Michael R. Larsen
Director & Actuary
September 20, 2016

Stochastic Reserve Model Using a Linear Mixed Effect Model

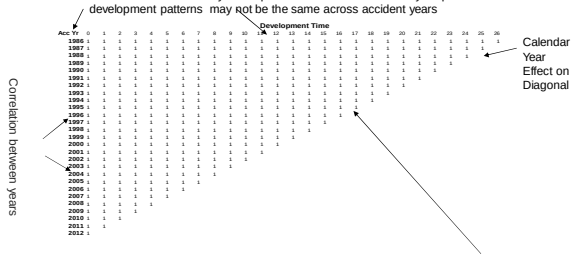
Agenda

- Why use a Mixed Model?
  - Reserve Modeling Problem
  - Mixed Model Description
  - Mixed Model Application to Reserve Model
- Stochastic Reserve Model Example
  - Analysis Flow Chart
  - Data Description
  - Exploratory Analysis
  - Model Using SAS PROC MIXED
  - Residual Review
  - Construction of Reserve Distribution

Why Use a Mixed Model?

Modeling Loss Triangle Issues

Non-constant variance by development time or correlation may be present and development patterns may not be the same across accident years



Three time dimensions, but any point on triangle uniquely determined by two leading to singular matrix-multi-collinearity problem

Triangle is only a sample of all accident years, some form credibility weighting may be needed

4

What is a Mixed Model?

- Generalized Linear Model with added features
- Includes Fixed & Random effects
  - Fixed effect completely contains all possible variable levels
  - Random effect recognizes one has a sample of possible values for a variable
- Uses Two Covariance Matrixes
  - Fixed & Random each have own matrix (matrices are interconnected with each other)
  - Iterative approach to solve: freeze one and optimize other until convergence
- Matrix structure options
  - Error Correlation options
  - Variance Modeling options

5

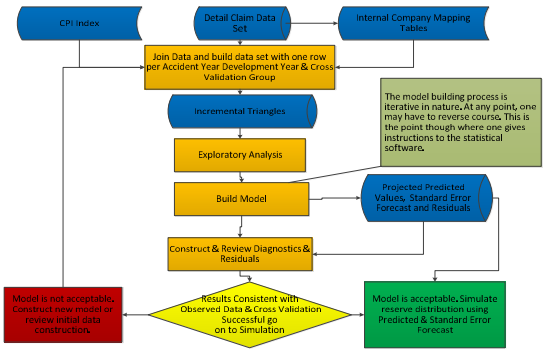
Mixed Model Features Applied to Reserving

- Fixed & Random effects
  - Random Effects designation induces credibility weighting
  - Classifying accident year as random effect eliminates multi-collinearity
  - Interaction and nested effects available to handle changing patterns
  - Interaction with Random Effects induces credibility weighted result
- Covariance Matrixes
  - Iterative solution automatically blends variance modeling with mean estimates by development, calendar and accident year effects
  - Range of options for variance modeling by development period
  - Range of options for correlation in errors by time

6

Stochastic Reserve Model Example

Analysis Flow Chart



 8

Data Description

Data Transformation

- Incremental Dollars to reduce correlation
- Normalize data to remove known trends
  - Exposure: Created Counts by accident year
  - Constant Dollars: CPI by calendar year
- Natural log transform to move data to lognormal scale and model as Normal distribution
- Insert small value for missing observations
- Assign claims to cross validation groups

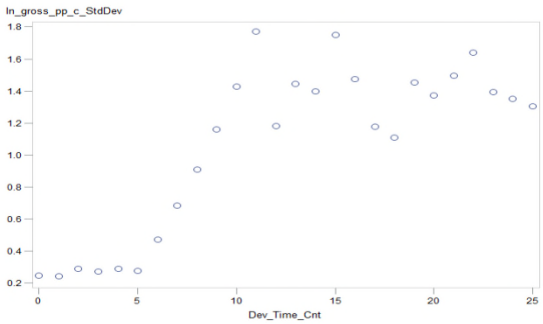
Model Description

- Three explanatory variables:
  - Calendar time
  - Accident Year
  - Development Year
- Explanatory variable categories:
  - Fixed Effects
    - Calendar Time
    - Development Year
    - Accident Year Groups
  - Random Effect
    - Accident year
- Dependent Variables
  - Sigma: Log-Linear Dispersion
  - Mu: Linear Mixed Model prediction

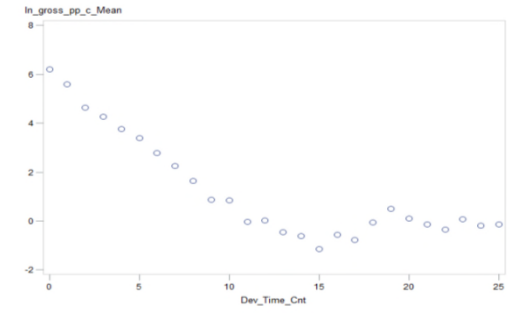
 9

Exploratory Data Analysis

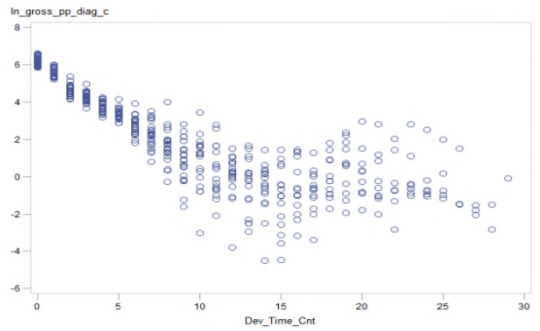
Standard Deviation of Natural Log Pseudo Pure Premium (Incremental Amount/Created Counts) All Accident Years After Normalization



Average of Natural Log Pseudo Pure Premium All Accident Years After Normalization

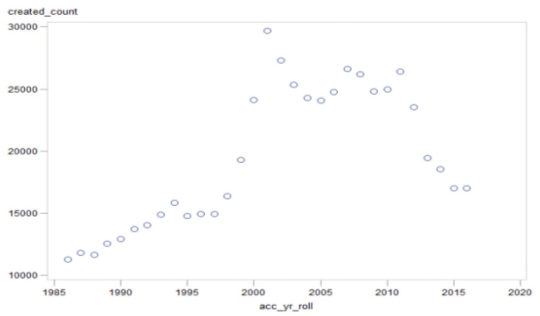


Natural Log Pseudo Pure Premium All Accident Years After Normalization



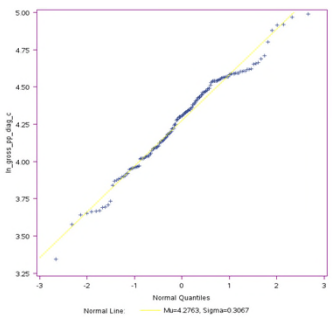
13

Exposure Across Accident Years (Created Claim Count at 12 Months)



14

Q/Q Plot to Evaluate Lognormal Assumption



15

Example of Linear Mixed Model Application

Model Structure

Mean Predicted Pure Premium

- Graphs indicate three time periods for development years:
  - Categorical type effect first three time periods (C_Dev_Time_2)
  - A quadratic effect for time periods 3 through 10 with spline at age 10 (C_Dev_Time_10)
  - Smaller slope after time 10 to ultimate (spline with Dev_Time_Cnt)
- Shift in accident years after 2005
- Cal_yr_time included as continuous variable: (calendar year –first calendar year in data)

Standard Error for Pure Premium

- Small standard error initially
- Increase in standard error stops at about age 10
- Model using log-linear dispersion approach:
  - Residual *exp(C_Dev_Time_10)
- Include accident year random effect variance
- No correlation effect

.

PROC MIXED Code to Produce Example

```
PROC MIXED DATA = WORK.SORTTempTableSorted COVTEST ASYCORR PLOTS(ONLY)=ALL METHOD=REML ;

CLASS cal_yr acc_yr_used acc_yr_roll Variance_Group_G Cal_Yr_Grp C_Dev_Time_20
C_Dev_Time_1 C_Dev_Time_2 C_Dev_Time_3 C_Dev_Time_4 acc_yr_grp acc_yr_grp_2 ;

BY cross_var;
```

Set up Cross-Validation with BY

```
MODEL ln_gross_pp_c=    cal_yr_time   Dev_Time_Cnt   acc_yr_grp_2
                        C_Dev_Time_10   C_Dev_Time_10*C_Dev_Time_10
                        C_Dev_Time_2 /
RESIDUAL VCIRY
HTYPE=3
SOLUTION
CL
ALPHA=0.05
DDFM=KENWARDROGER
INTERCEPT
```

Define Model

Use Kenward Rodger Degrees of Freedom

```
E3
OUTPM=WORK.MEAN_PP_TRAIN_GROSS(LABEL="Predicted means data set for WORK.TRAIN_GROSS_NEW" DROP=__FLAG)
OUTP=WORK.PRED_PP_TRAIN_GROSS(LABEL="Predicted values data set for WORK.TRAIN_GROSS_NEW" DROP=__FLAG);
```

Define Random & Fixed Effect Covariance Structure

```
RANDOM int /  CL subject=acc_yr_roll  ALPHA=0.05 TYPE=VC SOLUTION;

REPEATED  / SUBJECT=acc_yr_roll 'LOCAL =EXP(C_DEV_TIME_10 )  TYPE=VC ;
```

Variance and Fixed Effect Parameter Estimates

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Intercept	acc_yr_roll	0.000123	0.001034	0.12	0.4528
Residual		0			
EXP C_Dev_Time_10		0.5496	0.02141	25.67	<.0001
Residual		0.00453	0.000828	5.38	<.0001

Note both variance and the mean are modeled.

C_Dev_Time_X has constant value X after reaching age X

Solution for Fixed Effects

Effect	acc_yr_grp_2	C_Dev_Time_2	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept			5.0898	0.1365	259	37.28	<.0001	0.05	4.8209	5.3586
cal_yr_time			0.01914	0.002391	26.8	8.01	<.0001	0.05	0.01423	0.02405
Dev_Time_Cnt			-0.05941	0.01997	271	-2.97	0.0032	0.05	-0.09872	-0.02009
acc_yr_grp_2	GE_2004		0.1457	0.04128	25.5	3.53	0.0016	0.05	0.0608	0.2307
acc_yr_grp_2	lt_2004		0	.	.	.	.	.	.	.
C_Dev_Time_10			-0.1481	0.04861	299	-3.05	0.0025	0.05	-0.2437	-0.0524
C_Dev_Tim*C_Dev_Time			-0.02552	0.004646	406	-5.49	<.0001	0.05	-0.03465	-0.01639
C_Dev_Time_2		1	0.95	0.09372	229	10.14	<.0001	0.05	0.7653	1.1347
C_Dev_Time_2		2	0.6107	0.05759	170	10.6	<.0001	0.05	0.497	0.7244
C_Dev_Time_2		3	0	.	.	.	.	.	.	.

Random Effect Results (subset of Accident Years Shown)

Solution for Random Effects

Effect	acc_yr_roll	Estimate	Std Err Pred	DF	t Value	Pr > |t|	Alpha	Lower	Upper
Intercept	1986	-0.00308	0.02486	1	-0.12	0.9216	0.05	-0.3189	0.3128
Intercept	1987	0.001655	0.02517	1	0.07	0.9582	0.05	-0.3182	0.3215
Intercept	1988	0.00126	0.02545	1	0.05	0.9685	0.05	-0.3221	0.3246
Intercept	1989	-0.00037	0.0257	1	-0.01	0.9908	0.05	-0.3269	0.3262
Intercept	1990	-0.0003	0.0259	1	-0.01	0.9926	0.05	-0.3293	0.3287
Intercept	1991	-0.00258	0.02604	1	-0.1	0.937	0.05	-0.3334	0.3283
Intercept	1992	0.000774	0.02614	1	0.03	0.9811	0.05	-0.3314	0.3329
Intercept	1993	-0.00033	0.0262	1	-0.01	0.9921	0.05	-0.3332	0.3326
Intercept	1994	-0.00053	0.02622	1	-0.02	0.9871	0.05	-0.3337	0.3326

Estimates are adjustments to Intercept and are credibility weighted using what actuaries would call least squares credibility.

Diagnostics & Residual Review

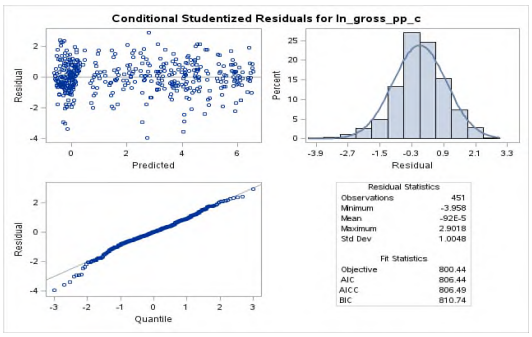
Standard Diagnostics

- Are residuals roughly Normal? (Can we safely use the t & F tests?)
- Have we dealt with non-constant variance successfully? (residuals vs. predicted chart)
- Are there variables which should be dropped? Type 3 tests

Residual & Expected Dollar Graph Review

- Are residuals roughly centered around zero for three time dimensions?
- Are incremental dollars actual vs. predicted roughly in synch?
- Do projected dollars look plausible?

Analysis of Residuals for Normality and Success on Modeling Variance



Overall Model Fit Diagnostics

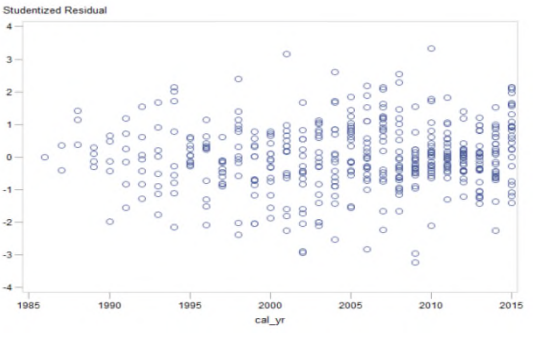
Fit Statistics	
-2 Res Log Likelihood	810.6
AIC (Smaller is Better)	816.6
AICC (Smaller is Better)	816.7
BIC (Smaller is Better)	820.9

AICC useful when comparing across models, since models need not be subsets of one another

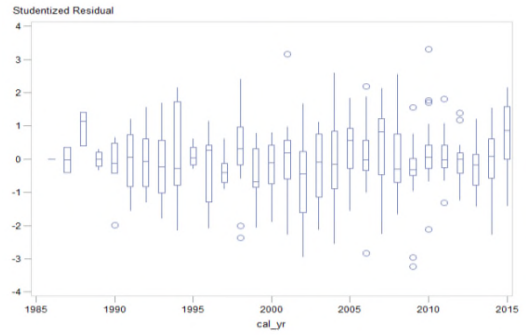
Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Intercept	1	258	3512.91	<.0001
cal_yr_time	1	26.8	64.1	<.0001
Dev_Time_Cnt	1	271	8.85	0.0032
acc_yr_grp_2	1	25.5	12.46	0.0016
C_Dev_Time_10	1	299	9.28	0.0025
C_Dev_Tim*C_Dev_Time	1	406	30.18	<.0001
C_Dev_Time_2	2	144	57.21	<.0001

Type 3 tests provide useful inter model comparison to see if a term should be dropped.

Scatter Plot of Calendar Year Residuals

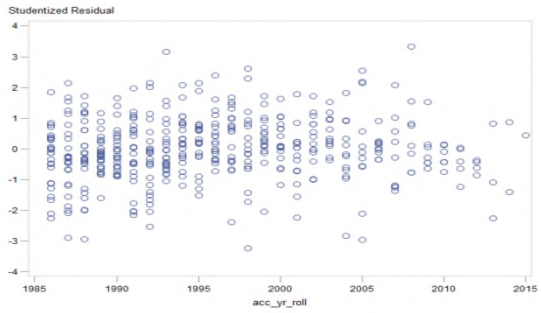


Box Plot of Calendar Year Residuals



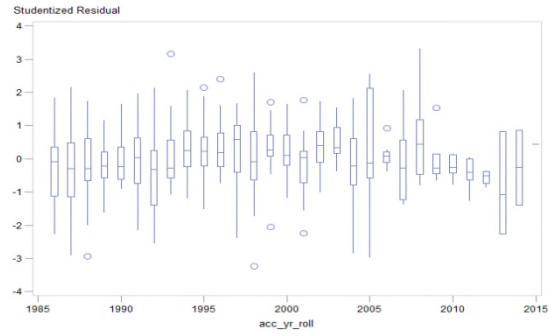
25

Scatter Plot of Residuals by Accident Year



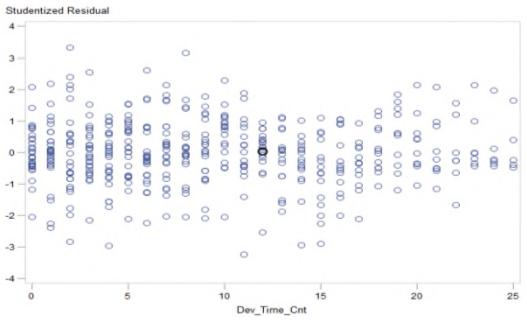
26

Box Plot of Accident Year Residuals

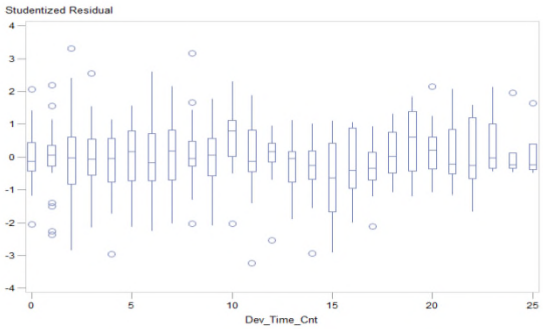


27

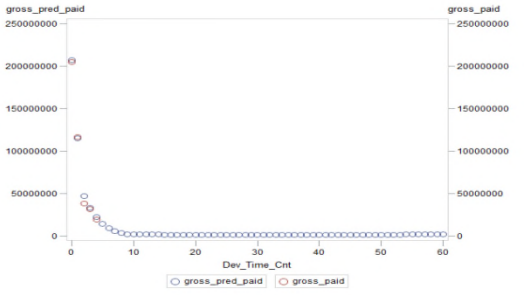
Scatter Plot of Residuals by Development Year



Box Plot of Residuals by Development Year



Actual to Projected in Real Dollars Example



Excerpt from Simulation Code to Build Reserve Distribution

```
G_chi_rand = RAND('CHISQUARE',G_DF);

G_StdErr_R = G_StdErrPred *((G_DF-
1)**.5)/( G_chi_rand**.5);

G_PP_Log = RAND
('NORMAL',G_Pred,G_StdErr_R);

G_PP= exp(G_PP_Log +ln_cpi_change) ;
```

Use Chi-Square distribution to describe standard error distribution. G_DF is the degrees of freedom assigned by SAS PROC MIXED. G_StdErrPred is the standard error of forecast.

G_Pred is the forecast pseudo pure premium on log scale. Use continuous inversion to pull one observation from distribution around mean predicted then invert from lognormal to real dollars accounting for CPI change originally removed. Have to multiply by created counts to derive paid dollars.

Repeat process many times to build empirical distribution.

CLOSING SLIDE

Conclusion

- Reserve models differ from class plan models
  - Typically, variance is not a simple function of the mean and is not constant
  - Easy to over fit (need cross validation)
  - Forecast error may indicate simpler model is better
- Software has improved
- Numerous Mixed Model textbooks (with examples included)
- CAS has published articles on regression models applied to reserving
- You can do this

Acknowledgment

- "Best Estimate for Reserves", Casualty Actuarial Society Forum Casualty Actuarial Society - Arlington, Virginia,1998: Fall, Pages 1-54, Ben Zehnwirth
- "Stochastic Loss Reserving Models Using Generalized Linear Models", Casualty Actuarial Society Monograph Series Number 3, 2016, Taylor, Greg and McGuire, Graine
- "Hierarchical Growth Curve Models for Loss Reserving", Casualty Actuarial Society Forum Casualty Actuarial Society - Arlington, Virginia, 2008: Fall, Pages 146-173, James C. Guszcza
- SAS for Mixed Models (Second Edition), SAS Institute Inc., 2006, Littell, Ramon, C.
- Generalized, Linear and Mixed Models, John Wiley & Sons, 2008, McCullough, Charles E.
- Generalized Linear Mixed Models Modern Concepts Methods and Applications, Chapman & Hall/ CRC Texts in Statistical Science, 2012, Stroup, Walter, W.