Interpolation Along a Curve

Joseph Boor, FCAS, Ph.D., CERA
Actuary
The Florida Office of Insurance Regulation

Presentation to 2014 Casualty Actuarial Society Annual Meeting

November 11, 2014

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.

Why does Loss Development Factor interpolation matter?

- Have accident year development, need policy year development.
- Have triangle at regular twelve month December 31 periods, need to develop data at, say May 31.
- Some advisory organization benchmark data , has annual periods triangulated as of, say, March 31. Need to convert LDFs for use on standard December 31 data.

Why does Increased Limit Factor interpolation matter?

- Have existing increased limit factor table set up, want to expand the options available to policyholders.

Interpolation methods using basic algebra

- Linear interpolation.
- Geometric interpolation - fit growth curve to both points.
  - Remember paid loss correction from Berquist-Sherman.
- Take 1/LDF (=%incurred, %paid), and linearly interpolate 1/LDF, then invert.
- Geometric interpolation of 1/LDF.

Problems with basic algebra methods

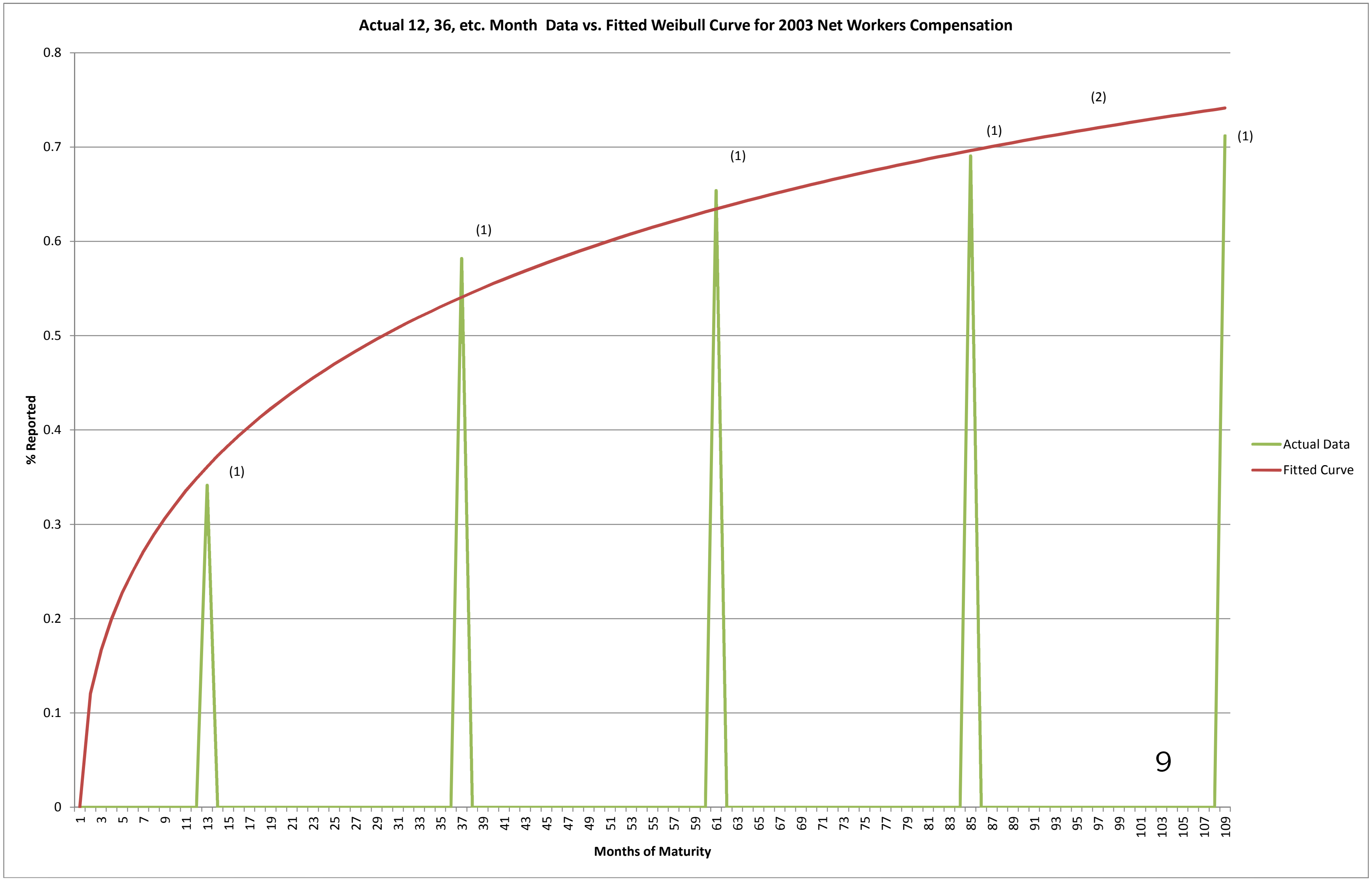
- Methods assume basic straight line or exponential curve
  - loss development incremental development is known to generally follow a “hump shaped” curve.
  - linear interpolation on ILFs fails the “Miccolis test”.

Alternatives–Curve Fitting

- Alternative is to fit a curve to the data, then read interpolated values off the curve.
  - Generally "hump shaped" Weibull probability mass function curve for loss levels emerging at each point.
    - % incurred or %paid then follows corresponding CDF
  - Generally, two parameter Pareto used for distribution of losses by size underlying ILFs.
    - ILF for limit "L" is $\frac{\int_0^L xf(x)dx+L(1-F(L))}{\int_0^B xf(x)dx+B(1-F(B))}$, where "$B$" = basic limit

Alternatives–Curve Fitting

- Curve fitting approach has some strong advantages, but it may have problems ...



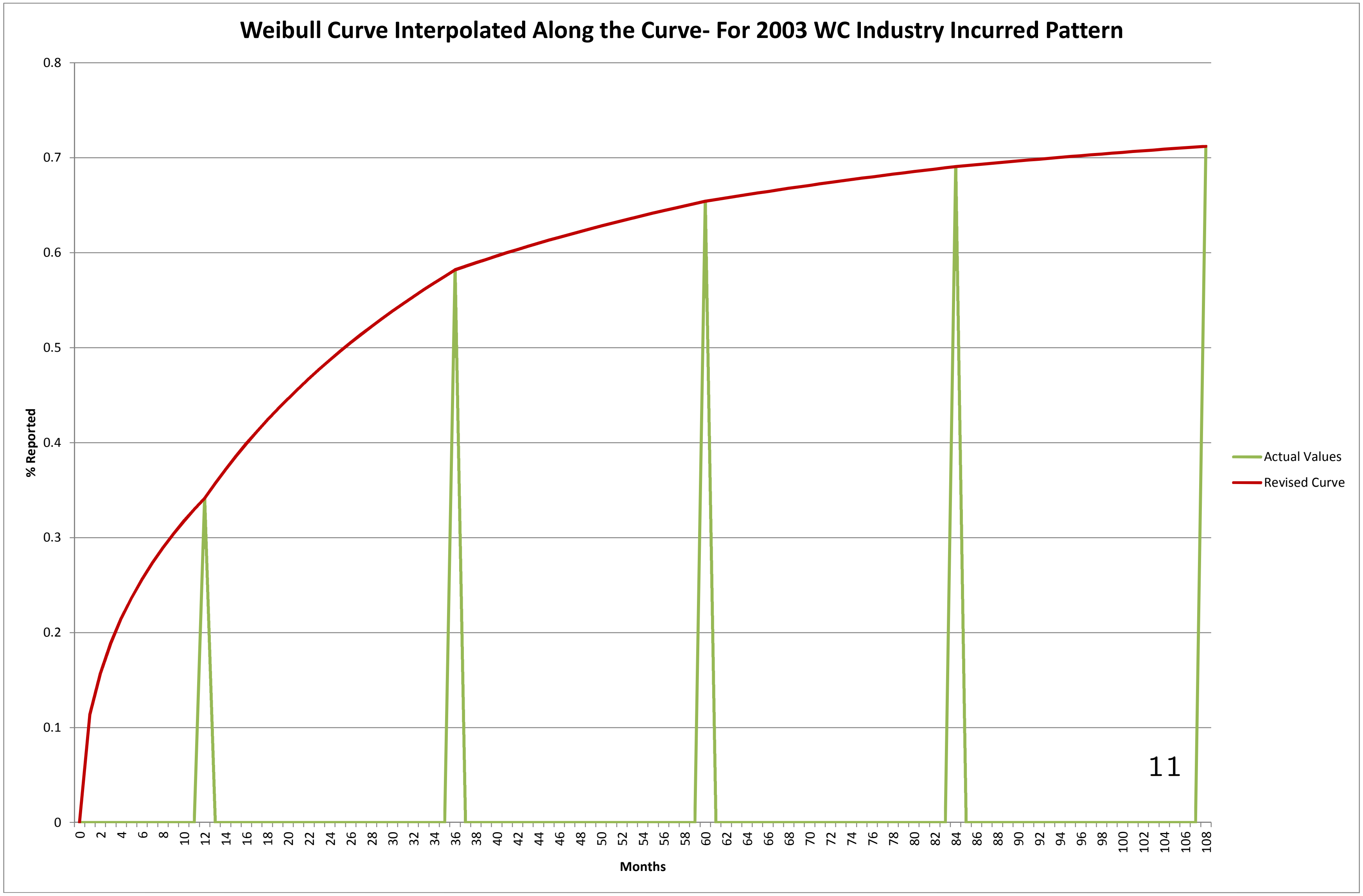
Actual 12, 36, etc. Month Data vs. Fitted Weibull Curve for 2003 Net Workers Compensation
% Reported
Months of Maturity
(1)
(1)
(1)
(1)
(2)
(1)
Actual Data
Fitted Curve

Correcting Fit Problems-Interpolation Along the Curve

- Stretch and rotate each segment of the fitted curve so that it hits the original data points exactly.

- if we have actual data points $d(t_0), d(t_1), d(t_2), \ldots, d(t_m)$ and a curve fitted to those points of $g(t)$, and we desire an estimate at $t^*, t_a < t^* < t_{a+1}, a \in 0, 1, 2, \ldots, m-1$, we take

$$\hat{d}(t^*) = d(t_a) + \frac{g(t^*) - g(t_a)}{g(t_{a+1}) - g(t_a)}[d(t_{a+1}) - d(t_a)].$$



Weibull Curve Interpolated Along the Curve- For 2003 WC Industry Incurred Pattern
% Reported
Months
Actual Values
Revised Curve

That is interpolating along the curve

Testing Against Alternatives

- vs. Straight linear interpolation, Straight Geometric Interpolation
- vs. Fitted Curve
  - for LDFs- Weibull regression
  - for ILFs - two parameter Pareto minimizing squared error in approximating known ILFs
- For LDFs - vs. linear, geometric interpolation of %paid/incurred

What is Weibull Regression?

- Weibull formula at time t for ILDFs is

$$1/ILDF(t) = \%Reported(t) = 1 - \exp(ct^b).$$

- So

$$1 - (1/ILDF(t)) = \%IBNR(t) = \exp(ct^b),$$

- or

$$\ln\left(\ln[1 - (1/ILDF(t))]\right) = \ln(c) + b\ln(t).$$

Weibull Regression-Part 2

- Given

$$\ln\left(\ln[1-(1/ILDF(t))]\right) = \ln(c) + b\ln(t),$$

- Just regress y values on left hand side to estimate Weibull parameters c and $\ln(b)$.

- For paid LDFs (PLDFs) regression works the same

$$\ln\left(\ln[1-(1/PLDF(t))]\right) = \ln(c) + b\ln(t),$$

LDF Testing Data

- Paid and (where possible) Incurred Development per 2003 Sked P - 11 Common Lines.
- Sample of 10 Small Co. Paid and (where possible) Incurred Development per (2011) Sked P.
- Use even (0,24, 48, etc.) LDFs to project odd (12, 36, etc.) LDFs and vice versa.

Batting Averages of the Methods - NAIC Data

- Winning Percentage of Intermediate LDF Value Estimates from Various Interpolation Methods vs. Interpolation Along the Curve

- Percentage of the Tests in Which Each Method Was Superior to Interpolation Along the Curve

Curve Fit to:	Number of Curves Fit	Winning % of Interp. Along the Curve	Geometric Interpolation	Linear Interpolation	Linear %Pd or Incrrd Interpolation	Geometric %Pd or Incrrd Interpolation	Unadjusted Weibull	Number of Times Weibull Outside Range
Even Maturity Paid LDFs	11	73 %	6 %	6 %	11 %	0 %	18 %	1
Odd Maturity Paid LDFs	11	77 %	5 %	5 %	9 %	2 %	14 %	5
Even Maturity Incrd LDFs	10	58 %	13 %	10 %	15 %	3 %	30 %	1
Odd Maturity Incrd LDFs	10	68 %	18 %	15 %	25 %	5 %	18 %	3
Straight Average		69 %	10 %	9 %	15 %	3 %	20 %	

Relative Errors of the Various Methods - NAIC Data

- Geometric Average Ratio of Squared Error of Intermediate LDF Value Estimates from Various Interpolation Methods to Squared Error of Interpolation Along the Curve

- Ratios Capped at 5% and 2000%

Curve Fit to:	Number of Curves Fit	Geometric Interpolation	Linear Interpolation	Linear % Pd or Incrrd Interpolation	Geometric % Pd or Incrrd Interpolation	Unadjusted Weibull
Even Maturity Paid LDFs	11	1034 %	1183 %	568 %	1781 %	316 %
Odd Maturity Paid LDFs	11	1277 %	2278 %	548 %	2931 %	527 %
Even Maturity Incrd LDFs	10	801 %	864 %	655 %	1401 %	190 %
Odd Maturity Incrd LDFs	10	694 %	943 %	366 %	1904 %	386 %
Straight Average		935 %	1235 %	524 %	1947 %	336 %

Batting Averages of the Methods - Small Company Data

- Winning Percentage of Intermediate LDF Value Estimates from Various Interpolation Methods vs. Interpolation Along the Curve

- Percentage of the Tests in Which Each Method Was Superior to Interpolation Along the Curve

Curve Fit to:	Number of Curves Fit	Winning % of Interp. Along the Curve	Geometric Interpolation	Linear Interpolation	Linear %Pd or Incrrd Interpolation	Geometric %Pd or Incrrd Interpolation	Unadjusted Weibull	Number of Times Weibull Outside Range
Even Maturity Paid LDFs	10	38 %	27 %	23 %	33 %	7 %	43 %	3
Odd Maturity Paid LDFs	10	50 %	20 %	20 %	25 %	8 %	30 %	4
Even Maturity Incrd LDFs	7	43 %	38 %	33 %	29 %	19 %	36 %	3
Odd Maturity Incrd LDFs	8	28 %	28 %	28 %	34 %	9 %	38 %	5
Straight Average		40 %	28 %	26 %	30 %	11 %	36 %	

Relative Errors of the Various Methods - Small Company Data

- Geometric Average Ratio of Squared Error of Intermediate LDF Value Estimates from Various Interpolation Methods to Squared Error of Interpolation Along the Curve

- Ratios Capped at 5% and 2000%

Curve Fit to:	Number of Curves Fit	Geometric Interpolation	Linear Interpolation	Linear % Pd or Incrrd Interpolation	Geometric % Pd or Incrrd Interpolation	Unadjusted Weibull
Even Maturity Paid LDFs	10	365 %	419 %	232 %	943 %	147 %
Odd Maturity Paid LDFs	10	473 %	658 %	302 %	1362 %	265 %
Even Maturity Incrd LDFs	7	238 %	255 %	336 %	423 %	110 %
Odd Maturity Incrd LDFs	8	341 %	406 %	177 %	1453 %	201 %
Straight Average		355 %	429 %	253 %	985 %	176 %

Conclusions - LDF Testing

- Interpolating along the Weibull curve is the most accurate option.
- Benefits vs. straight Weibull curve are reduced when loss volume is thin.

Testing of Pareto Estimation of ILFs

- Got seven sets of excess factors from NCCI-converted them to ILFs
- High granularity of NCCI factors made accurate interpolation too easy.
- Selected key values \$25,000, \$50,000, \$75,000, \$100,000, \$150,000, etc.
- Fitted curve to odds to predict evens and vice versa.

Testing of Pareto Estimation of ILFs - Curve Fitting

- Fit curve to minimize sum of squared errors vs. real data of (α, T) Pareto distribution at the points you're fitting to, the L's.

$$ILF(L, \$250,000) = \frac{\alpha - \left(\frac{T}{L}\right)^{\alpha-1}}{\alpha - \left(\frac{T}{\$250{,}000}\right)^{\alpha-1}}.$$

- \$250,000 pre-chosen as basic limit (where ILF = 1.)

Batting Averages of the Methods - NCCI ILF-type Data

- Winning Percentage of Intermediate ILF Value Estimates from Various Interpolation Methods vs. Interpolation Along the Curve

	Fitted Curve	Interp. Along the Curve	Linear Interpolation	Geometric Interpolation
	12 %	81 %	7 %	0

Relative Errors of the Various Methods - NCCI ILF-type Data

- Geometric Average Ratio of Squared Errors Relative to ILF Interpolation Along the Curve

	Fitted Curve	Interp. Along the Curve	Linear Interpolation	Geometric Interpolation
Sq. Error Ratio	703 %	100 %	592 %	740 %

Conclusions - ILF Testing

- Interpolating along the curve is the most accurate option.

Last Step - Cubic Splines

- Alternate view from numerical analysts
  - Cubic Splines is interpolation method of choice
    - Cubic polynomial between each two data points
    - Hit each data point
    - First and second derivatives match at each data point
    - Second derivative zero at outer endpoints
  - Details and calculation spreadsheet in the paper.

Skip to Relative Errors of Cubic Splines - NAIC Data

- Geometric Average Ratio of Error of Intermediate LDF Value Estimates from Cubic Splines vs. Interpolation Along the Curve

- Ratios Capped at 2000% Above and 5% Below

Curve Fit to:	Number of Curves Fit	Error Ratio
Even Maturity Paid LDFs	11	188 %
Odd Maturity Paid LDFs	11	133 %
Even Maturity Incrd LDFs	10	271 %
Odd Maturity Incrd LDFs	10	152 %
Straight Average		178 %

Relative Errors of Cubic Splines - Small Company Data

- Geometric Average Ratio of Error of Intermediate LDF Value Estimates from Cubic Splines vs. Interpolation Along the Curve

- Ratios Capped at 2000% Above and 5% Below

Curve Fit to:	Number of Curves Fit	Error Ratio
Even Maturity Paid LDFs	10	274 %
Odd Maturity Paid LDFs	10	169 %
Even Maturity Incrd LDFs	7	201 %
Odd Maturity Incrd LDFs	8	142 %
Straight Average		193 %

Relative Errors of Cubic Splines - NCCI ILF-type Data

- Geometric average ratio of error of intermediate LDF value estimates from cubic splines vs. interpolation along the curve = 254%.

- Standard capping used

Summary

- As confirmed by testing, on average interpolation along the curve produces the most accurate estimates of all the methods reviewed.

???