CLRS 2013
CLFM Estimates

Daniel Murphy
Trinostics LLC
with Emmanuel Bardis and Ali Majidi

Agenda

- CLFM and the R ChainLadder package
  - Finding the selection-consistent model
  - Graphing the link ratio function
  - A look at two diagnostic plots
  - Calculating IBNR and standard errors
  - Visualizing the estimated distribution of the predicted IBNR outcomes
- California Workers Comp data
- Questions for discussion

Users are demanding something be done! ☺

- Apr 8, 2009
  I am using the latest version of Chainladder in R 2.8.1 and have found it to be an excellent package indeed.

  There are occasions when the development factor may need to be selected as different from the output of the linear model. Is there a place in the MackChainLadder code where different development factors may be used?

  Thanks and Regards.

props, Markus!

Markus Gesmann

- Feb 27, 2013
  I agree with this proposal. We often have to choose specific coefficients. Could it be an option in the input of the functions bootchainladder and MackChainLadder?

  Thank you in advance.

CLFM in the ChainLadder Package

- ChainLadder (https://code.google.com/p/chainladder/)
  - A library of functions (a "package") for the R statistical environment (www.r-project.org)
  - Primarily targeted toward stochastic reserving
  - Originated and maintained by Markus Gesmann of Lloyds
    - Other contributing authors: Wayne Zhang and yours truly
  - Distributed under the GPL (General Public License)
    - Therefore, open-source, free to download, use, copy, modify, etc.
- Markus programmed the Mack method using linear regression models on the development periods
  - He used Barnett & Zehnwirth's ("Best Estimates for Reserves") delta (δ) notation for weighting the observations
    - So CLFM's α = Barnett & Zehnwirth's δ
  - He used Mack's recursive formula (1999 paper) to chain the standard error statistics together
    - Mack's formulas use alpha ($2-\delta$) for weighting the observations

Finding a selection-consistent model: CLFMdelta

- As explained by Manolis, a selection-consistent member of CLFM is a model whose expected value of the regression slope equals the actuary's selected RTR
- Use CLFMdelta(Triangle, selected, tolerance = .0005)
  - Triangle = loss data
  - selected = actuary's selected age-to-age factors
  - tolerance = proximity of found parameter to selected RTR
- selected = c(8.206, 1.624, 1.275, 1.175, 1.115, 1.042, 1.035, 1.018, 1.009)
- CLFMdelta(RAA, selected)

ChainLadder

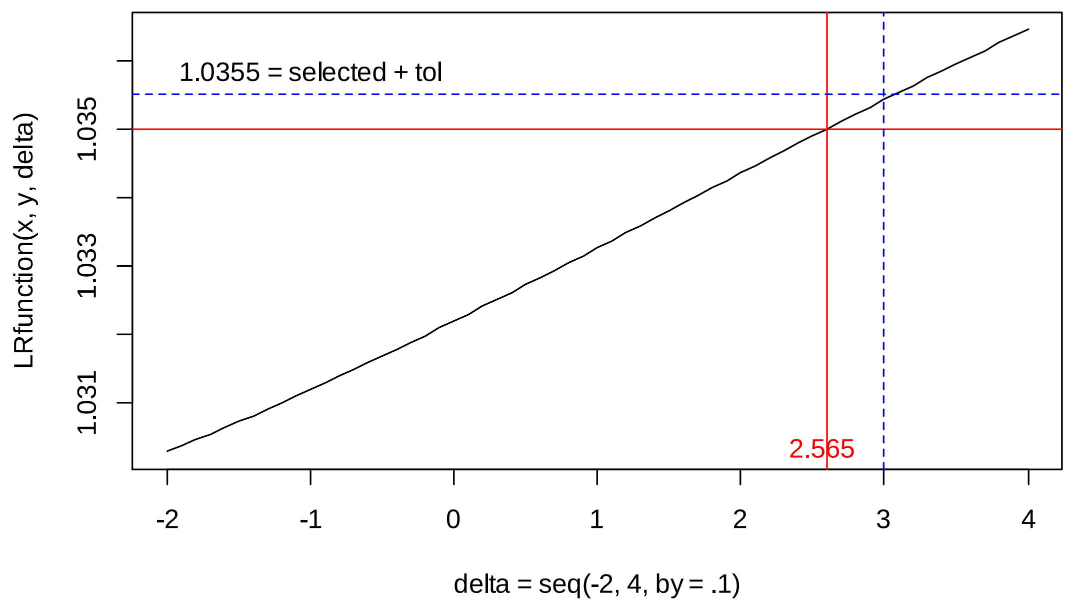
```
2.000000 1.000000 1.158150 1.305441 1.116562 1.000000 3.000000 2.000000 1.000000
```

Paper

```
2.000    1.000    1.158    1.305    1.117    1.000    2.565    2.005    2.005
```

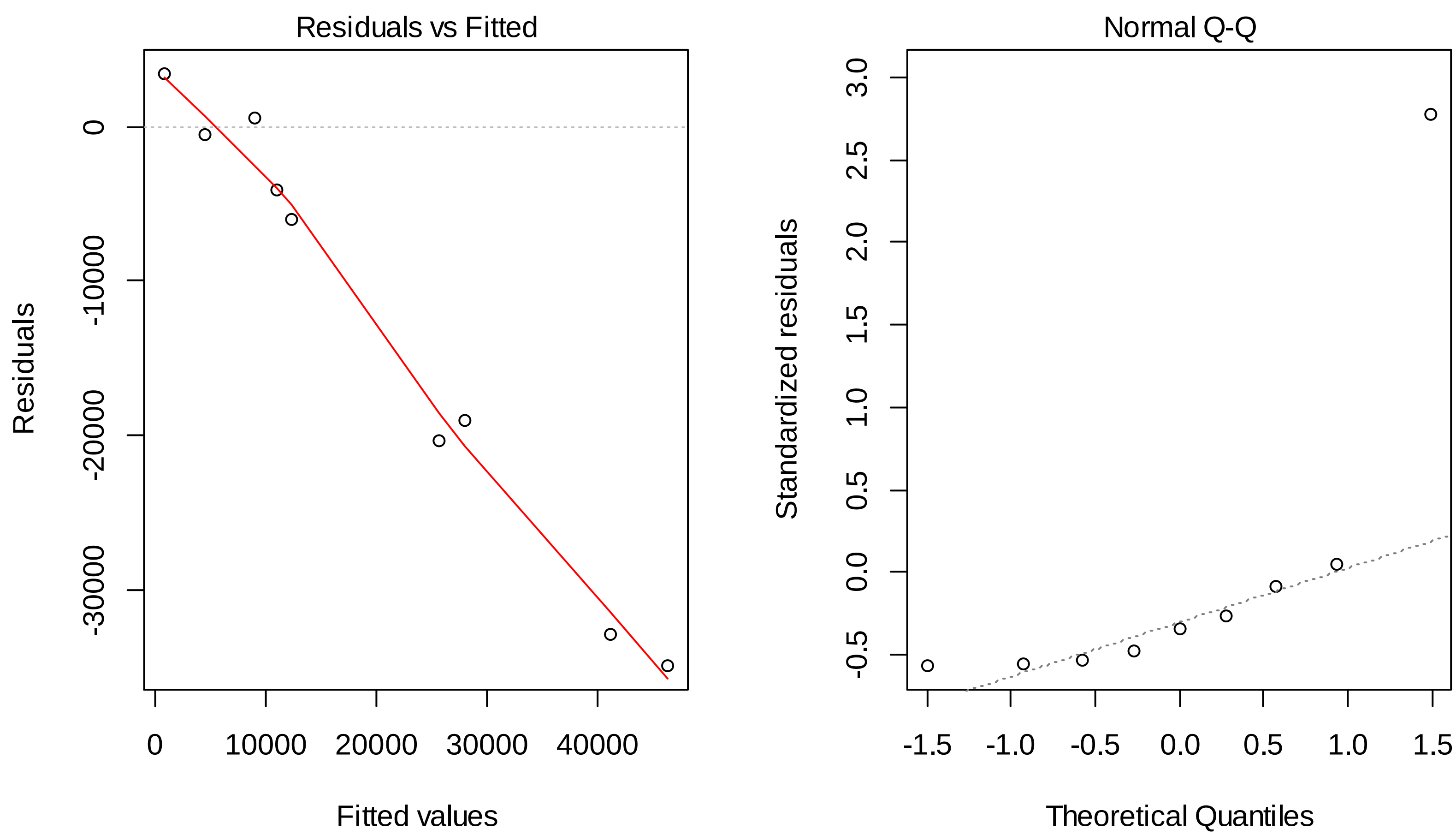
Visualizing the search for α: ChainLadder's LRfunction

- LRfunction(x, y, delta) ⇐ B&Z's δ notation
  - x = beginning value of loss during a development period
  - y = ending value of loss during a development period
  - delta = a real number or a vector of real numbers
- Here, x & y are the column 7 & 8 losses for development period 7-8



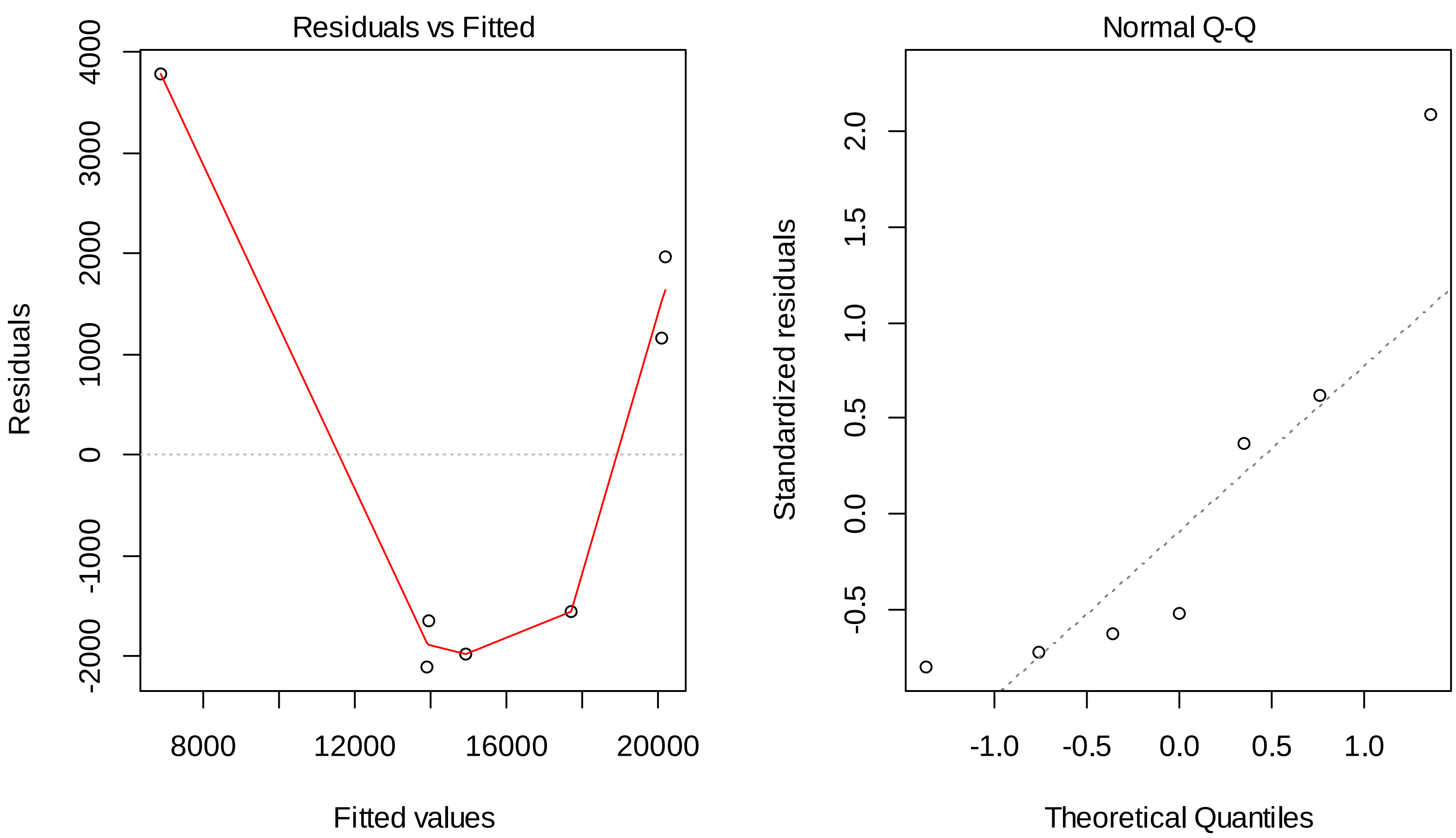
Inspect the residuals: Built-in diagnostics for linear model fits

Development period 1-2 (alpha = 2)



Two diagnostics (cont.)

Development period 3-4 (alpha = 1.158)



Running ChainLadder's `MackChainLadder` function with alpha = 2-delta gets close to a full CLFM implementation

```
> MackChainLadder(RAA, alpha = 2 - CLFMdelta(RAA, selected))
MackChainLadder(Triangle = RAA, alpha = 2 - CLFMdelta(RAA, selected))

     Latest Dev.To.Date Ultimate   IBNR Mack.S.E CV(IBNR)
1981 18,834      1.0000   18,834      0 0.00e+00      NaN
1982 16,704      0.9909   16,858    154 4.21e-01  0.00273
1983 23,466      0.9734   24,108    642 6.19e+02  0.96367
1984 27,067      0.9400   28,793  1,726 8.06e+02  0.46670
1985 26,180      0.9022   29,018  2,838 1.50e+03  0.53034
1986 15,852      0.8092   19,591  3,739 1.98e+03  0.52918
1987 12,314      0.6886   17,881  5,567 2.18e+03  0.39144
1988 13,112      0.5401   24,276 11,164 5.60e+03  0.50196
1989  5,395      0.3327   16,217 10,822 6.42e+03  0.59279
1990  2,063      0.0405   50,887 48,824 8.17e+04  1.67315

              Totals
Latest:   160,987.00
Dev:            0.65
Ultimate: 246,463.48
IBNR:      85,476.48
Mack S.E.: 82,651.02
CV(IBNR):       0.97
```

ChainLadder

Table 3. CLFM calculations for representative entries

AY/Age	Estimated Ultimate	Current Diagonal	Estimated Unpaid	Total Risk	CV
All	246,387	160,987	85,400	82,838	97.0%

Paper

- ✓ Although the `MackChainLadder` function does not provide for the psi-function process risk adjustment, the bottom line CVs are virtually identical

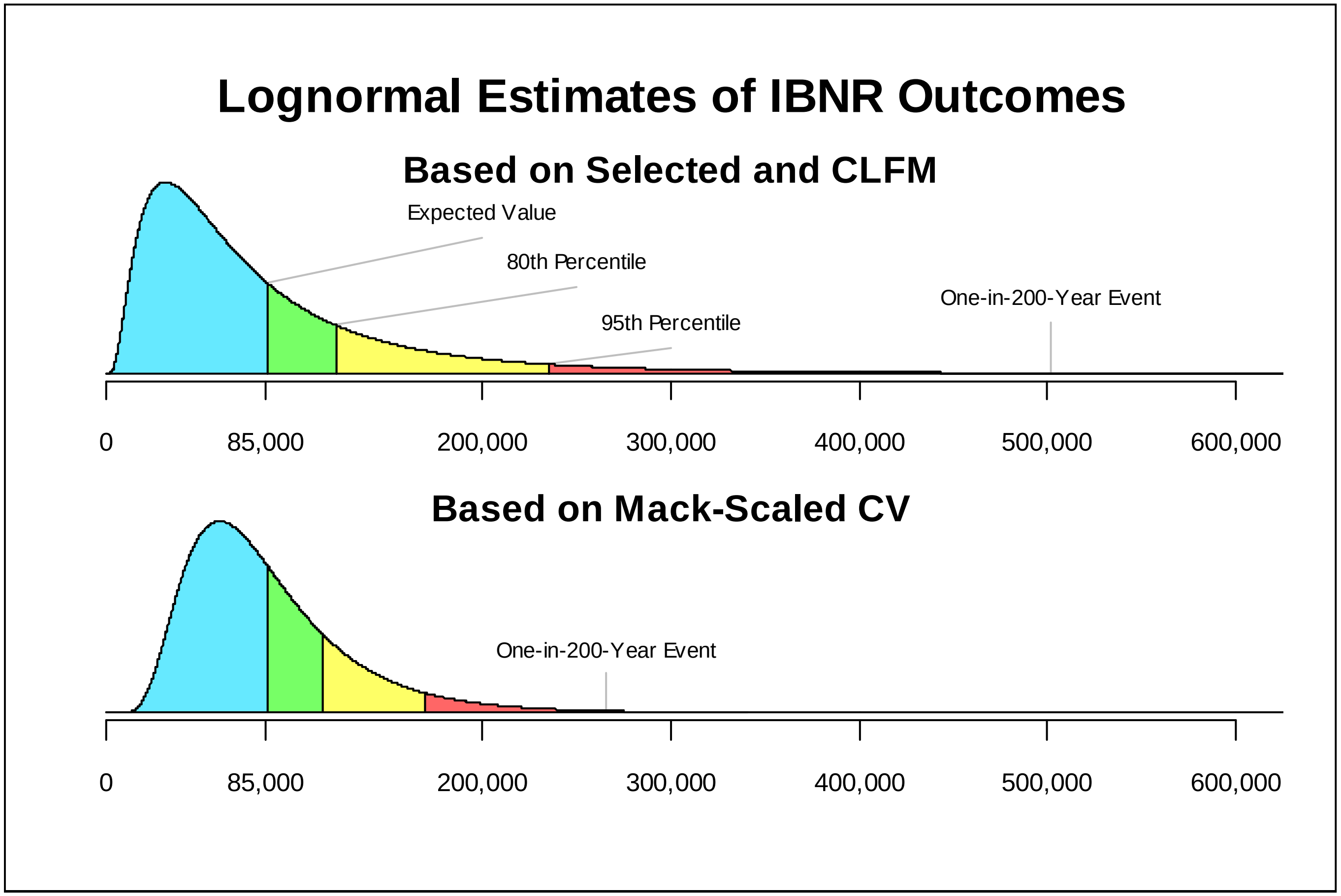
Running "Vanilla" Mack Method

```
> MackChainLadder(RAA)
MackChainLadder(Triangle = RAA)

     Latest Dev.To.Date Ultimate   IBNR Mack.S.E CV(IBNR)
1981 18,834       1.000   18,834      0        0      NaN
1982 16,704       0.991   16,858    154      143    0.928
1983 23,466       0.974   24,083    617      592    0.959
1984 27,067       0.943   28,703  1,636      713    0.436
1985 26,180       0.905   28,927  2,747    1,452    0.529
1986 15,852       0.813   19,501  3,649    1,995    0.547
1987 12,314       0.694   17,749  5,435    2,204    0.405
1988 13,112       0.546   24,019 10,907    5,354    0.491
1989  5,395       0.336   16,045 10,650    6,332    0.595
1990  2,063       0.112   18,402 16,339   24,566    1.503

              Totals
Latest:   160,987.00
Dev:            0.76
Ultimate: 213,122.23
IBNR:      52,135.23
Mack S.E.: 26,880.74
CV(IBNR):       0.52
```

IBNR is smaller
CV is smaller

- Can fit a lognormal to the mean and standard error
  - CLFM: use the IBNR and Mack S.E. on the previous page
  - Mack-scaled: use the IBNR on the previous page and standard error = IBNR on previous page times CV on this page (not recommended; non-cohesive model)

Visualizing the predictive IBNR distributions



WC Indemnity Paid Dollars

Paid Indemnity Loss Development ($millions)												
	Age (months)											
Acc Year	12	24	36	48	60	72	84	96	108	120	...	372
1979											...	410
1980											...	490
...				...	...	...	...	...	...	...		
2001			2,454	3,244	3,715	4,001	4,205	4,348	4,452	4,528		
2002		1,438	2,563	3,306	3,726	4,006	4,190	4,320	4,406	4,486		
2003	434	1,464	2,482	3,100	3,497	3,749	3,910	4,028	4,132	4,227		
2004	392	1,142	1,738	2,148	2,397	2,573	2,699	2,809	2,908			
2005	322	880	1,331	1,644	1,843	1,988	2,108	2,207				
2006	311	890	1,370	1,683	1,911	2,083	2,224					
2007	320	929	1,438	1,791	2,042	2,230						
2008	322	942	1,486	1,888	2,171							
2009	287	881	1,424	1,822								
2010	292	921	1,500									
2011	299	956										
2012	325											

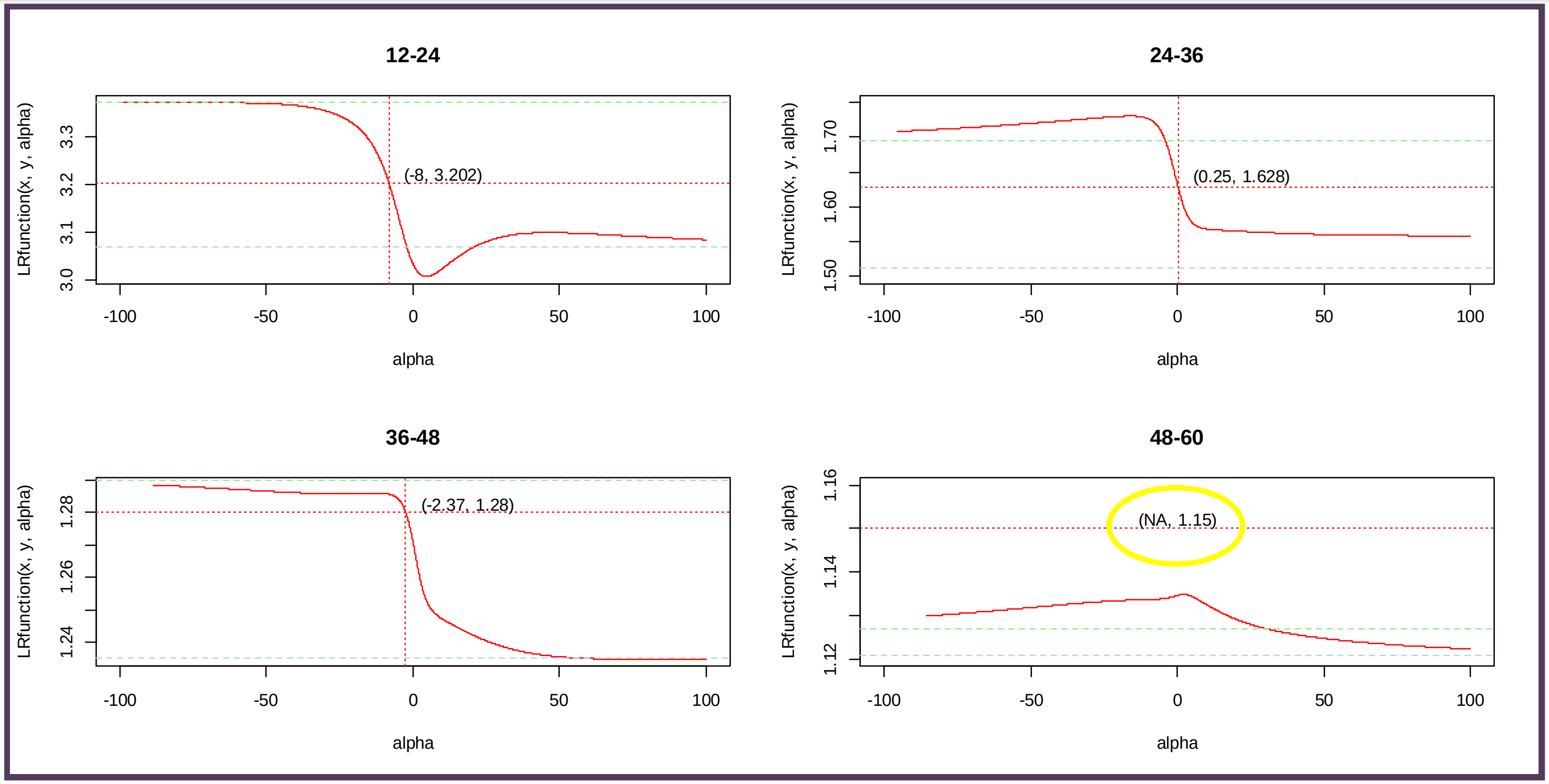
- California workers comp data evaluated 12/31/2012
- The green shaded cell is the observation with the minimum beginning value in that development period
- The blue shaded cell is the maximum beginning value

Link Ratios

Acc Year	24/12	36/24	48/36	60/48	72/60	84/72	96/84	108/96	120/108
2001			1.322	1.145	1.077	1.051	1.034	1.024	1.017
2002		1.782	1.290	1.127	1.075	1.046	1.031	1.020	1.018
2003	3.370	1.696	1.249	1.128	1.072	1.043	1.030	1.026	1.023
2004	2.914	1.522	1.236	1.116	1.073	1.049	1.041	1.035	
2005	2.734	1.512	1.235	1.121	1.079	1.060	1.047		
2006	2.866	1.539	1.229	1.135	1.090	1.068			
2007	2.905	1.547	1.246	1.140	1.092				
2008	2.927	1.577	1.271	1.150					
2009	3.069	1.616	1.280						
2010	3.154	1.628							
2011	3.202								
Selected	3.202	1.628	1.280	1.150	1.092	1.068	1.047	1.035	1.019

- The industry committee’s decision is to select the most recent factor
- The green cell in each column is the link ratio corresponding to the observation with the minimum beginning value
- The blue cell is corresponds to the observation with the maximum beginning value

Link Ratio Function
First Four Development Periods



- Red horizontal dotted line: selected value
- Red vertical dotted line: value of alpha such that LRfunction(alpha) = selected vale
- Asymptotes are at the link ratios of the AY with the minimum and maximum beginning values
  - Link ratios between asymptotes termed “reasonable” in paper
  - A less restrictive definition appears possible – an unsolved problem at this time

Project the development of unpaid loss to age 48 months

```
> triangle
      12   24   36   48
2001  NA   NA 2454 3244
2002  NA 1438 2563 3306
2003 434 1464 2482 3100
2004 392 1142 1738 2148
2005 322  880 1331 1644
2006 311  890 1370 1683
2007 320  929 1438 1791
2008 322  942 1486 1888
2009 287  881 1424 1822
2010 292  921 1500   NA
2011 299  956   NA   NA
2012 325   NA   NA   NA
```

trapezoid

```
> library(ChainLadder)
> delta <- CLFMdelta(Triangle = triangle,
+    selected = c(3.202, 1.628, 1.28))
> MackChainLadder(triangle,
+    alpha = 2 - delta,
+    est.sigma = "Mack",
+    mse.method = "Independence")
```

```
     Latest Dev.To.Date Ultimate  IBNR Mack.S.E CV(IBNR)
2001  3,244       1.000    3,244     0      0.0      NaN
2002  3,306       1.000    3,306     0      0.0      NaN
2003  3,100       1.000    3,100     0      0.0      NaN
2004  2,148       1.000    2,148     0      0.0      NaN
2005  1,644       1.000    1,644     0      0.0      NaN
2006  1,683       1.000    1,683     0      0.0      NaN
2007  1,791       1.000    1,791     0      0.0      NaN
2008  1,888       1.000    1,888     0      0.0      NaN
2009  1,822       1.000    1,822     0      0.0      NaN
2010  1,500       0.781    1,920   420     97.2    0.231
2011    956       0.480    1,992 1,036    171.1    0.165
2012    325       0.150    2,168 1,843    342.7    0.186

                Totals
Latest:      23,407.00
Dev:              0.88
Ultimate:    26,706.60
IBNR:         3,299.60
Mack S.E.:      402.17
CV(IBNR):         0.12
```

$\frac{1}{3.202 \cdot 1.628 \cdot 1.28} = 0.150$

Coefficient of Variation = 0.12

- Note that the default Mack Method using weighted average link ratios results in a CV of 0.09, which is 25% less than the CV indicated by the actual selected factors
- As of this writing, ChainLadder's S.E. calculation
  - limits alpha to the range [-4, 8]
  - does not yet reflect the PSI function adjustment

Possible questions for discussion

1. Under what circumstances might it be reasonable to expect the standard error of cumulative developed losses to be inversely proportional to the beginning value of loss ($\alpha<0$)?

2. What is the difference between the Chain Ladder method and the Loss Development method?

3. [per 2nd post on slide 1]
   Is it appropriate to carry out the England and Verrall bootstrap method given a triangle and an arbitrary set of selected link ratios? Why or why not?

Thanks

- To my co-authors Manolis and Ali Majidi for being the brains behind our paper

In abstentia

- To the many reviewers for their time, patience, and constructive comments, and their dedication to the CAS