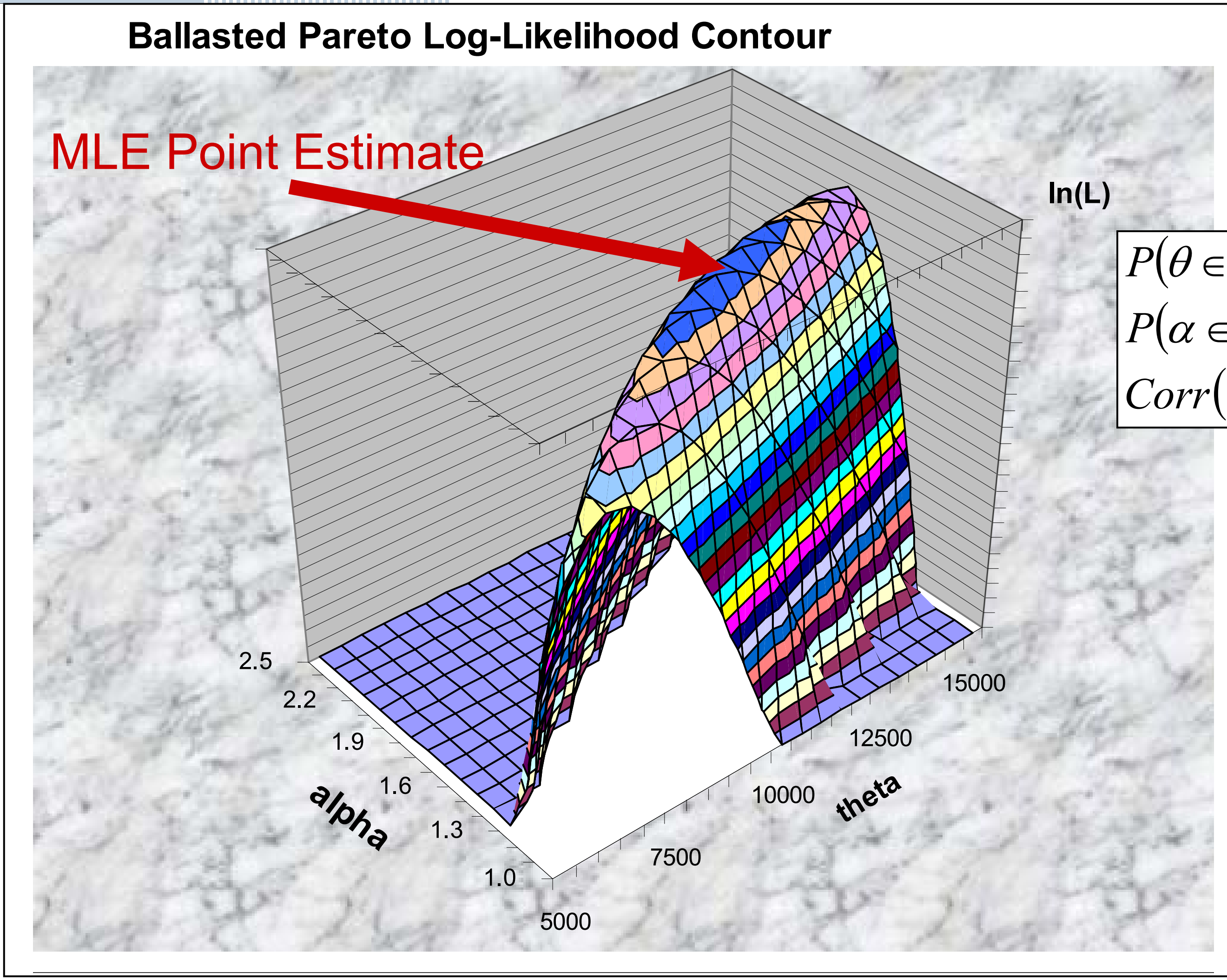
You can now estimate the parameter standard deviations and correlations

$$\sigma_\theta = \sqrt{\frac{1{,}402{,}166{,}287}{n}}; \quad \sigma_\alpha = \sqrt{\frac{11.43641}{n}}; \quad \rho_{\theta,\alpha} = \frac{115{,}173.7}{n \times \sigma_\theta \times \sigma_\alpha}$$

$$n = 1{,}000; \quad \sigma_\theta = 1{,}184.131; \quad \sigma_\alpha = 0.106941; \quad \rho_{\theta,\alpha} = 0.909512914$$

Fitting Size of Loss Distributions

Likelihood Contour for a Ballasted Pareto



Based on the results of the covariance matrix and assuming the parameters are normally distributed, we now assume the following

$$P(\theta \in 10{,}000 \pm 3 \times 1184.131) \approx 99.9\%$$

$$P(\alpha \in 1.4057 \pm 3 \times 0.1069) \approx 99.9\%$$

$$Corr(\theta, \alpha) = 0.9095$$

Fitting Size of Loss Distributions

Using an Extremal Pareto rather than a Ballasted Pareto

There is an alternate parameterization for the Ballasted Pareto called the Extremal Pareto

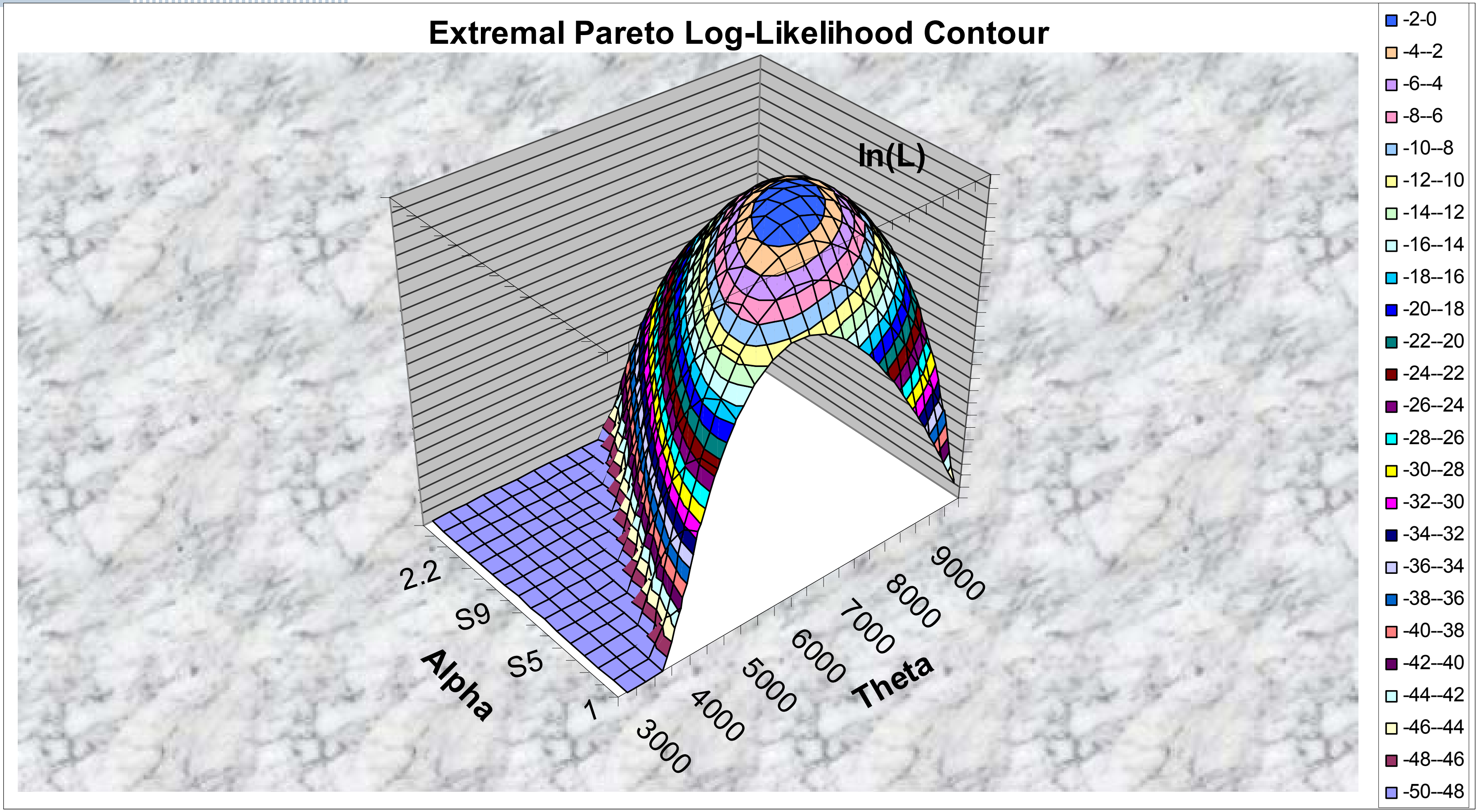
$$F_{EP}\left(x;\theta^*,\alpha\right)=F_{BP}\left(x;\theta^*\alpha,\alpha\right)=1-\left(1+\frac{x}{\theta^*\alpha}\right)^{-\alpha};\ \theta^*_{EP}=\frac{\theta_{BP}}{\alpha}$$

$$\tilde{\theta}^*=7108.2963;\quad \tilde{\alpha}=1.4057$$

While there is still significant volatility around the parameters the correlation between θ^* and α has been significantly reduced as can be seen in the next graph

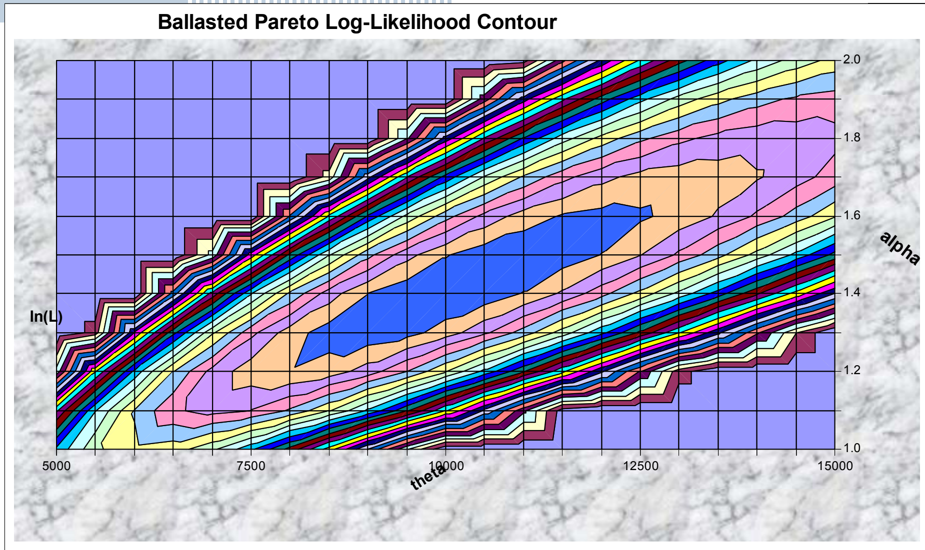
Fitting Size of Loss Distributions

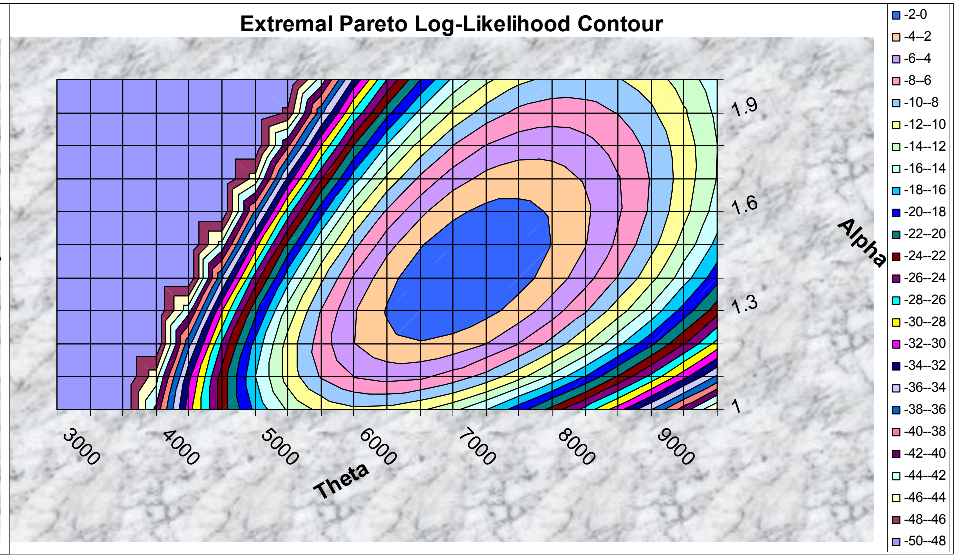
Likelihood Contour for an Extremal Pareto



Fitting Size of Loss Distributions

Comparing Correlations between Ballasted and Extremal Paretos



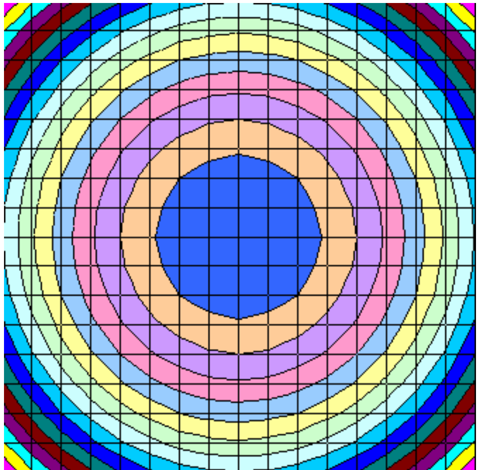


- If you are ignoring parameter uncertainty, then the Extremal Pareto does not improve your modeling (as the point estimate is the same)
- If you model parameter uncertainty including the correlation between parameters, then the Extremal Pareto does not improve your analysis.
- Since few curve fitting packages estimate the parameter correlation, but many calculate t-statistic for parameters, therefore you can estimate the parameter standard deviations.
- If you are modeling the parameter uncertainty (stddev) but not the correlation, then the Extremal Pareto can give better results.

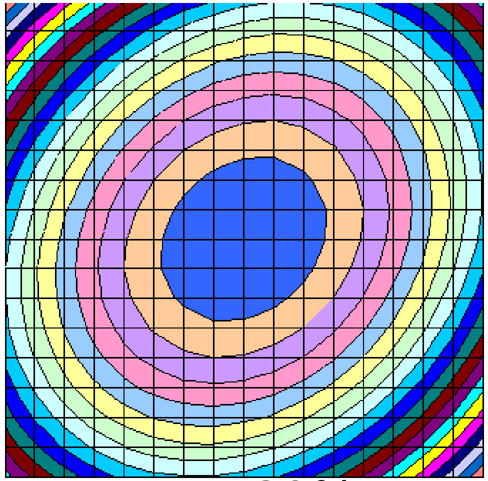
Fitting Size of Loss Distributions

Ballasted Pareto compared to Bi-Variate Normal Distribution

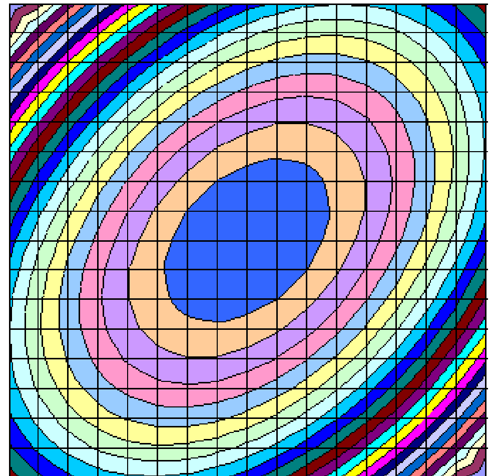
Bivariate Normal Likelihood Countures with varying correlation



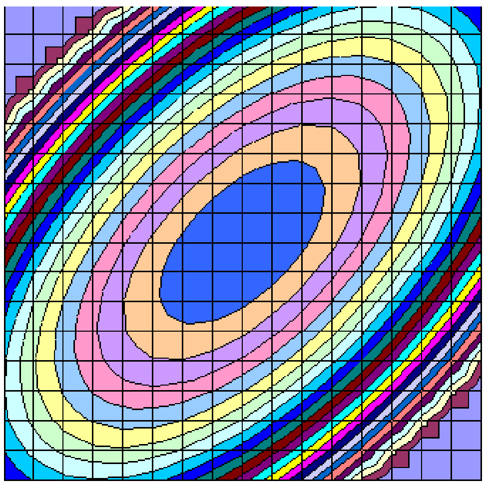
ρ=0%



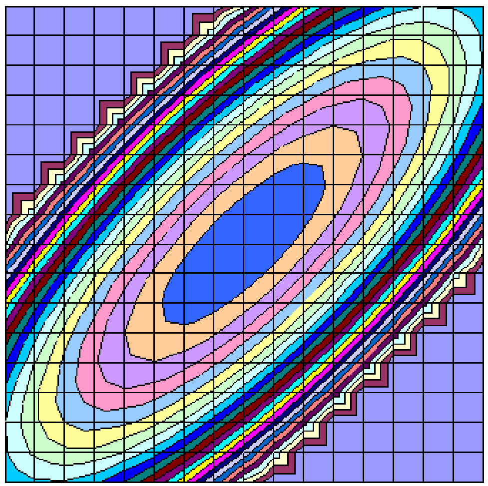
ρ=20%



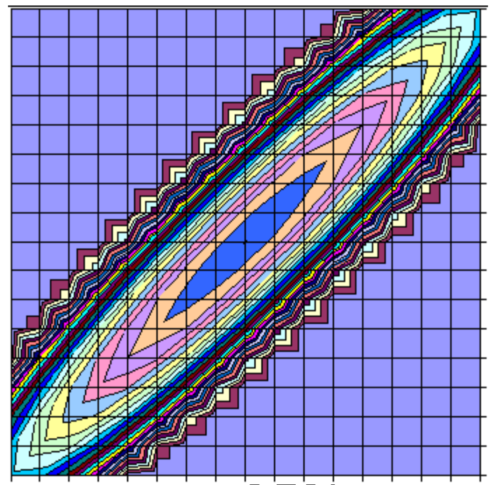
ρ=40%



ρ=60%



ρ=80%



ρ=95%

Section 3

Modeling Multivariate Parameter Uncertainty

Modeling Multivariate Parameter Uncertainty

- First you need to be able to simulate from a multivariate distribution on the parameters
- Since the MLE estimates are asymptotically normally distributed, the multivariate Normal is a natural choice
  - Some believe and have shown that a multivariate Log-Normal is superior particularly when parameters are constrained to be positive
- Any correlating function or copula could be used that you think is appropriate
- The following example is based on the Multivariate Normal

Modeling Multivariate Parameter Uncertainty

Reflecting the parameter uncertainty and parameter correlation with a multivariate normal

First you need to factor the correlation matrix V to solve for the lower diagonal matrix C such that

$$V = CC'$$

Choleski factorization is generally used

$$\begin{bmatrix} 1 & \rho_{\theta,\alpha} \\ \rho_{\theta,\alpha} & 1 \end{bmatrix} = \begin{bmatrix} c_{1,1} & 0 \\ c_{2,1} & c_{2,2} \end{bmatrix} \begin{bmatrix} c_{1,1} & c_{2,1} \\ 0 & c_{2,2} \end{bmatrix}$$

$$1 = c_{1,1}^2; \qquad c_{1,1} = 1$$

$$\rho_{\theta,\alpha} = c_{1,1}c_{2,1}; \qquad c_{2,1} = \rho_{\theta,\alpha}$$

$$1 = c_{2,1}^2 + c_{2,2}^2; \qquad c_{2,2} = \sqrt{1-\rho_{\theta,\alpha}^2}$$

$$C = \begin{bmatrix} 1 & 0 \\ \rho_{\theta,\alpha} & \sqrt{1-\rho_{\theta,\alpha}^2} \end{bmatrix}$$

Klugman, Panjer, Willmot – Loss Models from data to decisions, pp 613

Modeling Multivariate Parameter Uncertainty

Reflecting the parameter uncertainty and parameter correlation with a multivariate normal

Second generate two independent standard normal deviates

$$z_1 \approx Normal(0,1); \;\; z_2 \approx Normal(0,1)$$

If V was the correlation matrix

$$Z^{'} = C \times Z$$

$$\begin{bmatrix} z_1^{'} \\ z_2^{'} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho_{\theta,\alpha} & \sqrt{1-\rho_{\theta,\alpha}^2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

Then new z'_1 and z'_2 are the correlated standard normal deviates

$$z_\theta = \mu_\theta + \sigma_\theta z_1^{'}; \quad z_\alpha = \mu_\alpha + \sigma_\alpha z_1^{'}$$

Then new z_θ and z_α are the correlated parameter estimate

Trivia, if you evaluate the Normal Distribution at the estimated z's

$$u_1^{'} = F(z_1^{'}); \;\; u_1^{'} = F(z_1^{'})$$

Then u'_1 and u'_2 are uniform deviates correlated by a normal copula deviates. Which is one reason the Normal copula is popular

Modeling Multivariate Parameter Uncertainty

Reflecting the parameter uncertainty and parameter correlation with a multivariate normal

Alternatively you can work directly with covariance V to solve for the lower diagonal matrix C such that

$$V = CC'$$

Choleski factorization is generally used

$$\begin{bmatrix} \sigma_\theta^2 & \rho_{\theta,\alpha}\sigma_\theta\sigma_\alpha \\ \rho_{\theta,\alpha}\sigma_\theta\sigma_\alpha & \sigma_\alpha^2 \end{bmatrix} = \begin{bmatrix} c_{1,1} & 0 \\ c_{2,1} & c_{2,2} \end{bmatrix} \begin{bmatrix} c_{1,1} & c_{2,1} \\ 0 & c_{2,2} \end{bmatrix}$$

$$\sigma_\theta^2 = c_{1,1}^2; \qquad c_{1,1} = \sigma_\theta$$

$$\rho_{\theta,\alpha}\sigma_\theta\sigma_\alpha = c_{1,1}c_{2,1}; \qquad c_{2,1} = \rho_{\theta,\alpha}\sigma_\alpha$$

$$\sigma_\alpha^2 = c_{2,1}^2 + c_{2,2}^2; \qquad c_{2,2} = \sigma_\alpha\sqrt{1-\rho_{\theta,\alpha}^2}$$

$$C = \begin{bmatrix} \sigma_\theta & 0 \\ \rho_{\theta,\alpha}\sigma_\alpha & \sigma_\alpha\sqrt{1-\rho_{\theta,\alpha}^2} \end{bmatrix}$$

Klugman, Panjer, Willmot – Loss Models from data to decisions, pp 613

Modeling Multivariate Parameter Uncertainty

Reflecting the parameter uncertainty and parameter correlation with a multivariate normal

Second generate two independent standard normal deviates

$$z_1 \approx Normal(0,1);\ \ z_2 \approx Normal(0,1)$$

If V was the covariance matrix

$$\begin{bmatrix} z_\theta \\ z_\alpha \end{bmatrix} = \begin{bmatrix} \sigma_\theta & 0 \\ \rho_{\theta,\alpha}\sigma_\alpha & \sigma_\alpha\sqrt{1-\rho_{\theta,\alpha}^2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} \mu_\theta \\ \mu_\alpha \end{bmatrix}$$

Then new z_θ and z_α are the correlated normal deviates

$$z_\theta \approx Normal(\mu_\theta, \sigma_\theta);\ \ z_\alpha \approx Normal(\mu_\alpha, \sigma_\alpha)$$

Klugman, Panjer, Willmot – Loss Models from data to decisions, pp 613

Modeling Multivariate Parameter Uncertainty

Reflecting the parameter uncertainty and parameter correlation with a multivariate normal applied to the Ballasted Pareto example

In order to model the uncertainty and correlation of a Ballasted Pareto's parameter You need the MLE estimates of the parameter

$$\tilde{\theta} = 10{,}000; \quad \tilde{\alpha} = 1.4057;$$

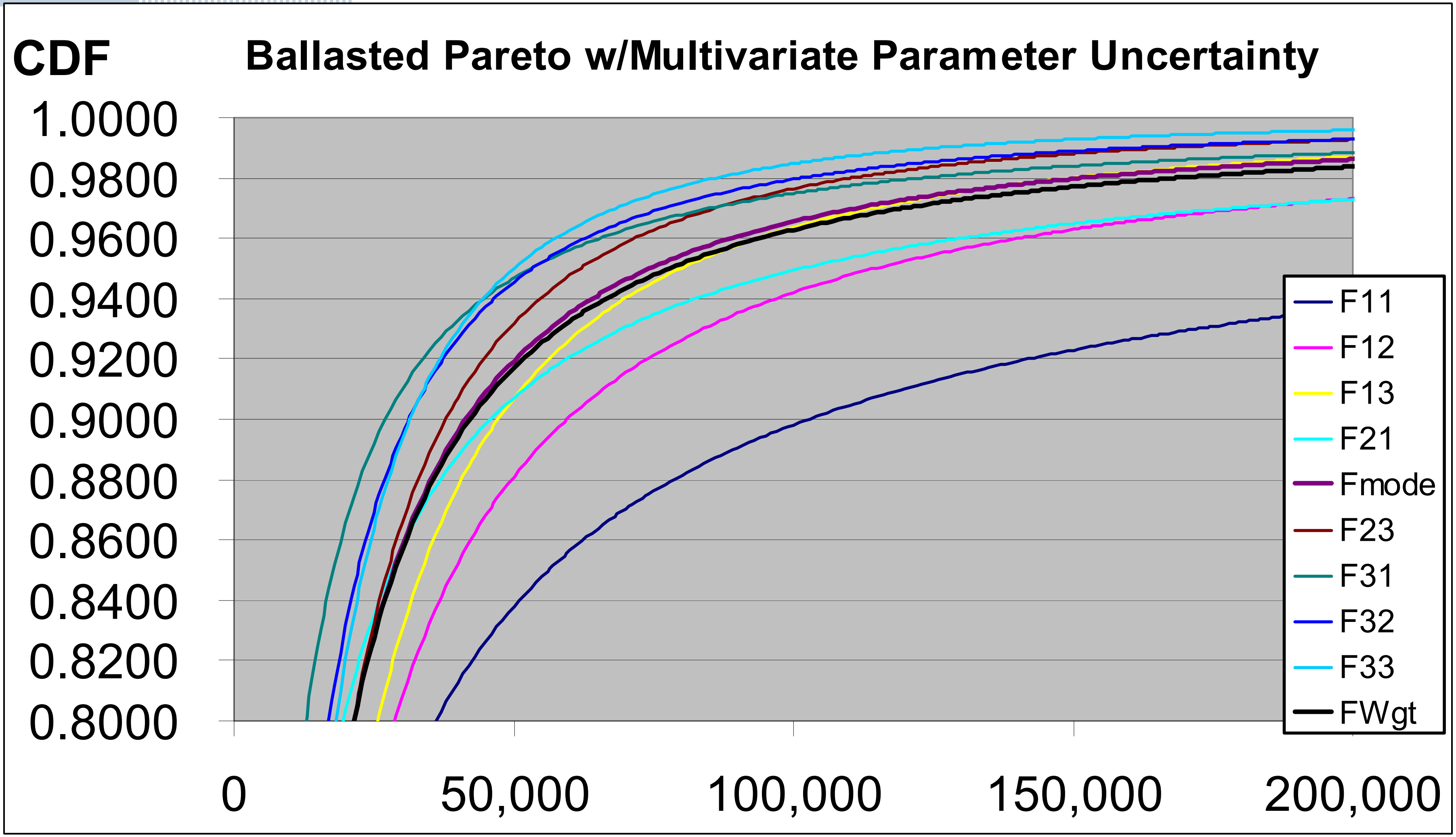
$$\sigma_{\theta} = 1{,}184.131; \quad \sigma_{\alpha} = 0.106941; \quad \rho_{\theta,\alpha} = 0.909512914$$

Next you need to be able to generate some MultiVariate Normals with the above means, standard deviations, and correlation coefficient

Now I will do a similar comparison between the Numerical Integration mixing methods and Simulation

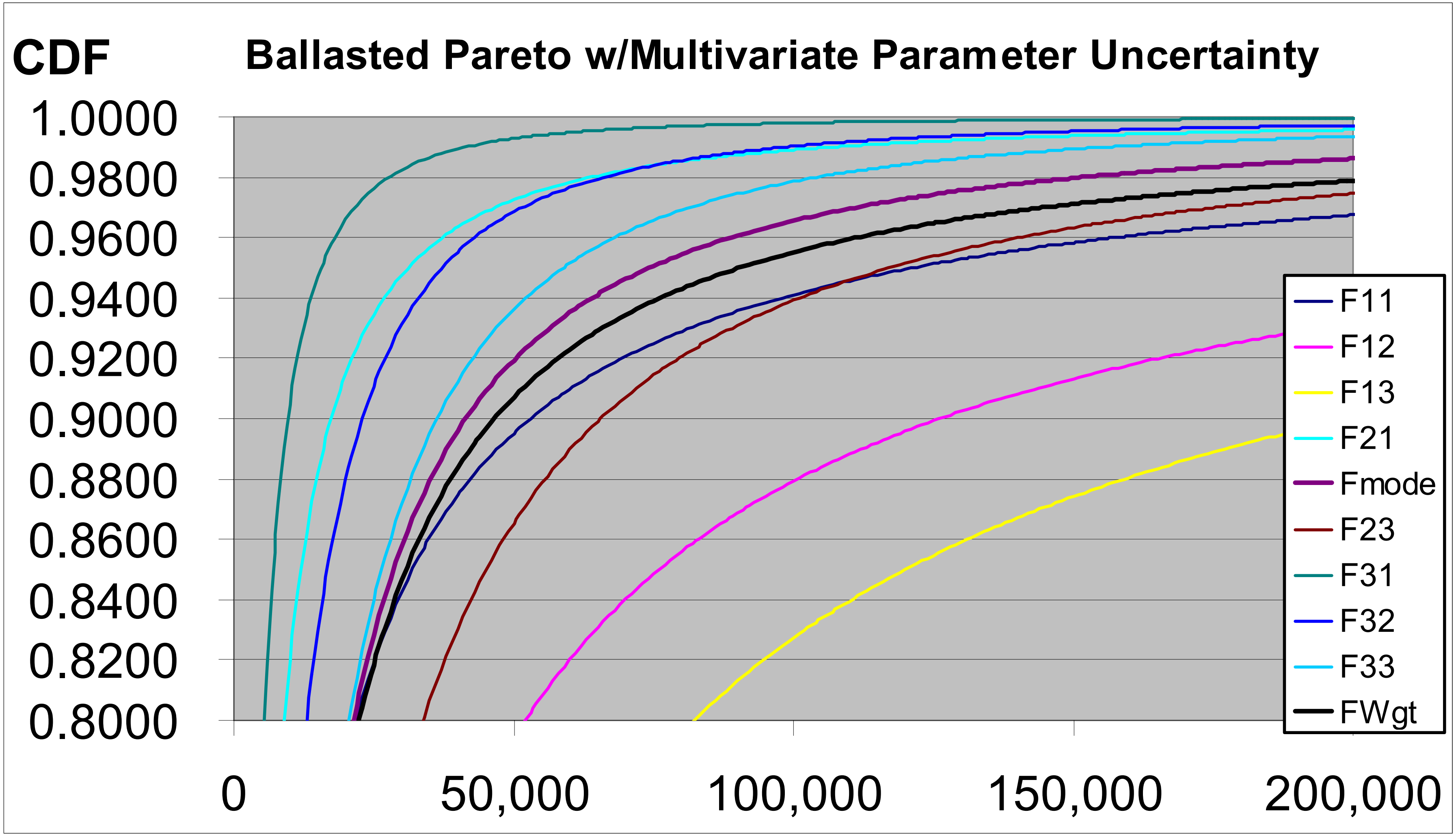
Modeling Multivariate Parameter Uncertainty

Multivariate 3pt Gaussian Integration Including Correlation



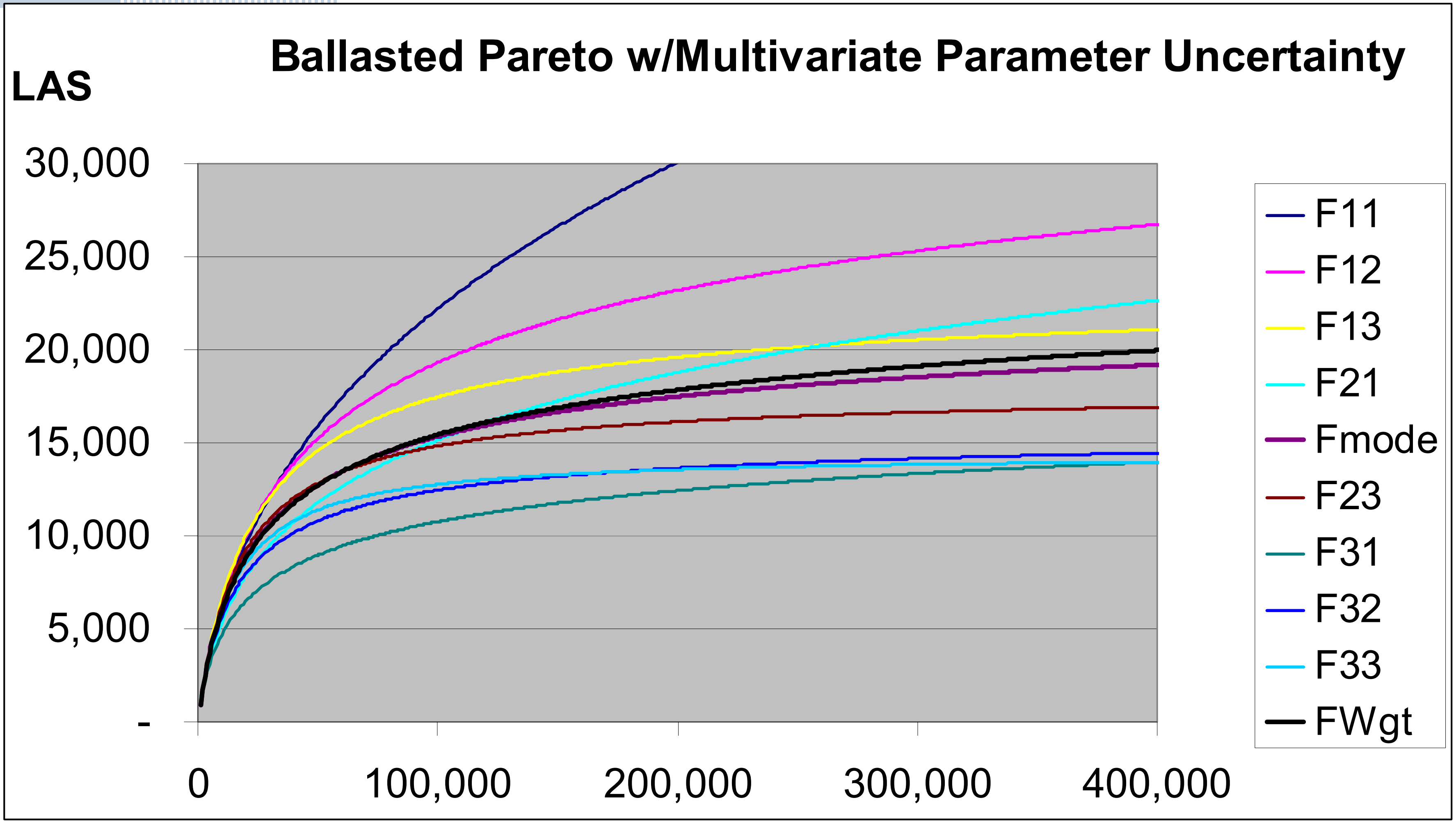
Modeling Multivariate Parameter Uncertainty

Multivariate 3pt Gaussian Integration Excluding Correlation



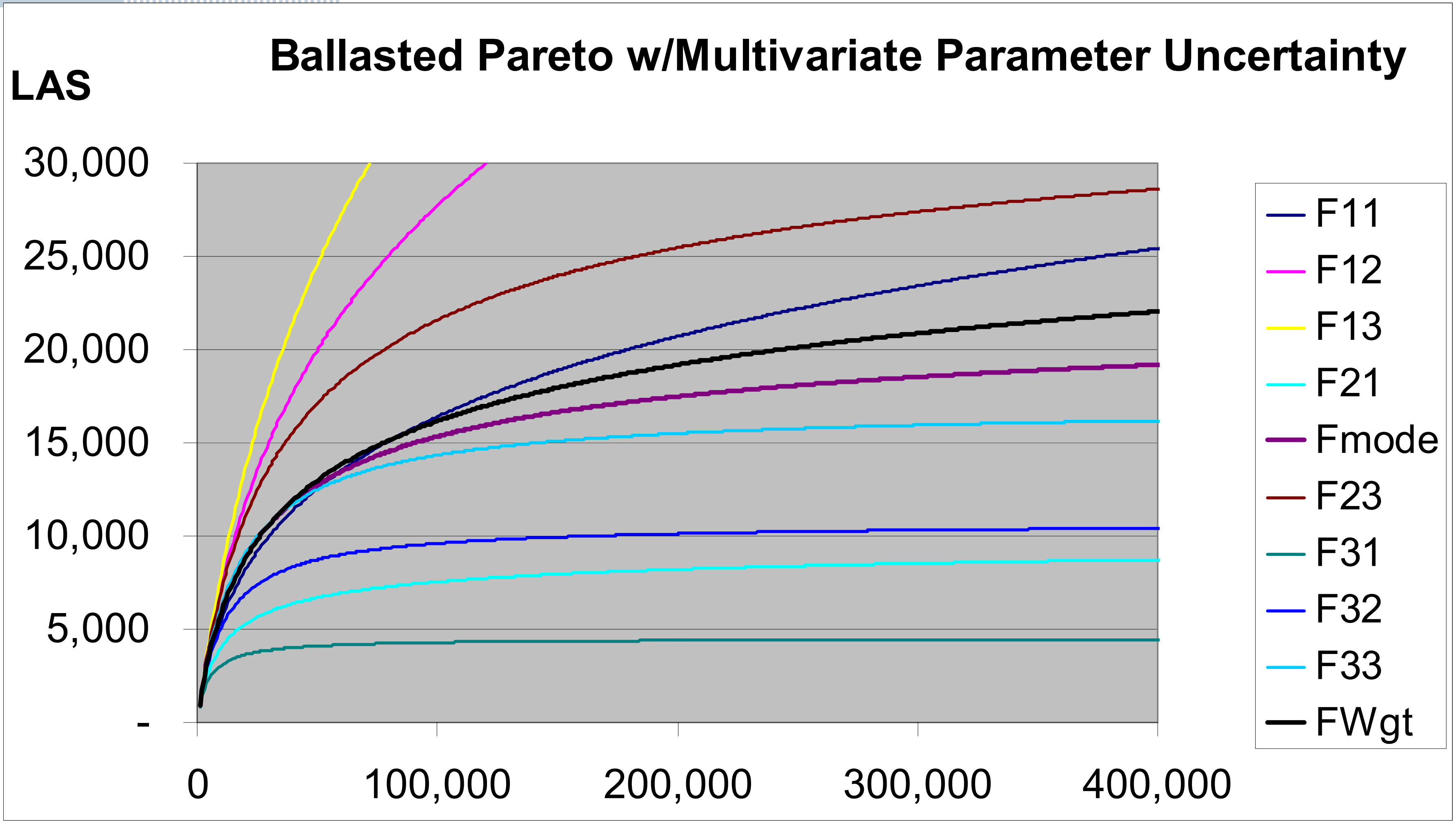
Modeling Multivariate Parameter Uncertainty

Multivariate 3pt Gaussian Integration Including Correlation



Modeling Multivariate Parameter Uncertainty

Multivariate 3pt Gaussian Integration Excluding Correlation



Uncertainty in Reinsurance Pricing

Summary

- The required tools to include the effects of parameter uncertainty are available in common tools like spreadsheets
- The most difficult step is to estimate the information matrix and that can be approximated by numerical differentiation
- Matrix Functions like Inverse, Multiplication, Transpose are built into most spreadsheets and Choleski Factorization is not difficult to solve
- Beyond parameter uncertainty this also gives all the tools needed for a normal copula
- Uncertainty can be built into many more of the actuarial models than is commonly done
- Much progress has been made on DFA modeling, it should expand into other areas