GUY CARPENTER

PARAMETER RISK REVISITED
Variance Papers: Parameter Risk

NOVEMBER 16, 2016

Avraham Adler, FCAS, CERA, MAAA
Senior Vice President
GC Analytics®

Morristown, NJ

MARSH & McLENNAN COMPANIES

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.**

Parameter Risk Revisited
Agenda

- What is parameter risk?
- Van Kampen's bootstrap approach
- Maximum Likelihood
- Hierarchical Bayesian
- Comparison

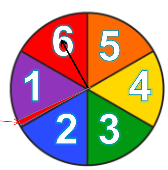
 October 10, 2016 2

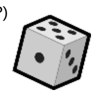
Section 1

Parameter Risk Revisited
What is parameter risk?

What is parameter risk?
Overview of various risk types

- Process Risk
  - Spinner with six sections, which one is chosen?
  - Expected value – 3.5
- Parameter Risk
  - Finite sample size
    - Current snapshot is not accurate
    - 0.1% chance of 10,000
    - Mean actually 5.1665
  - Changing parameters over time
    - Current snapshot is accurate
    - Number 6 wedge is expanding over time (ink spreading?)
- Model Risk
  - Not a spinner but a six-sided die





What is parameter risk?
Problem statement

- Are the observed results:
  - Reasonable results from probable parameters?
  - Outlier results from improbable parameters?

Section 2

Parameter Risk Revisited
Van Kampen's bootstrap approach

Van Kampen's bootstrap approach
Reinsurance pricing problem

- Reinsurance contract is a 2.5% excess 75% LR stop loss
- There are ten years of loss ratio observations, of which one attached the cover
- What is the expected LR in the stop loss?
- If any of the ten observed years were outliers, true expected results of stop loss can be very different

Year	GULR	Ln(GULR)	Agg Stop LR
1	58.4%	-0.5376	0.0%
2	64.5%	-0.4388	0.0%
3	67.4%	-0.3953	0.0%
4	52.6%	-0.6415	0.0%
5	58.4%	-0.5376	0.0%
6	64.5%	-0.4388	0.0%
7	78.4%	-0.2440	2.5%
8	70.6%	-0.3488	0.0%
9	62.0%	-0.4786	0.0%
10	64.5%	-0.4388	0.0%
Emp. Mean	64.1%	-0.4500 (μ)	0.25%
Emp. StDev	7.12%	0.1100 (σ)	0.79%
Emp. Skew	0.5003		
Fitted LR	64.2%		0.235%
Emp. LoL			10.0%
Fitted LoL			9.4%

Van Kampen's bootstrap approach
Bootstrap procedure

- Calculate summary statistics of observed data
  - Mean, Standard Deviation, and Skew
- Generate sets of parameters from uniform grid
  - Van Kampen assumed a lognormal distribution
- For each μ and σ pair generate 10,000 sets of 10 years of data
- For each 10 year block calculate summary statistics
  - If all three statistics are "close" to empirical values, deem set "viable"
- Weight each "viable" parameter set by number of its observations which are "close"
- Calculate expected results by weighting outcome of each parameter set by its viability
  - Weighted average now takes into account many parameter sets!

Van Kampen's bootstrap approach
Updated process and results

- In 2003, using MS Excel, process took over 8 hours to generate 10,000 sets of 10-year blocks for each of 3,950 pairs of μ and σ
- In 2014, using R and Rcpp, process for 4,029 parameter sets took under two minutes!
- Fine-grain grid of 38,081 parameter sets took 17 minutes!

Statistic	Original	Coarse Mesh	Fine Mesh
Ground-up LR	64.15%	64.36% · (+0.3%)	64.4% · (+0.4%)
Agg Stop LR	0.235%	0.318% · (+35.5%)	0.336% · (+43.2%)

- Parameter risk makes a very big difference for the stop loss!

Section 3

Parameter Risk Revisited
Maximum Likelihood

Maximum Likelihood
Properties of MLE

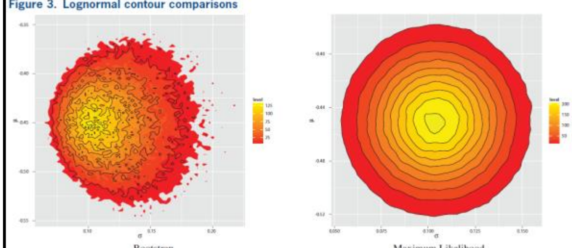
- One property of MLE is its *asymptotic normality*
  - Under general conditions, as sample size increases, the distribution of the estimators tends to multivariate normal
- Can use Hessian at point of convergence to estimate SD of—and correlations between—the parameters
- Given parameter estimates, can use multivariate normal to generate pairs of correlated values and use those to estimate loss
- Do this a gazillion times (Monte Carlo simulation) and the empirical average result is an estimate containing parameter risk

Maximum Likelihood
Results under lognormal assumption

- Lognormal estimate shows no correlation between parameters
  - Known property of normal distribution: mean and sample variance are independent
- Monte Carlo simulation showed effectively zero change in both ground-up and agg stop loss
- Generated lognormal parameters are relatively evenly and equally distributed
  - Estimated loss ratio very different from bootstrap results

Maximum Likelihood
Different parameter distribution between bootstrap and MLE

Figure 3. Lognormal contour comparisons



- If lognormal parameters by nature are independent, why are results different?

Maximum Likelihood
Results under other distributional assumptions

- Other distributions investigated include gamma, Weibull, Burr, and Inverse Burr.
- Many “untempered” observations were not mathematically impossible yet patently ridiculous
  - Loss ratio of ten quattuorsexagintillion percent (10^{196}%) anyone?
  - That's1000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000%!
- Actuarial judgement applied:
  - Results trimmed by 0.1% (top and bottom 500 observations removed)
  - Illegal parameter sets removed (e.g. Burr parameters < 0)
  - LR cap of 300% implemented (no more ten quattuorsexagintillion ☹)

Maximum Likelihood
Model comparison: technique

- Models compared using Akaike's Information Criterion with small-sample bias correction (AICc)
  - Can be used to compare different models built on the **exact same** data
  - The magnitude of the value is irrelevant—it is the **difference** between the values which is important
- Rule of thumb for differences from minimum ("best model"):
  - 0 – 2: Substantial support for second model
  - 4 – 7: Less support for second model
  - 10+: Essentially no support for second model

Maximum Likelihood
Model comparison: results

Model	AICc	ΔAICc	Adjusted GULR	Adjusted Agg Stop LR
Lognormal	-20.1072	0.0000	64.15%	0.23%
Gamma	-20.0431	0.0641	64.30%	0.27%
Weibull	-18.2925	1.8147	63.73%	0.27%
Inverse Burr	-15.7649	4.3423	106852.69%	0.37%
Burr	-15.7579	4.3493	304090355.01%	0.47%
IB – "Valid"			63.56%	0.35%
Burr – "Valid"			66.85%	0.44%

- Note that for Burr & Inverse Burr, AICc is not "so" bad, but results are completely unreasonable
- Even if adjusted to be "valid", results remain extremely unlikely
- Agg Stop protected by being limited.
  - **Important: When simulating with parameter risk, limit your range!**

Maximum Likelihood
Estimation on log scale

- Solving for log of parameters prevents any from being negative in normal space
- For this data set, results of MLE in log-space are disappointing
  - Gamma now shows almost no change
- Time to try something new

Section 3

Parameter Risk Revisited
Bayesian Analysis

Bayesian Analysis
Properties of simple Bayesian model

- Explicit formulation of *a priori* distribution of the parameters
  - Parameters are not fixed—only the data
  - Weakly informative prior on parameters allows for exploration of parameter space without random walk being forced to go to where it doesn't want and prevented from going to where it wants!
- Modern programs make coding and running much easier
  - JAGS, Stan
  - Both have packages that allow them to be called directly from R
- Same distributional families used in MLE analysis tested

Bayesian Analysis
Bayesian model comparison criteria

- Various information-criterion for comparing models
  - DIC (Speighalter et al.)
    - Not fully Bayesian as relies on drop-in estimates
    - used in other Variance paper (written mainly in 2012)
  - WAIC (Wantanabe)
    - More fully Bayesian
    - used in this paper (written mainly in 2014)
  - PSIS-LOO (Vehtari et al.)
    - More fully Bayesian
    - Less subject to asymptotic bias than WAIC
    - will probably use going forward

Bayesian Analysis
Simple fit comparisons

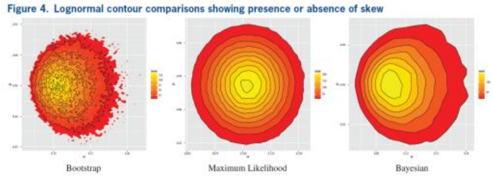
- Same distributional families used in MLE analysis tested

Model	WAIC	ΔWAIC	Adjusted GULR	Adjusted Agg Stop LR
Lognormal	-21.32	0.00	64.44%	0.34%
Gamma	-20.44	0.88	64.09%	0.19%
Weibull	-19.38	1.94	64.15%	0.32%
Inverse Burr	-20.38	0.94	64.76%	0.34%
Burr	-19.52	1.80	63.97%	0.30%

- Lognormal looks very close to bootstrap
- Most model criterion are much closer and results are similar
- What happened to the gamma, agg stop loss ratio went down?!
  - Hold this thought

Bayesian Analysis
Parameter probability comparison: lognormal distribution

Figure 4. Lognormal contour comparisons showing presence or absence of skew



- Bayesian parameter contour plot shows the skew absent from MLE
- No longer "symmetric"
- Estimated loss ratios are no longer "balance out" thus final weighted estimates show change

Bayesian Analysis
Parameter probability comparison: lognormal distribution II

- So important, it deserves to be shown twice!
- Maximum Likelihood results, assuming asymptotic normality, constrained by "prison of ellipticity"
- Bayesian parameters free to explore the parameter space

Figure 6. Superimposed logn... plots

Bayesian Analysis
Approximate Bayesian computation

- Closeness of bootstrap and fully Bayesian model not an accident
- Bootstrap is actually an example of approximate Bayesian computation
  - Full Bayesian model needs likelihood even using MCMC
  - Approximate Bayesian computation (ABC) doesn't even need likelihood
    - It needs data, a generative model, priors, and a matching criterion
- ABC algorithm:
  - Generate parameters from a prior
  - Generate data using sampled parameters
  - Retain or reject parameters based on matching criterion
  - Calculate posterior predictive distribution from "retained" data
- Exactly what Van Kampen did:
  - Matching criterion was "closeness of summary statistics"

 October 10, 2016

Bayesian Analysis
Properties of hierarchical Bayesian model

- Hierarchical models allow for explicit formulation of parameter dependence
- Generate correlated pairs of μ and σ by using multivariate normal
- Multivariate normal parameters allowed to explore their own parameter space based on data, unlike MLE case

Hierarchical Lognormal Model

```
data {
  int < lower = 0> N; // number of years
  vector < lower = 0.0 >[N] LR; // loss ratios
}
parameters {
  vector [2] mu; // multivariate normal (MVN ) prior mean : implied (0 ,0)
  corr_matrix [2] Sigma_corr ; // MVN prior correlation ( implied uniform )
  vector < lower = 0.0 >[2] Sigma_scale ; // prior scale for cov: implied (0 ,0)
  vector [2] Pars ; // correlated lognormal parameter vector
}
transformed parameters {
  real mu_post ;
  real < lower = 0.0 > sigma_post ;
  mu_post <- Pars [1];
  sigma_post <- exp( Pars [2]); // ensure non - negativity
}
model {
  matrix [2, 2] Sigma ; // will become MVN prior covariance
  Sigma <- diag_matrix ( Sigma_scale ) * cholesky_decompose ( Sigma_corr );
  Pars ~ multi_normal_cholesky (mu , Sigma ); // faster to use cholesky
  LR ~ lognormal ( mu_post , sigma_post );
}
generated quantities {
  real LR_post ; // distribution of LR based on posterior parameters
  real ASL_post ; // corresponding ASL
  vector [N] PointPosteriors ; // log pointwise predictive density
  for (i in 1:N) {
    PointPosteriors [i] <- exp( lognormal_log (LR[i], mu_post , sigma_post ));
  }
  LR_post <- lognormal_rng ( mu_post , sigma_post ); // stochastic observation of LR
  ASL_post <- fmin (0.025 , fmax ( LR_post - 0.725 , 0.0));
}
```

Simple Lognormal Model

```
data {
  int < lower = 0> N; // number of years
  vector < lower = 0.0 >[N] LR; // loss ratios
}
parameters {
  real mu; // prior for mu
  real < lower = 0.0 > sigma ; // non - negative prior for sigma
}
model {
  mu ~ normal (0.0 , 2.0);
  sigma ~ uniform (0.0 , 1.0);
  LR ~ lognormal (mu , sigma );
}
generated quantities {
  real LR_post ; // distribution of LR based on posterior parameters
  real ASL_post ; // corresponding ASL
  vector [N] PointPosteriors ; // log pointwise predictive density
  for (i in 1:N) {
    PointPosteriors [i] <- exp( lognormal_log (LR[i], mu , sigma ));
  }
  LR_post <- lognormal_rng (mu , sigma ); // stochastic observation of LR
  ASL_post <- fmin (0.025 , fmax ( LR_post - 0.725 , 0.0));
}
```

 October 10, 2016

Bayesian Analysis
Hierarchical fit comparisons

- Compare lognormal with gamma

Model	WAIC	ΔWAIC	Adjusted GULR	Adjusted Agg Stop LR
Lognormal	-21.44	0.00	64.25%	0.30%
Gamma	-20.32	0.12	64.21%	0.30%

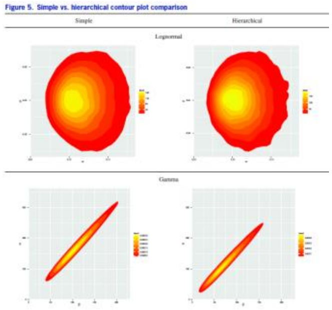
- Why did gamma fall back in line?

 October 10, 2016

Bayesian Analysis
Hierarchical Bayesian model insights

- Lognormal parameters are not correlated
  - Hierarchical model does not improve results
- Gamma parameters highly correlated
  - Hierarchical model shows clear change, leading to improvement

Figure 5. Simple vs. hierarchical contour plot comparison



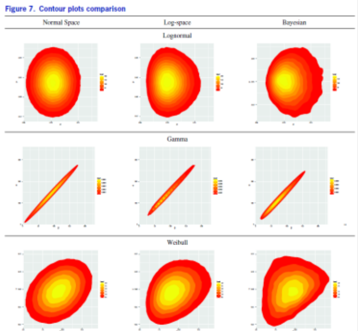
Section 4

Parameter Risk Revisited
Comparisons

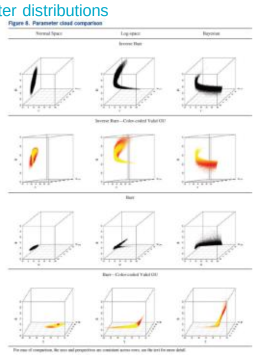
Comparisons
Contour plots: two-parameter distributions

- Normal Space tends to be elliptical and symmetric
- Lognormal space still shows some ellipticity
- Bayesian space can be non-elliptic
  - Weibull is almost triangular

Figure 7. Contour plots comparison



Comparisons
Parameter cloud: three-parameter distributions



- Normal Space tends to be ellipsoidal and symmetric
- Lognormal space better, but can be considered "stretched" ellipsoid
- Bayesian space can be weird
- Brighter colors mean larger expected GULR
  - Bayesian space shows structure

Comparisons
Goodness-of-fit and resulting estimates

	Maximum Likelihood			Bayesian		
Family	AICc	GU Effect	ASL Effect	WAIC	GU Effect	ASL Effect
Lognormal	-20.11	-0.004%	-1.97%	-21.44	0.16%	26.09%
Gamma	-20.04	0.23%	14.99%	-21.32	0.09%	25.25%
Weibull	-18.29	-0.66%	16.64%	-19.38	0.004%	34.63%
Inverse Burr	-15.76	166,464%	59.46%	-20.39	0.96%	44.02%
Burr	-15.76	474,023,000%	99.08%	-19.52	-0.28%	29.78%

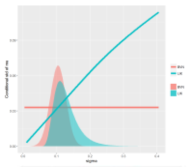
- Bayesian results more similar to each other and much less extreme than maximum likelihood
- In hindsight, under MLE, looks as if distortion of symmetrical lognormal "offset" distortion of highly correlated gamma, leading to similar GoF
- Bayesian scores of lognormal and gamma very similar—as are results!

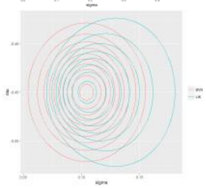
Comparisons
Recent developments—HOT OFF OF THE PRESSES I

- Analysis by John A. Major, Director of Actuarial Research, GC Analytics®
- Yes, the skew is what drives the change to the loss ratios
- **Q: How does it do that?**
- **A**: Standard errors of parameters calculated under MLE are **too small**
- Remember, (log)likelihood is a surface
  - Standard error of parameters is measured by Hessian at small area around maximum
  - Multivariate normal assumes parameters are homosecedastic
    - MVN mu and sigma always have same standard error
  - MLE/MVN appears to miss features of overall likelihood surface
  - Bayesian procedure can explore the overall space, and thus recover the moments, more freely

Comparisons
Recent developments—HOT OFF OF THE PRESSES II

- Conditional s.e. of bivariate normal mu (red) is constant regardless of value of sigma
- Bayesian posterior distribution of mu & sigma (blue) shows heteroscedasticiy
  - The s.e. of mu is not constant!
- Substituting the proper moments for mu and sigma would allow a MVN calculation to get close to the correct loss ratios
  - One cannot easily obtain proper moments from MLE
- Bayesian analysis does this automatically!





Figures courtesy of John A. Major

Comparisons
Recap

- MLE estimation, be it in normal or lognormal space does not fully explore the curvatures of the (log)likelihood space
  - Asymptotic results may not be good enough for a particular finite sample
- Bayesian model, especially hierarchical, can explore parameters space more freely, even for smaller samples

***THE* BRUCE DICKINSON RULE OF PARAMETER ESTIMATION**

Appendix A

Parameter Risk Revisited
References

References

- Adler, Avraham, "Estimating the Parameter Risk of a Loss Ratio Distribution—Revisited," *Variance* 9:1, **2015**, pp. 114–139.
  - All graphs, figures, and charts contained herein are derived from the above paper and personal conversations
- "More Cowbell," *Saturday Night Live*, NBC, April 8, 2000, Television
- Van Kampen, Charles E., "Estimating the Parameter Risk of a Loss Ratio Distribution," *Casualty Actuarial Society Forum,* **2003***, Spring*, 177-213

Appendix B

Parameter Risk Revisited
Disclaimers

About Guy Carpenter

Disclaimer

- Guy Carpenter & Company, LLC provides this presentation for general information only. The information contained herein is based on sources we believe reliable, but we do not guarantee its accuracy, and it should be understood to be general insurance/reinsurance information only. Guy Carpenter & Company, LLC makes no representations or warranties, express or implied. The information is not intended to be taken as advice with respect to any individual situation and cannot be relied upon as such. Please consult your insurance/reinsurance advisors with respect to individual coverage issues.
- Statements concerning tax, accounting, legal or regulatory matters should be understood to be general observations based solely on our experience as reinsurance brokers and risk consultants, and may not be relied upon as tax, accounting, legal or regulatory advice, which we are not authorized to provide. All such matters should be reviewed with your own qualified advisors in these areas.
- Readers are cautioned not to place undue reliance on any historical, current or forward-looking statements. Guy Carpenter & Company, LLC undertakes no obligation to update or revise publicly any historical, current or forward-looking statements, whether as a result of new information, research, future events or otherwise.
- Guy Carpenter is committed to adhering to antitrust laws, and cautions all recipients that this report is intended solely to provide general industry knowledge. Under no circumstances shall it be used as a means for representatives of competing companies, and/or firms, to reach any understanding whatsoever, whether it be about specific pricing of specific products, if particular products should be marketed to the public, or the terms under which products are marketed.
- Guy Carpenter undertakes to keep confidential all information concerning the business and affairs of its clients that may be obtained or received as a result interaction with such clients. Guy Carpenter will not, without the client's prior written consent (such consent not to be unreasonably withheld or delayed), disclose such information, in whole or in part, to any other person other than affiliates, employees, agents, professional advisers, or sub-contractors involved in the provision or receipt of (re)insurance brokerage services, or in accordance with normal (re)insurance broking practice to (re)insurers and their agents. The provisions of this paragraph will not apply to the information to the extent that it is (i) already lawfully in a party's possession on the date of its disclosure; (ii) in the public domain other than as a result of a breach of this clause; or (iii) required to be disclosed pursuant to legal, or regulatory requirements. Guy Carpenter may use for its own internal purposes, and include and disclose to third parties, on an anonymous and aggregate basis, information relating to (re)insurance transactions in benchmarking, modelling and other analytics offerings derived from such information.
- This document or any portion of the information it contains may not be copied or reproduced in any form without the permission of Guy Carpenter & Company, LLC, except that clients of Guy Carpenter & Company, LLC need not obtain such permission when using this report for their internal purposes.
- The trademarks and service marks contained herein are the property of their respective owners.

GUY CARPENTER