STATE AUTO®
Insurance Companies

Mixture Distribution and Its Applications on P&C Insurance Data

Luyang Fu, Ph.D., FCAS, MAAA
Doug Pirtle, FCAS

May 2011

Antitrust Notice

- The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.

- Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.

- It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.

Agenda

- Introduction
- Mixture Distribution
- Finite Mixture Model
- Case Study
- Conclusions
- Q&A

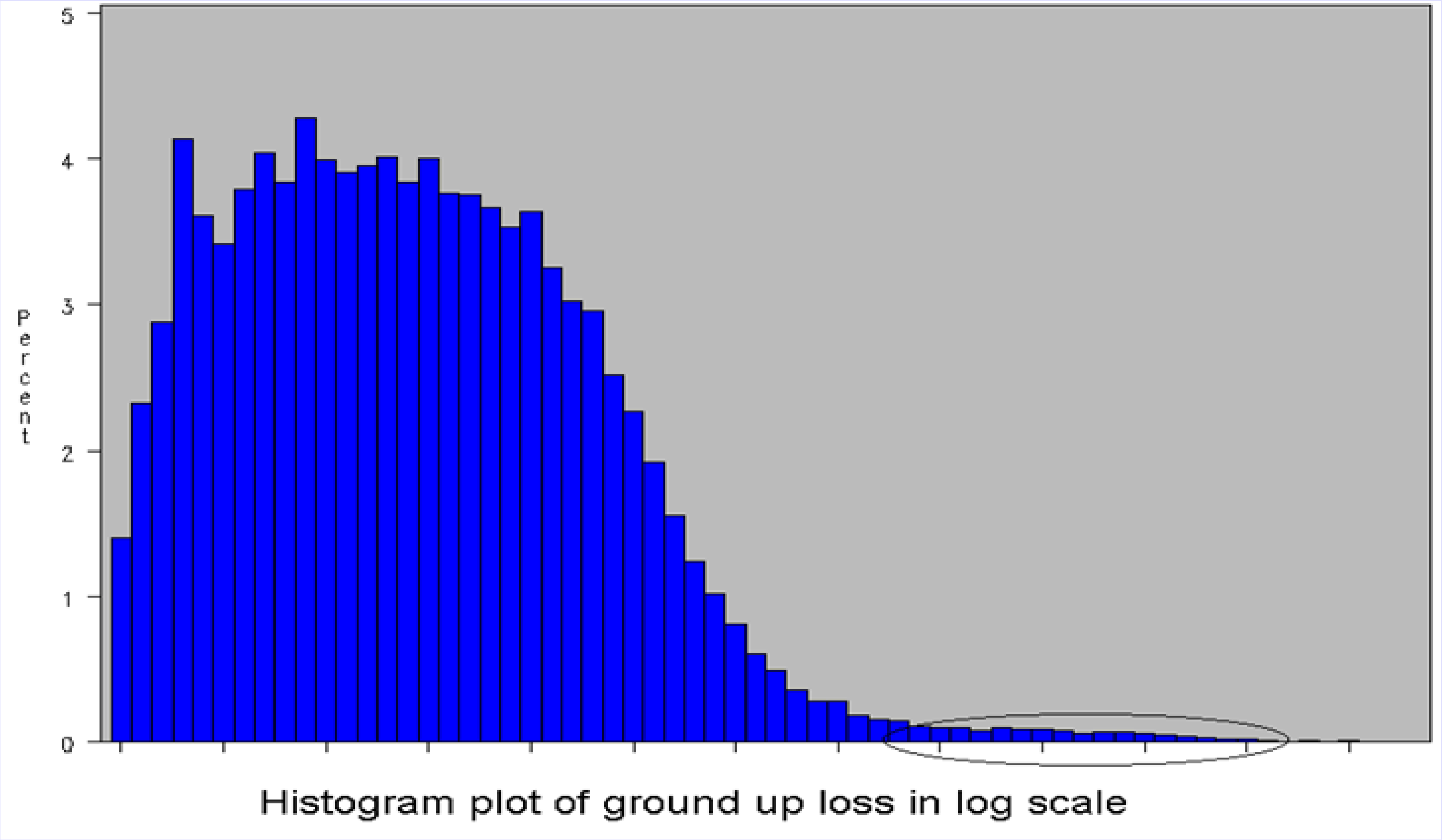
Introduction

Skewed Insurance Data

- Skewed and asymmetric
- Heavy tails
- Mixed: typical and extreme
  - Investment return: normal and crisis
  - Claim amount: typical and large losses

Introduction

HO by-peril example: heavier tail than lognormal

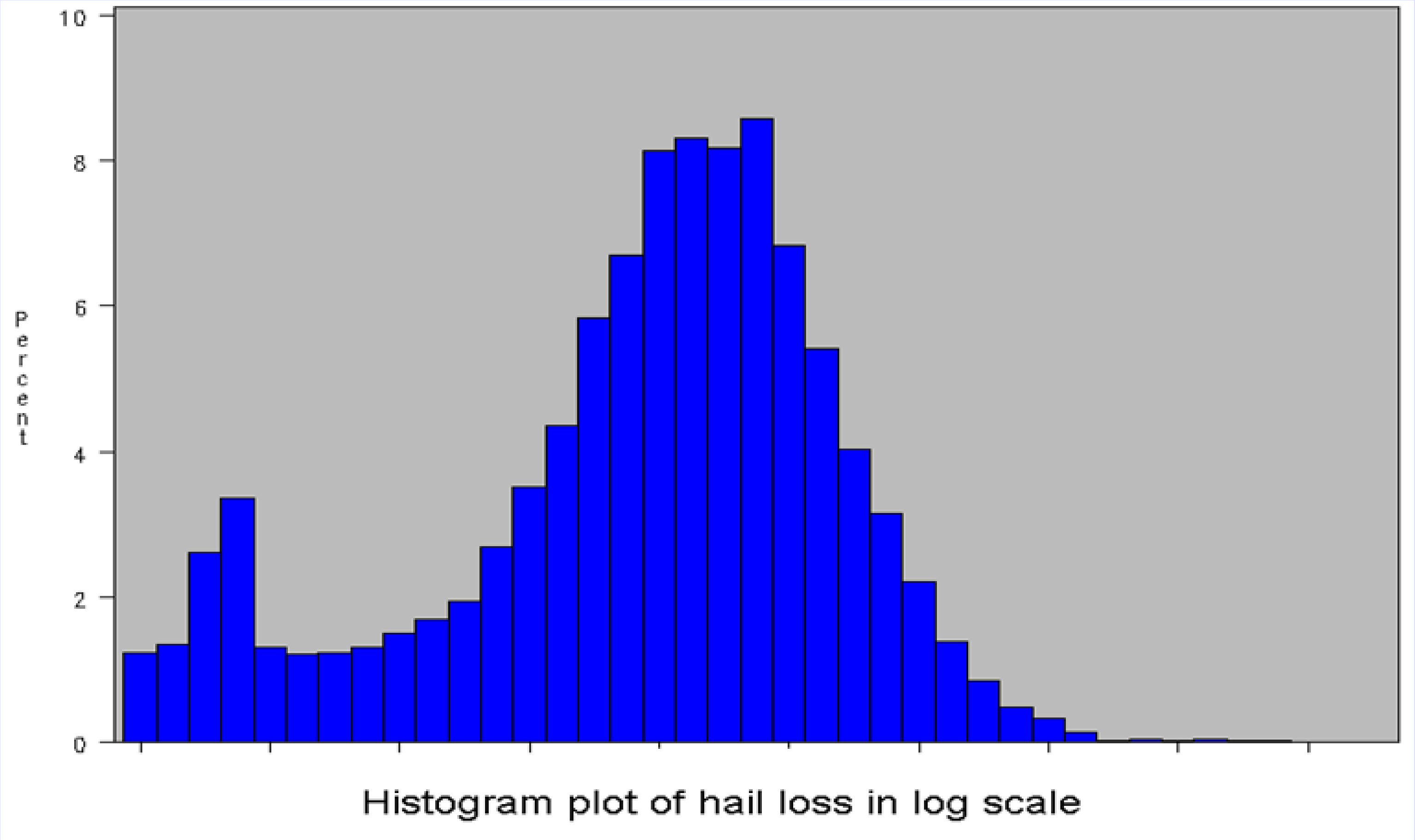


Introduction

HO by-peril example: multiple peaks

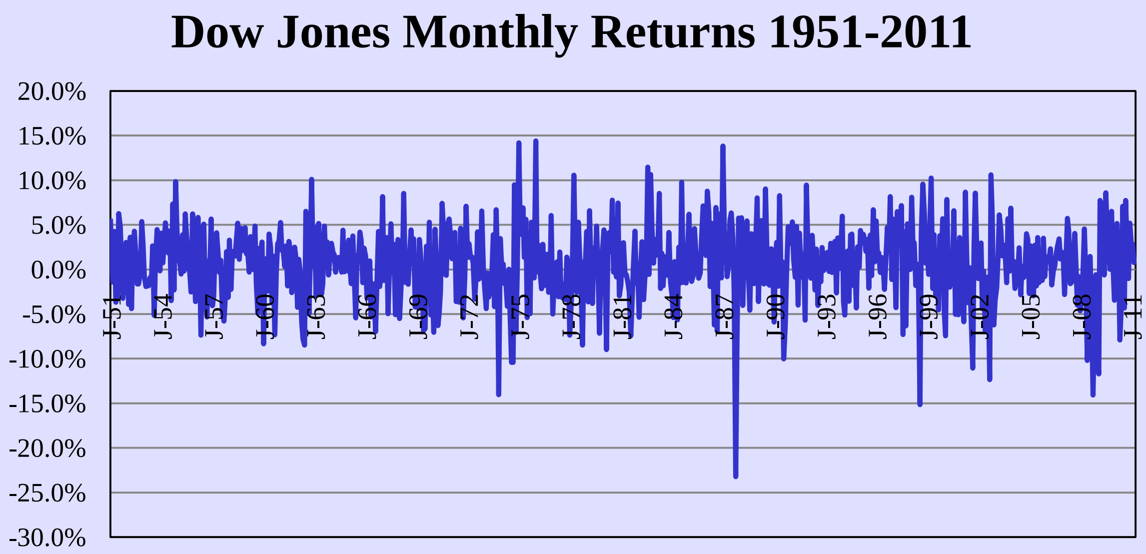
Introduction

HO by-peril example: multiple peaks



Introduction

Investment example in DFA



- Assuming normal distribution, the likelihood of monthly loss over 14.1% (largest monthly drop in Deep Recession) is 0.02%; actual observation is 0.55%.

Mixture Distribution

- Single distribution does not fit insurance data well
- A combination of multiple distributions can represent data better
- Mixture distributions:

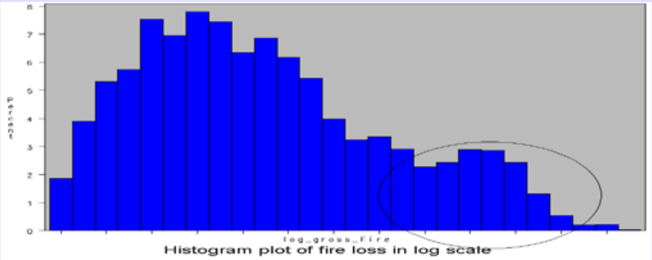
$$f(x, \pi_1, \pi_2, \ldots \pi_n, \beta_1, \beta_2, \ldots \beta_n) = \sum_i^n \pi_i \cdot f_i(x, \beta_i)$$

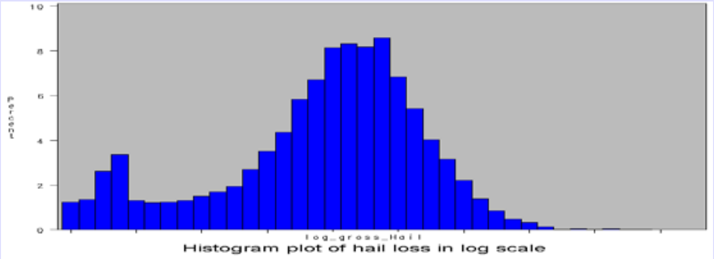
$$where \quad \sum_i^n \pi_i = 1$$

Mixture Distribution

Typical mixture distributions in insurance

- Claims count: Zero + Poisson
- Claim amount: gamma + lognormal or gamma + Pareto





Peril	π	α	β	μ	σ
Fire	0.785	0.51	10500	11.5	0.83
Hail	0.148	1.19	520	8.8	0.61

Mixture Distribution

- Regime-Switching Models of Equity Returns;
- Two lognormal distributions with low and high volatilities;
- Two regimes may switch by a matrix of transition probabilities;
- Hamilton (1990), Hardy (2001), Ahlgrim, D'Arcy, and Gorvett (2004).

	Low Volatility	High Volatility
Mean	0.96%	-2.20%
Standard Deviation	3.59%	7.17%
Probability of Switching	3.37%	30.87%

The likelihood of penetrating -14.1% by regime-switching model is 0.41%.

Finite Mixture Model

$$f(y \mid X; \pi_1, \pi_2, \ldots \pi_n, \theta_1, \theta_2, \ldots \theta_n) = \sum_i^n \pi_i \cdot f_i(X, \theta_i)$$

$$where \quad \sum_i^n \pi_i = 1$$

- y: response variable; X: explanatory variables
- A finite mixture model can be thought as a mixture of multiple GLMs
  - $f_1(y \mid X; \theta_1)$ is a GLM for smaller fire loss assuming gamma
  - $f_2(y \mid X; \theta_2)$ is a GLM for large fire loss assuming lognormal
- Often named as latent class model in economics

Finite Mixture Model

- Improvements on GLM
  - Expand distribution assumptions:
    Single exponential-family distribution vs. mixture
  - Expand model structure:
    Single regression formula vs. multiple models
  - Better fits on insurance data with heavy-tails, multimodal , excessive zeros, and other complex error distributions

AOI Group	**5%** Deductible Factors for Hail	
	GLM gamma	FMM
2	0.359	0.419
18	0.187	0.348

Finite Mixture Model

Numerical Solution

- Solving maximum likelihood function

$$\underset{\pi,\theta}{Max}\sum_{j=1}^{N}\log(\sum_{i=1}^{n}\pi_i f_i(y_j \mid X_j;\theta_i))$$

with constraint $\pi_i > 0 \quad and \quad \sum_{i}^{n}\pi_i = 1$

- EM (Expectation-Maximization) Algorithm
- Quasi-Newton Method
- Bayesian MCMC

Case Study: Data Description

- Simulated Hurricane Model Output
- 8,500 of 10,000 years with hurricane losses.
- Mean Aggregate Severity = $57,000,000
- Standard Deviation = $136,000,000
- Skewness = 6.5

- Positive skewness suggests an asymmetric distribution
  - Lognormal
  - Gamma

Case Study: Simple Distributions Fit Poorly

- Lognormal: Determine Parameters
  - Maximum Likelihood Estimation (MLE)
  - Method of Moments (MOM)
- Intuitive Test: MLE and MOM parameter estimates differ implying Lognormal is not a good fit.
- Chi-Square Test:
  - Critical Value at 95% = 11.1
  - Test Statistic Value = 419.0
  - Since 419.0>11.1 we reject the null hypothesis that the data were drawn from a Lognormal distribution with the fitted parameters.

Case Study: Simple Distributions Fit Poorly

Lognormal MLE

- Mean of log(loss) is 16.03 and Standard deviation is 2.50
  - Implied Mean = $ 207,000,000
  - Implied Stdev = $4,681,000,000
  - Max observed value = $3,053,000,000
- Excess small losses (81 losses <=$3000) make the error from model misspecification extreme.
  - Lognormal assumes log(loss) are symmetric
  - Log($3000)=8.01. The symmetric point on the other side of mean is 24.05, or $27,800,000,000
  - The losses are positively skewed with a heavy right tail; log(loss) is negatively skewed with heavy left tail. Lognormal cannot address this specific shape of distribution.

Case Study: Simple Distributions Fit Poorly

- Gamma: Determine Parameters
  - MLE fit
  - MOM fit
- Intuitive Test: MLE and MOM parameter estimates differ implying Gamma is not a good fit.
- Chi-Square Test:
  - Critical Value at 95% = 11.1
  - Test Statistic Value = 683.3
  - Since 683.3>11.1 we reject the null hypothesis that the data were drawn from a Gamma distribution with the fitted parameters.

Case Study: Mixed Distributions Fit Better

- Mixed Gamma-Lognormal: Determine Parameters
  - Density:

$$f(x,\alpha_1,\beta_1,\pi_1,\mu_2,\sigma_2) = \pi_1 * f_1(x,\alpha_1,\beta_1) + (1-\pi_1) * f_2(x,\mu_2,\sigma_2)$$

  - Likelihood:

$$L(\alpha_1,\beta_1,\pi_1,\mu_2,\sigma_2) = \prod_{i=1}^{8500} f(x_i,\alpha_1,\beta_1,\pi_1,\mu_2,\sigma_2)$$

  - Log-Likelihood:

$$l(\alpha_1,\beta_1,\pi_1,\mu_2,\sigma_2) = \sum_{i=1}^{8500} \ln(f(x_i,\alpha_1,\beta_1,\pi_1,\mu_2,\sigma_2))$$

Case Study: Mixed Distributions Fit Better

- Mixed Gamma-Lognormal: MLE Parameters

$$\alpha_1 = .446, \beta_1 = 57.9M$$
$$\pi_1 = 0.884$$
$$\mu_2 = 19.221, \sigma_2 = 0.789$$

- Intuition: Aggregate Severity is drawn from:
  - 88.4% of time Gamma (Mean=26M, Stdev=39M)
  - 11.6% of time Lognormal (Mean=304M, Stdev=282M)

- Match to 1st two moments:
  - Mean of mixture matches data within 0.2%.
  - Standard deviation of mixture matches data within -0.7%.

Case Study: Mixed Distributions Fit Better

- Mixed Gamma-Lognormal: Significance?

- Likelihood Ratio Test 95% Critical Value=7.8
  - Mixed vs. Gamma Test Statistic = 668
  - Mixed vs. Lognormal Test Statistic = 1331

- Since test statistics > critical value the mixed distribution provides a significantly better fit to the data than either of the simple distributions.

Case Study: Fitting Mixtures

- Tools Available to Fit Mixed Distributions
  - Microsoft Excel SOLVER
  - R
  - SAS
  - Other
- Steps to Fit Mixed Distributions
  - Write the Mixed Density Function
  - Specify Initial Parameter Values
  - Write the Log-Likelihood Function
  - Maximize the Log-Likelihood by Changing Parameters

Case Study: Fitting Mixtures

- Mixed Gamma-Gamma:
  - Density:

$$f(x,\alpha_1,\beta_1,\pi_1,\alpha_2,\beta_2) = \pi_1 * f_1(x,\alpha_1,\beta_1) + (1-\pi_1) * f_2(x,\alpha_2,\beta_2)$$

  - Specify Initial Parameter Values
  - Likelihood:

$$L(\alpha_1,\beta_1,\pi_1,\alpha_2,\beta_2) = \prod_{i=1}^{8500} f(x_i,\alpha_1,\beta_1,\pi_1,\alpha_2,\beta_2)$$

  - Log-Likelihood:

$$l(\alpha_1,\beta_1,\pi_1,\alpha_2,\beta_2) = \sum_{i=1}^{8500} \ln(f(x_i,\alpha_1,\beta_1,\pi_1,\alpha_2,\beta_2))$$

Case Study: Fitting Mixtures

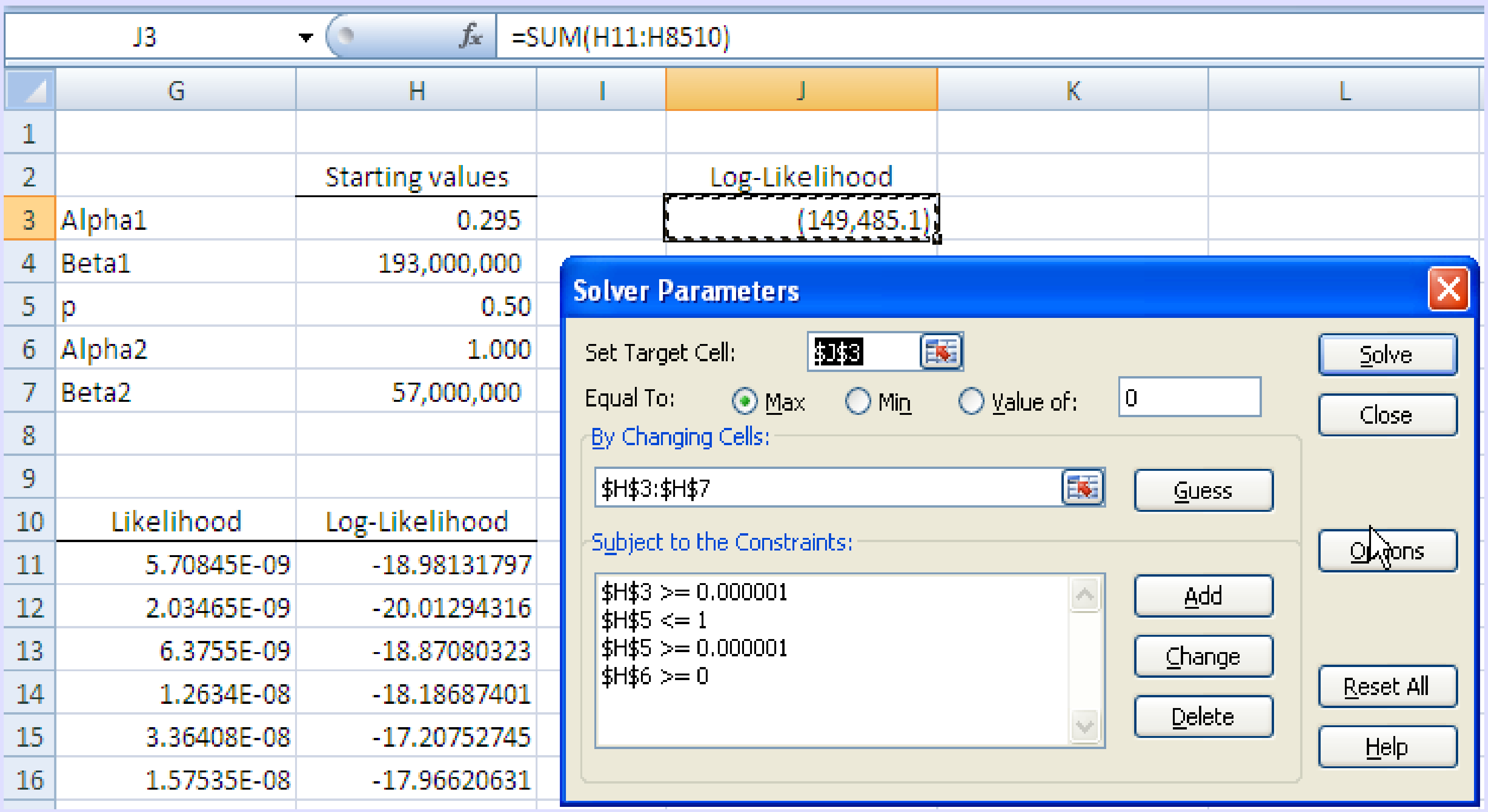
- Maximize Log-Likelihood: Excel SOLVER

=$H$5*GAMMADIST(E11,$H$3,$H$4,FALSE)+(1-$H$5)*GAMMADIST(E11,$H$6,$H$7,FALSE)

D	E	F	G	H	I	J	K
				Starting values		Log-Likelihood	
			Alpha1	0.295		(150,165.8)	
			Beta1	193,000,000			
			p	0.50			
			Alpha2	1.000			
			Beta2	57,000,000			
Year	AggLoss		Likelihood	Log-Likelihood			
1	46,452,953		=$H$5*GAMMADIST(E11,$H$3,$H$4,FALSE)+(1-$H$5)*GAMMADIST(E11,$H$6,$H$7,FALSE)				
2	108,518,889		2.03465E-09	-20.01294316			

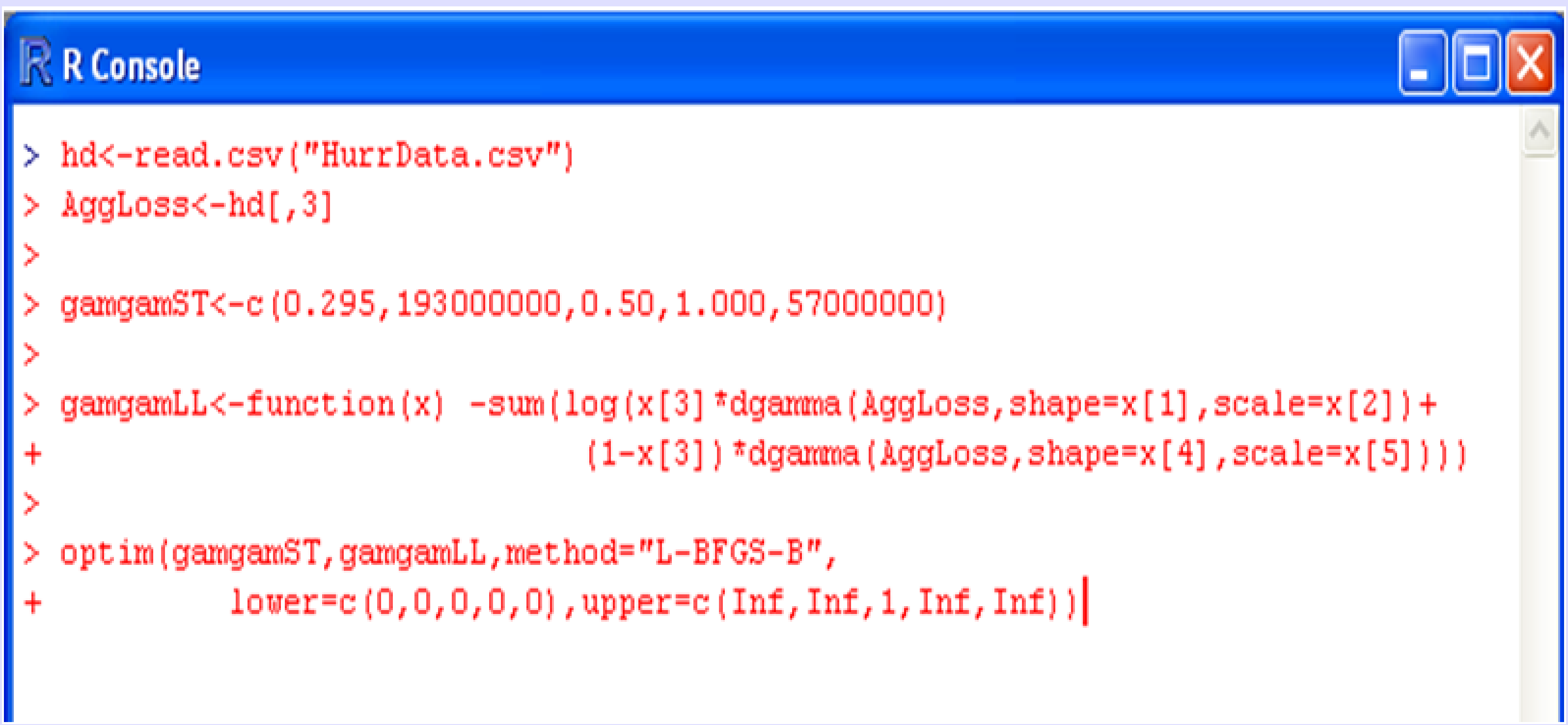
Case Study: Fitting Mixtures

- Maximize Log-Likelihood: Excel SOLVER



Case Study: Fitting Mixtures

- Maximize Log-Likelihood: R
  - http://www.r-project.org/



R Console

```
> hd<-read.csv("HurrData.csv")
> AggLoss<-hd[,3]
>
> gamgamST<-c(0.295,193000000,0.50,1.000,57000000)
>
> gamgamLL<-function(x) -sum(log(x[3]*dgamma(AggLoss,shape=x[1],scale=x[2])+
+                             (1-x[3])*dgamma(AggLoss,shape=x[4],scale=x[5])))
>
> optim(gamgamST,gamgamLL,method="L-BFGS-B",
+          lower=c(0,0,0,0,0),upper=c(Inf,Inf,1,Inf,Inf))
```

Case Study: Fitting Mixtures

- Parameter Risk: Sample Data
  - The second distribution could have low credibility.
  - Sensitivity test with slight data changes
  - Parameter uncertainties in cat modeling firms (AIR, RMS, EQECAT)

- Parameter Risk: Initial Values
  - Could lead to local maxima
  - Try different starting values
    - Start with 90%/10% weights
    - Use same distribution to infer starting means such as a mixture of 2-Gamma distributions.

Case Study: Fitting Mixtures

- Parameter Risk: Robustness
  - Remove 81 losses less than $3000, and refit MLE lognormal and gamma-lognormal distributions.
  - For lognormal, the fitted mean <u>decreased</u> by 29%; the fitted standard deviation <u>decreased</u> by 54%.

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\ln(x_i) \qquad E[X] = e^{\mu+\sigma^2/2}$$

  - For gamma-lognormal, the fitted mean <u>increased</u> 2%, the fitted standard deviation <u>decreased</u> by 0.1%.
  - Mixture distribution is more robust!

Case Study: Implications

- Expected Reinsurance Recovery
  - Low credibility for high layers
    - Hurricane output only contained 56 losses over $800M.
    - Only 5 losses over $1.6B.
  - Fitted distribution can help evaluate cost for higher layers

- Alternative Tail Estimates
  - Percentiles/VaR
  - TVaR

Conclusions

- Insurance data are skewed and heavy tailed.
- Single distribution in general cannot fit data well.
- Mixture distribution can represent insurance data with excess zeroes, multiple modes, and heavy tails.
- Finite mixture model improves GLM by assuming mixture distribution.
- Many insurance applications: ERM (PML, TVaR), asset management, reinsurance (cat, per risk), high deductible (worker comp, property), predictive modeling (frequency, severity).

Q & A

STATE AUTO®
Insurance Companies