CP4: Fitting and Bootstrapping GLMs for Incremental Development Triangles

Thomas Hartl, PwC LLP

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.

Overview

- Session is based on two call papers
  - Fitting a GLM to Incomplete Development triangles
    - Detailed description of model and how to go about fitting it in MS Excel using Visual Basic
  - Bootstrapping GLMs for Development Triangles using Deviance Residuals
    - Algorithm for rescaling deviance residuals and case study of bootstrapping with Pearson residuals vs bootstrapping with deviance residuals

Objectives

- Understand issues encountered when fitting a regression model to an incomplete development triangle
- Understand nature of bootstrapping
- Understand some practical limitations encountered when bootstrap based on residual resampling is employed

Fitting a GLM to Incomplete Development Triangles

- Outline of presentation
  - Description of the model
  - Issues encountered when dealing with incomplete triangles
  - Quick introduction to graph theory
  - What can be learned about the model for a particular development triangle

Description of the model

- Multiplicative factorial GLM for incremental development amounts (using exposure and development period parameters)
- Reserve projection based on out-of-sample projection of future incremental development amounts
- Fit is accomplished using pseudo-likelihood framework – i.e. model is specified by choice of variance function

Description of the model

- Multiplicative GLM ⇨ log link function
- Factorial model ⇨ discrete parameters
- Out-of-sample projection ⇨ we fit a regression model to past development amounts
- Pseudo-likelihood ⇨ fitting procedure only depends on second moment assumptions

Description of model

- Model is linear on log scale:

γ	$\gamma+\beta_2$	$\gamma+\beta_3$	$\gamma+\beta_4$	$\gamma+\beta_5$
$\gamma+\alpha_2$	$\gamma+\alpha_2+\beta_2$	$\gamma+\alpha_2+\beta_3$	$\gamma+\alpha_2+\beta_4$	
$\gamma+\alpha_3$	$\gamma+\alpha_3+\beta_2$	$\gamma+\alpha_3+\beta_3$		
$\gamma+\alpha_4$	$\gamma+\alpha_4+\beta_2$			
$\gamma+\alpha_5$				

Issues with incomplete triangles

- Not enough data points for all parameters

γ	$\gamma+\beta_2$	$\gamma+\beta_3$	X	$\gamma+\beta_5$
$\gamma+\alpha_2$	$\gamma+\alpha_2+\beta_2$	$\gamma+\alpha_2+\beta_3$	X	
$\gamma+\alpha_3$	$\gamma+\alpha_3+\beta_2$	$\gamma+\alpha_3+\beta_3$		
$\gamma+\alpha_4$	$\gamma+\alpha_4+\beta_2$			
$\gamma+\alpha_5$				

Issues with incomplete triangles

- Choice of reference cell matters after all

X	X	X	X	X
$\gamma+\alpha_2$	$\gamma+\alpha_2+\beta_2$	$\gamma+\alpha_2+\beta_3$	$\gamma+\alpha_2+\beta_4$	
$\gamma+\alpha_3$	$\gamma+\alpha_3+\beta_2$	$\gamma+\alpha_3+\beta_3$		
$\gamma+\alpha_4$	$\gamma+\alpha_4+\beta_2$			
$\gamma+\alpha_5$				

Issues with incomplete triangles

- Data splits into unrelated regions

X	X	$\gamma+\beta_3$	$\gamma+\beta_4$	$\gamma+\beta_5$
X	X	$\gamma+\alpha_2+\beta_3$	$\gamma+\alpha_2+\beta_4$	
X	X	$\gamma+\alpha_3+\beta_3$		
$\gamma+\alpha_4$	$\gamma+\alpha_4+\beta_2$			
$\gamma+\alpha_5$				

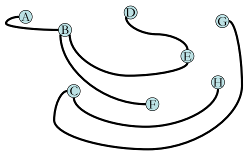
Issues with incomplete triangles

- Exact fit cells

X	X	$\gamma+\beta_3$	$\gamma+\beta_4$	***$\gamma+\beta_5$***
X	X	$\gamma+\alpha_2+\beta_3$	$\gamma+\alpha_2+\beta_4$	
$\gamma+\alpha_3$	$\gamma+\alpha_3+\beta_2$	***$\gamma+\alpha_3+\beta_3$***		
$\gamma+\alpha_4$	$\gamma+\alpha_4+\beta_2$			
$\gamma+\alpha_5$				

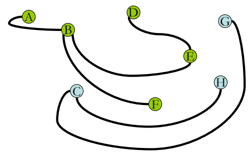
Quick intro to graph theory

- A graph is a collection of NODES which are pair-wise connected by EDGES



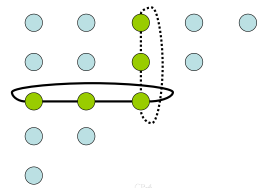
Quick intro to graph theory

- Maximal connected components
  - A, B, D, E & F
  - C, G & H



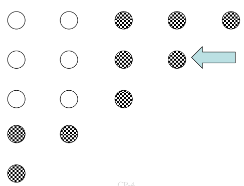
Quick intro to graph theory

- Development triangles as graphs:
  - All cells in a row are pair-wise connected
  - All cells in a column are pair-wise connected



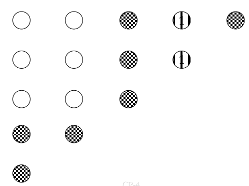
Quick intro to graph theory

- Breadth first search for triangles
  - Start with all included cells untested
  - Pick one cell to start with



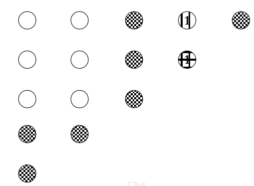
Quick intro to graph theory

- Breadth first search for triangles (step 3)
  - Mark all cells in column of first untested cell with component counter and column tested flag



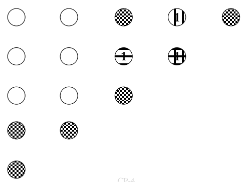
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



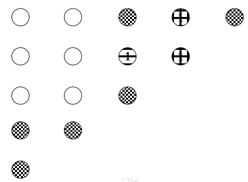
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



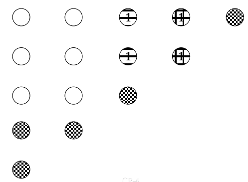
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



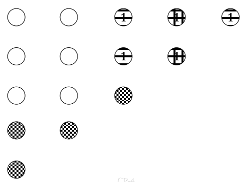
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



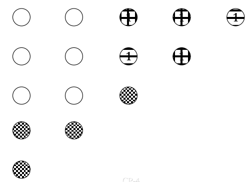
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



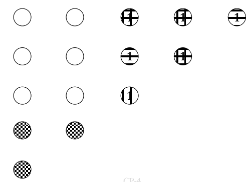
Quick intro to graph theory

- Breadth first search for triangles (step 6)
  - Loop over row tested cells: mark cell as done and mark other cells in column as column tested



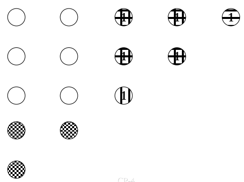
Quick intro to graph theory

- Breadth first search for triangles (step 6)
  - Loop over row tested cells: mark cell as done and mark other cells in column as column tested



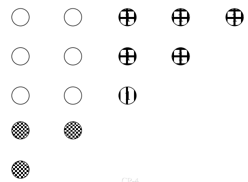
Quick intro to graph theory

- Breadth first search for triangles (step 6)
  - Loop over row tested cells: mark cell as done and mark other cells in column as column tested



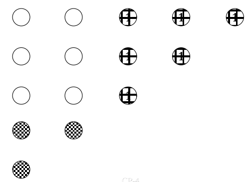
Quick intro to graph theory

- Breadth first search for triangles (step 6)
  - Loop over row tested cells: mark cell as done and mark other cells in column as column tested



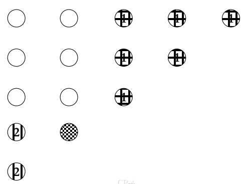
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



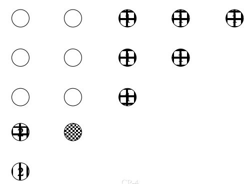
Quick intro to graph theory

- Breadth first search for triangles (step 3)
  - Mark all cells in column of first untested cell with component counter and column tested flag



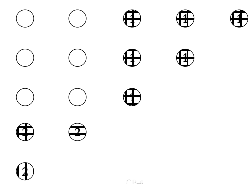
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



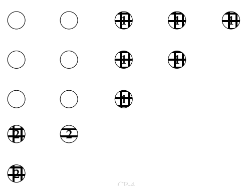
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



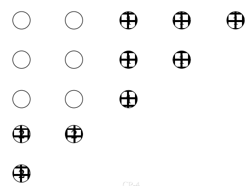
Quick intro to graph theory

- Breadth first search for triangles (step 5)
  - Loop over column tested cells: mark cell as done and mark other cells in row as row tested



Quick intro to graph theory

- Breadth first search for triangles (step 6)
  - Loop over row tested cells: mark cell as done and mark other cells in column as column tested

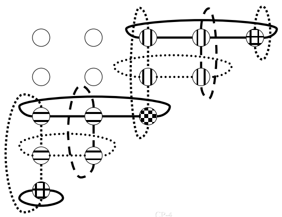


What do we learn?

- We can use the Breadth First algorithm to find the maximal connected components of an incomplete development triangle ⇨ <u>Projecting future development amounts is only possible within the row and column range of each maximal connected component</u>
- For each connected component we can also analyze what each cell contributes to our knowledge of the inherent variability

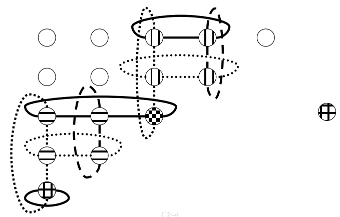
What do we learn?

- Within a maximal connected component there are three different types of nodes



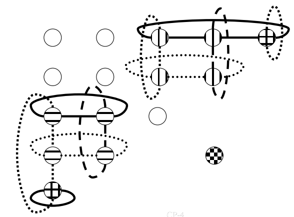
What do we learn?

- Effect of removing a single parameter cell



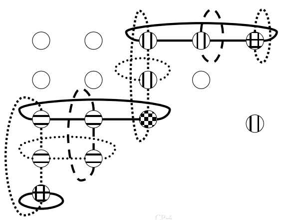
What do we learn?

- Effect of removing a critical connector cell



What do we learn?

- Effect of removing a regression cell



What do we learn?

- Single parameter cells and critical connector cells are exact fit cells ⇨ <u>no information about variability</u> for these cells
- Fit for connected components of regression cells is independent of what is going on in rest of triangle ⇨ can be used to <u>split regression fit into isolated subcomponents</u> (if there are any critical connector cells)

What else is in the call paper?

- Section 3 covers how to fit a GLM using MS Excel based Visual Basic code
- Section 4 covers how to calculate and plot standardized residuals
- Spreadsheet with illustrative implementation of algorithms discussed in call paper is available from author at request

Illustrative spreadsheet

- Input 10 x 10 triangle
- Select data points to include in model
- Analyze graph topology of incomplete triangle
- Choose variance function
- Fit GLM to incomplete triangle
- Study standardized residual plots
- Bootstrap range of reserve outcomes using Pearson residuals of Deviance residuals

Bootstrapping GLMs for Development Triangles using Deviance Residuals

- Not covered in presentation: Newton-Raphson algorithm for rescaling deviance residuals based on identity variance function
- Covered in presentation: case study of bootstrapping with Pearson residuals vs bootstrapping with deviance residuals

Bootstrapping GLMs for Development Triangles using Deviance Residuals

- Outline of presentation
  - What is bootstrapping?
  - Linear rescaling with Pearson residuals
  - Non-linear rescaling with Deviance residuals
  - Demonstration I: negative resampling values
  - Demonstration II: non-linear rescaling not possible
  - What do we learn?

What is bootstrapping?

- Approximates the distribution of a function that depends on sampled data
- Assumes that data is randomly distributed according to specified stochastic model
- Uses observed error structure to approximate random distributions of model

Any distributions derived are conditional on specified stochastic model being correct

Bootstrapping and Stochastic Reserving

- Reserves are a function of development triangle
- Get bootstrap distribution of reserve estimates by repeatedly resampling triangle
- Above only gives parameter uncertainty
- To approximate distribution of reserve outcomes we also need process error
- Can approximate process error using the same resampling procedure used for triangle

Bootstrapping and Heteroscedasticity

- Use resampling of standardized residuals to adjust for non-constant error structure
- Multiple definitions for residuals available
- Residual rescaling is the inverse process of residual standardization
- Want to approximate distributions of data points ⇨ resampling distributions should be consistent with stochastic model assumptions

Rescaling Example

- Data set Taylor & Ashe (1983)

357,848	766,940	610,542	482,940	527,326	574,398	146,342	139,950	227,229	67,948
352,118	884,021	933,894	1,183,289	445,745	320,996	527,804	266,172	425,046	
290,507	1,001,799	926,219	1,016,654	750,816	146,923	495,992	280,405		
310,608	1,108,250	776,189	1,562,400	272,482	352,053	206,286			
443,160	693,190	991,983	769,488	504,851	470,639				
396,132	937,085	847,498	805,037	705,960					
440,832	847,631	1,131,398	1,063,269						
359,480	1,061,648	1,443,370							
376,686	986,608								
344,014									

Rescaling Example

- Data set Taylor & Ashe (1983) – Fitted Values

140,801	*338,807*	*431,201*	*358,694*	*242,579*	*197,553*	**185,516**	**116,383**	**211,622**	**67,948**
293,186	*705,487*	*897,876*	*746,898*	**505,115**	**411,359**	**386,295**	**242,341**	**440,653**	*141,486*
396,579	*954,279*	*1,214,515*	**1,010,295**	**683,246**	*556,426*	**522,523**	**327,803**	*596,051*	*191,382*
214,098	*515,178*	**655,669**	*545,419*	**368,858**	**300,393**	**282,090**	*176,968*	*321,785*	*103,319*
307,853	**740,778**	**942,791**	**784,261**	**530,383**	**431,937**	*405,619*	*254,464*	*462,697*	*148,564*
343,763	**827,188**	**1,052,766**	**875,744**	**592,251**	*482,321*	*452,933*	*284,146*	*516,669*	*165,893*
386,316	**929,583**	**1,183,083**	**984,148**	*665,564*	*542,025*	*509,000*	*319,320*	*580,625*	*186,429*
442,821	**1,065,549**	**1,356,128**	*1,128,096*	*762,913*	*621,305*	*583,450*	*366,025*	*665,551*	*213,697*
400,230	**963,064**	*1,225,695*	*1,019,595*	*689,536*	*561,548*	*527,333*	*330,821*	*601,538*	*193,143*
344,014	*827,792*	*1,053,534*	*876,383*	*592,684*	*482,673*	*453,264*	*284,354*	*517,046*	*166,014*

Rescaling Pearson Residuals

- Definition residuals:

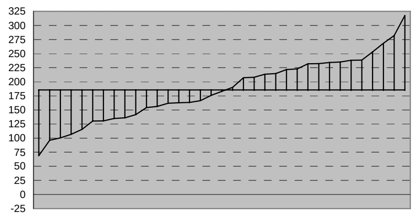
$$r_{\mathrm{P}} = \frac{y - \hat{y}}{\sqrt{V(\hat{y})}}$$

- Definition resampling distribution:

$$\mathbf{y}_{\mathrm{P}}^{*} = \hat{y} + \sqrt{V(\hat{y})} \cdot \mathbf{s}$$

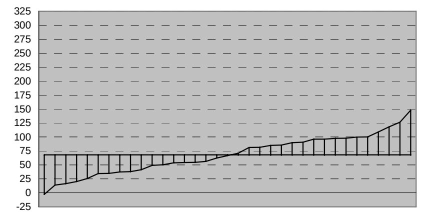
Rescaling Pearson Residuals

- Resampling distribution - fitted mean of 185,586



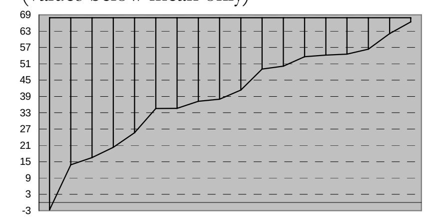
Rescaling Pearson Residuals

- Resampling distribution - fitted mean of 67,948



Rescaling Pearson Residuals

- Resampling distribution - fitted mean of 67,948 (values below mean only)



Rescaling Deviance Residuals

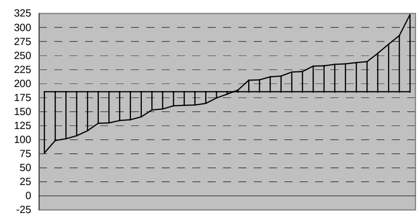
- Definition residuals (identity variance function):

$$r_{\mathrm{D}} = \mathrm{sign}(y - \hat{y}) \cdot \sqrt{2\left(y \cdot \log\left(y / \hat{y}\right) - y + \hat{y}\right)}$$

- Definition resampling distribution:
  - No closed form expression available
  - Substitute **s** for r_{D} in above equation and numerically solve for y
  - Need slight correction to make sure model assumption about variance function is satisfied

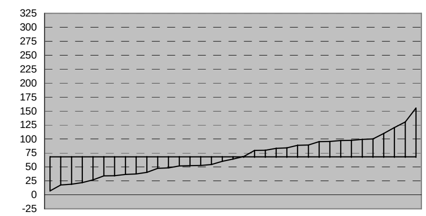
Rescaling Deviance Residuals

- Resampling distribution - fitted mean of 185,586



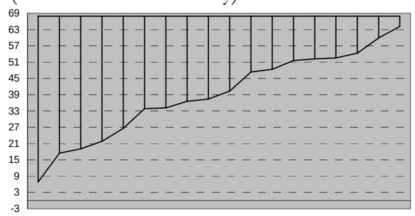
Rescaling Deviance Residuals

- Resampling distribution - fitted mean of 67,948



Rescaling Deviance Residuals

- Resampling distribution - fitted mean of 67,948 (values below mean only)



Demonstration I

- Negative resampling values
  - Top right cell is only cell for which we get a negative resampling value
  - Can directly compare bootstrapping results with Pearson and deviance residuals for model excluding top right cell
  - Bootstrapping with deviance residuals is also possible for model including top right cell

Demonstration I

- Bootstrapping results excluding top right corner
  - Pearson residuals (10,000 iterations)

Accident Period	Modeled Reserve	Bootstrap Projection	Sim. Future Development	Standard Pred. Error	5%-ile Sim. Outcome	95%-ile Sim. Outcome
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	596,051	603,398	595,127	166,522	(254,940)	288,038
4	498,753	504,064	498,789	135,273	(214,047)	231,751
5	1,122,779	1,134,746	1,125,780	224,917	(345,901)	394,656
6	1,736,070	1,751,181	1,734,825	302,852	(467,485)	522,686
7	2,616,534	2,640,194	2,612,849	407,758	(613,245)	724,636
8	4,127,340	4,164,901	4,132,367	586,633	(892,087)	1,040,074
9	4,956,065	4,990,267	4,959,138	801,618	(1,232,452)	1,417,929
10	5,087,731	5,161,854	5,082,052	1,393,141	(2,030,612)	2,510,119
Total	**20,741,324**	**20,950,606**	**20,740,927**	**2,504,915**	**(3,645,668)**	**4,603,584**

Demonstration I

- Bootstrapping results excluding top right corner
  - Deviance residuals (10,000 iterations)

Accident Period	Modeled Reserve	Bootstrap Projection	Sim. Future Development	Standard Pred. Error	5%-ile Sim. Outcome	95%-ile Sim. Outcome
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	596,051	601,425	595,682	165,133	(254,824)	283,543
4	498,753	502,686	497,483	135,937	(213,619)	233,681
5	1,122,779	1,130,897	1,122,776	225,374	(348,761)	388,058
6	1,736,070	1,748,560	1,735,691	300,235	(460,344)	514,240
7	2,616,534	2,636,940	2,619,302	409,205	(630,112)	709,919
8	4,127,340	4,156,304	4,128,423	582,196	(885,958)	1,016,114
9	4,956,065	5,002,022	4,962,549	802,422	(1,215,009)	1,405,855
10	5,087,731	5,169,300	5,088,560	1,404,841	(2,048,259)	2,501,894
Total	**20,741,324**	**20,948,135**	**20,750,465**	**2,530,813**	**(3,764,693)**	**4,599,894**

Demonstration I

- Bootstrapping results including top right corner
  - Deviance residuals (10,000 iterations)

Accident Period	Modeled Reserve	Bootstrap Projection	Sim. Future Development	Standard Pred. Error	5%-ile Sim. Outcome	95%-ile Sim. Outcome
1	-	-	-	-	-	-
2	141,486	148,558	141,810	99,435	(142,027)	181,427
3	787,433	802,512	786,345	227,758	(332,132)	415,547
4	602,073	612,774	600,459	168,556	(252,157)	302,197
5	1,271,343	1,290,547	1,271,004	266,900	(394,291)	476,089
6	1,901,963	1,926,750	1,906,391	343,783	(513,444)	607,984
7	2,802,963	2,834,990	2,804,315	448,446	(679,871)	795,858
8	4,341,037	4,384,089	4,338,730	639,559	(958,332)	1,144,621
9	5,149,209	5,209,231	5,145,549	844,468	(1,259,637)	1,509,566
10	5,253,745	5,354,869	5,249,988	1,444,013	(2,074,331)	2,567,191
Total	**22,251,251**	**22,564,319**	**22,244,592**	**2,868,629**	**(4,054,094)**	**5,235,817**

Demonstration II

- Data set Taylor & Ashe (1983) – Fitted Values

254,672	*611,704*	*774,193*	*665,389*	*434,726*	**320,588**	**299,529**	**184,715**	**283,087**	**67,948**
332,131	*797,756*	*1,009,667*	*867,770*	**566,950**	**418,096**	**390,632**	**240,897**	**369,188**	*88,615*
359,730	*864,049*	*1,093,569*	**939,880**	**614,062**	**452,839**	**423,093**	**260,915**	*399,867*	*95,978*
223,757	*537,449*	**680,214**	*584,618*	**381,955**	**281,672**	**263,169**	*162,293*	*248,723*	*59,700*
311,253	**747,608**	**946,198**	**813,221**	**531,310**	**391,814**	*366,076*	*225,754*	*345,981*	*83,044*
343,043	**823,968**	**1,042,841**	**896,282**	**585,577**	*431,833*	*403,467*	*248,812*	*381,319*	*91,526*
384,679	**923,974**	**1,169,412**	**1,005,065**	*656,650*	*484,245*	*452,436*	*279,011*	*427,600*	*102,635*
444,666	**1,068,059**	**1,351,772**	*1,161,796*	*759,049*	*559,759*	*522,990*	*322,520*	*494,280*	*118,640*
400,741	**962,553**	*1,218,240*	*1,047,030*	*684,067*	*504,464*	*471,327*	*290,660*	*445,454*	*106,920*
344,014	*826,299*	*1,045,792*	*898,818*	*587,234*	*433,055*	*404,608*	*249,516*	*382,397*	*91,785*

Demonstration II

- Difference to previous example: the two data points in column 6 excluded for demo I
- Minimum value for fitted values is 67,948
- Lower bound for deviance residuals is therefore -368.64 = $(2*67{,}948)^{0.5}$ [derived in paper]
- Unscaled deviance residual of -530.16 for cell (3,6) is below this bound [equation 3.7 in paper]
- Unable to rescale residual

What do we learn?

- Limited scope of "distribution free" resampling
- Reconsider parametric bootstrapping
  - Makes distributional assumptions
  - Avoids inconsistencies with model
  - Still captures correlations among parameter estimates that are difficult to calculate explicitly
- Further research into "robust" resampling schemes is required

Contact Information

- Spreadsheet with illustrative implementation of algorithms discussed in call papers is available from author at request
- thomas.hartl@us.pwc.com
- 617-530-7524

Selected References

- Anderson, D., et al., "A Practitioner's Guide to Generalized Linear Models—A CAS Study Note"
- Davison, A.C, and D.V. Hinkley, "Bootstrap Methods and Their Application"
- England, P.D., and R.J. Verrall, "Predictive Distributions of Outstanding Liabilities in General Insurance"
- McCullagh, P., and J.A. Nelder, "Generalized Linear Models"
- Pinheiro, Paulo J R, et al., "Bootstrap Methodology in Claim Reserving"

- PLEASE REFER TO FULL BIBLIOGRAPHIES IN CALL PAPERS