9th Webinar on General Insurance

Stochastic Reserving in Practice

JV Prasad,
Palreddy Vishnuvardhan,
Subhasree Nigamma

Institute of Actuaries of India

Agenda

- Introduction to stochastic claims reserving
- Illustration of ODP Bootstrap including B-F and Cape Cod approaches
- ODP Model
- Sample diagnostics & Practical considerations
- Q&A

Credits/ Acknowledgements

- The contents of this presentation are based extensively on “Using the ODP Bootstrap Model: A Practitioner’s Guide by Mark R. Shapland” Casualty Actuarial Society, Monograph Series No. 4

- However, any errors or omissions are exclusively those of the presenters

Agenda

- Introduction to stochastic claims reserving
- Illustration of ODP Bootstrap including B-F and Cape Cod approaches
- ODP Model
- Sample diagnostics & Practical considerations
- Q&A

Claims reserving in India

- Traditionally actuaries in India have focused on a deterministic point estimate in estimating their ultimate claim liabilities
  - Some exceptions to this exist
- Regulations and practice standards of IAI and IRDAI expect the valuation to be prudent but do not insist on quantification of this "prudence"
  - At present ALSM 2016 is silent on any margin or extent of prudence in the estimation of claim reserves
  - As per para 8.4 of APS 21, "The Appointed Actuary should allow for margin for adverse deviation, as applicable, to reflect the degree of confidence the Appointed Actuary has in the expected level of the parameter and his / her perception about the extent of such deviation"

Need for claims distribution

- Significant judgements in applying IFRS 17
  - **(119)** *“An entity shall disclose the confidence level used to determine the risk adjustment for non-financial risk. If the entity uses a technique other than the confidence level technique for determining the risk adjustment for nonfinancial risk, it shall disclose the technique used and the confidence level corresponding to the results of that technique”*
- Proposed Indian Risk Based Capital Framework
- Providing additional comfort to regulator/ shareholder/ investment analysts
- Building a systematic framework for claims reserving

Poll Question 1

With respect to “modeling” claims which approach would you prefer:

a. Using closed form distributions like Lognormal
b. Using regression-based approaches like GLM
c. Using nonparametric methods like bootstrap
d. I am not sure

Stochastic claims reserving

- Claims reserving process depends on estimating the extent of claim "yet to be paid" based on historical patterns
  - Estimated "unpaid factor" applied on cumulative payments till date for a particular cohort (e.g. Accident Year)
- Therefore, the purpose of developing a model for "claims reserving" is to quantify the variability inherent in the claims settlement process
  - That is the variability around the point estimate of "unpaid factor"
- As such a direct model of claims cost (such as lognormal model of claims) would not suffice the purpose since it captures inherent randomness in claims (variability in frequency and severity)

Stochastic claims reserving

- A more appropriate model would be one based on regression techniques (specifically Poisson regression)
  - Which would model the incremental claims for each cohort year for every year of future development
- In simple terms a Poisson regression model (with log link function) can be fitted to incremental claims data where the accident year and development year are predictors
- From the model fitted on the data, estimates of incremental claims can be obtained
- In ODP since mean of the predicted response is equal to sample mean of the response this approach yields same results as chain ladder method

Agenda

Institute of Actuaries of India

- Introduction to stochastic claims reserving
- Illustration of ODP Bootstrap including B-F and Cape Cod approaches
- ODP Model
- Sample diagnostics & Practical considerations
- Q&A

ODP Bootstrap approach

- To avoid sampling from a GLM model using the feature explained in the previous slide the ODP Bootstrap approach has been developed
- This is an algorithmic approach which has the following advantages:
  - It uses development factors and hence can serve as a bridge between the deterministic estimation of ultimate losses and the stochastic model
  - This makes it easier to relate to and explainable to external audience
  - It overcomes limitations of a GLM model like negative increments

Bootstrap approach[1]

- Bootstrapping is a form of nonparametric Monte Carlo methods that estimate the distribution of a population by resampling from the data
- These methods are often done when the sample is the only information available, and the distribution is unknown.
- The objective is to repeatedly resample from the data itself to get a better idea of the distribution or the parameter of interest

1. Analysis Of Bootstrap Techniques For Loss Reserving; A Thesis Submitted to the Graduate Faculty of the North Dakota State University of Agriculture and Applied Science By Taryn Ruth Chase

The ODP Bootstrap Algorithm

- Take a typical accident year triangle as set out below:

AY/DY	**1**	**2**	**3**
1	$c_{1,1}$	$c_{1,2}$	$c_{1,3}$
2	$c_{2,1}$	$c_{2,2}$	
3	$c_{3,1}$		

- From this upper triangle of observed cumulative claim amounts, the all year weighted average development factors can be estimated (w: being accident year and d: development year)

$$\hat{F}(d) = \frac{\sum_{\mathrm{w}=1}^{n-d+1} c(\mathrm{w},\mathrm{d})}{\sum_{w=1}^{n-d+1} c(\mathrm{w},\mathrm{d}-1)}$$

The ODP Bootstrap Algorithm

- Estimating the "cumulative fitted triangle"
  - The final diagonal of observed values are kept the same
  - The remaining diagonals are found by recursively dividing the fitted values at time "d" by development factors at time "d-1"
  - *Instead of estimating the incremental values in the lower right quadrant multiplying the latest diagonal by prospective development factors; we recreate the upper left quadrant by dividing the latest diagonal by the preceding development factors*
  - From the "cumulative fitted triangle" the "incremental fitted triangle" is obtained by differencing
  - Note: If all year weighted average DF is used the result will be identical to GLM with Poisson link function
  - From the "incremental fitted triangle" estimated above the residuals are derived

The ODP Bootstrap Algorithm

- Estimating the residual
  - It is representation of variance (process and model)
  - Various authors have suggested use of "Pearson residuals"
  - Unscaled Pearson residuals: $r = \frac{Observed - Fitted}{\sqrt{Fitted}}$
- Scale factor : $\emptyset = \frac{\sum r^2}{n-p}$
  - Where n is number of observations in the triangle
  - p is number of parameters to be estimated; typically 2 * number of years - 1
  - And summed over the n residuals over the upper triangle
  - Used in bringing process variance into projected unpaid after bootstrapping
- Other versions of residuals: Scaled Pearson Residuals, Standardized Pearson Residuals, Standardized Hetero-Adjusted Pearson Residuals

The ODP Bootstrap Algorithm

- Resampling the residual
  - Once the residuals "r" are estimated they need to be sampled with replacement
  - This can be done by ordering the residuals and using a random number generator to sample the residual consistent with its position
  - Once the upper left quadrant is populated with resampled residuals
    - The incremental values can be determined by
    - $r * \sqrt{Fitted} + Fitted = Observed$ *
  - This produces the "pseudo-triangle"
  - This can be converted into cumulative triangle
  - Using Chain ladder method, the future claims can be estimated
  - From this determine the estimate of unpaid claims

The ODP Bootstrap Algorithm

Institute of Actuaries of India

- Process Variance: When expected future incremental claims are being projected for each accident & development period,
  - assume that each future expected incremental value is the mean and the mean times the scale parameter is the variance of a gamma distribution
  - generate a random value from the resultant gamma distribution
- The Simulation:
  - The process set out in the preceding two slides can be repeated any number times to get "N" estimates of unpaid claims
    - Where N is the number of simulation runs
  - Values from the simulation will be the distribution of unpaid claims
- Various diagnostics and checks can be undertaken on the residuals which are discussed here subsequently

The ODP Bootstrap Algorithm

- Paid vs Incurred triangles
  - If incurred triangle is used, then it will result in the distribution of IBNR and not unpaid: this may not be desirable if the requirement is to assess the variance of unpaid and not just IBNR
  - However, it may be desirable to use Incurred figures as opposed to paid in certain situations
  - In such a case one alternative is to:
    - Run paid and incurred ODP bootstrap
    - Take the payment pattern from every iteration of paid bootstrap
    - Take the ultimate value from every iteration of incurred bootstrap
    - Divide the two to get an estimate of unpaid and run N times

Poll Question 2

During reserving for the recent accident years especially for long tailed LoB like TP, which method do you use?

- Bornheutter Ferguson Method
- Frequency Severity Method
- ULR
- Chain Ladder method
- Any other

The ODP B-F bootstrap

- In CL method, the estimate of Ultimate Claims depends greatly on the value of the paid or incurred till date.
  - A problem if CDF is very high, e.g., Paid Motor TP for recent AYs
- In the event the estimate of ultimate incorporates the Bornheutter-Ferguson method; the ODP bootstrap may need to be modified
  - The ODP part of the simulation remains as such
  - From the pseudo-triangle estimate the development factor for each simulation run
  - This provides the percent-unpaid estimate for the B-F method for each of the N simulation runs

The ODP B-F bootstrap

- One needs to estimate parameters of the a-priori loss ratio used in the B-F estimation process
  - Consider any (say LogNormal) distribution with appropriate parameterisation
  - Obtain a random sample of a-priori loss ratio for relevant AYs for every run of the simulation
- Obtain estimate of unpaid by blending the a-priori loss ratio with the unpaid obtained from ODP simulation
- This can be repeated for N simulation runs to get B-F ODP distribution of unpaid claims
- In case of changes to mix/premium rates this can be modified using Cape Cod method

Agenda

Institute of Actuaries of India

- Introduction to stochastic claims reserving
- Illustration of ODP Bootstrap including B-F and Cape Cod approaches
- ODP Model
- Sample diagnostics & Practical considerations
- Q&A

The ODP model

- Over-Dispersed Poisson (ODP) Model
  - A special case of GLMs
  - Log-link function
  - Error distribution is Poisson allowing for "over-dispersion"
- ODP Distribution
  - Assume Y is a random variable. Y has ODP distribution, if there exists a value φ such that Y/φ follows Poisson with parameter = $(E[Y] / \varphi)$
  - For a Poisson Distribution with parameter λ, both Mean & Variance = λ
  - Therefore, $Var[Y] = \varphi * E[Y]$ ($\varphi > 1$, if over-dispersion)

ODP Model Specifications

- Fit an ODP model with Incremental claims as response variable:
  - Log of Expected value of Incremental claims for i^{th} accident year, in j^{th} development period = $\alpha_i + \beta_j$, where
    - “i” ranges from 1 to n, where n is number of origin (typically, accident) periods
    - “j” ranges from 2 to n (if origin and development periods have same duration)
  - Variance of Incremental claims is φ times expected value
- May use R to fit the model
  - To get parameter estimates for α’s and β’s (and also for φ)
  - Using Maximum Likelihood Estimation (MLE) method
- Using estimates for α’s and β’s, project future **expected** incremental claims
  - It matches with Chain Ladder method, where age-to-age development factors are based on Column Sum Averages

ODP model in R

Sample Dataset:

- a 10x10 triangle
- The records in this table pertain to 1st development period & two last periods
- “AY” and “DY” are accident period and development period variables.
- Import this dataset with AY & DY as Categorical variables (to get parameter estimates for all “levels”)

AY	DY	Incremental Claim
2014	1	4,768
2015	1	5,367
2016	1	6,466
2017	1	5,851
2018	1	4,953
2019	1	5,556
2020	1	12,325
2021	1	7,097
2022	1	15,383
2023	1	15,384
…………………………		
…………………………		
2014	9	10,238
2015	9	15,280
2014	10	11,156

```
install.packages(“statmod”)
library(statmod)

Model_name <- summary(glm(formula = Incremental_Claim ~ AY + DY , family = tweedie(var.power = 1.0, link.power = 0), data = dataset_name))

Summary(Model_name)

fitted(Model_name)
```

- ODP is a special case of Tweedie family of distributions (when p = 1, denoted by “var.power”)
- “Link.power = 0” indicates log link function
- “fitted” function fits values for triangle (past, not future): in turn residuals

ODP model in R

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.50896    0.18317  46.454  < 2e-16 ***
AY2015       0.06981    0.12750   0.548 0.587400
AY2016       0.18186    0.12717   1.430 0.161315
AY2017       0.20820    0.12942   1.609 0.116414
AY2018       0.23628    0.13263   1.781 0.083281 .
AY2019       0.44662    0.13170   3.391 0.001702 **
AY2020       0.65065    0.13411   4.852 2.36e-05 ***
AY2021       0.64615    0.15031   4.299 0.000125 ***
AY2022       0.91933    0.17771   5.173 8.84e-06 ***
AY2023       1.13212    0.37428   3.025 0.004570 **
DY10         0.81077    0.42480   1.909 0.064314 .
DY2          1.79534    0.16793  10.691 1.01e-12 ***
DY3          2.01028    0.17008  11.820 5.95e-14 ***
DY4          1.90514    0.17521  10.874 6.33e-13 ***
DY5          1.69990    0.18454   9.212 5.32e-11 ***
DY6          1.42473    0.20015   7.118 2.30e-08 ***
DY7          1.20616    0.22084   5.462 3.64e-06 ***
DY8          0.96577    0.25703   3.757 0.000608 ***
DY9          0.90952    0.30565   2.976 0.005197 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Tweedie family taken to be 1638.901)

    Null deviance: 679958  on 54  degrees of freedom
Residual deviance:  63560  on 36  degrees of freedom
AIC: NA
```

AY	DY	Incremental	Fitted
2014	1	4,768	4,959
2015	1	5,367	5,318
2016	1	6,466	5,948
2017	1	5,851	6,107
2018	1	4,953	6,281
2019	1	5,556	7,751
2020	1	12,325	9,505
2021	1	7,097	9,463
2022	1	15,383	12,435
2023	1	15,384	15,384
…………………………			
…………………………			
2014	9	10,238	12,314
2015	9	15,280	13,204
2014	10	11,156	11,156

Scale Parameter estimate (1638.901)

Example: Projection of **Expected** Incremental Claims for AY2020 in DY6 = exp (8.50896 + 0.65065 + 1.42473) = 39511

Agenda

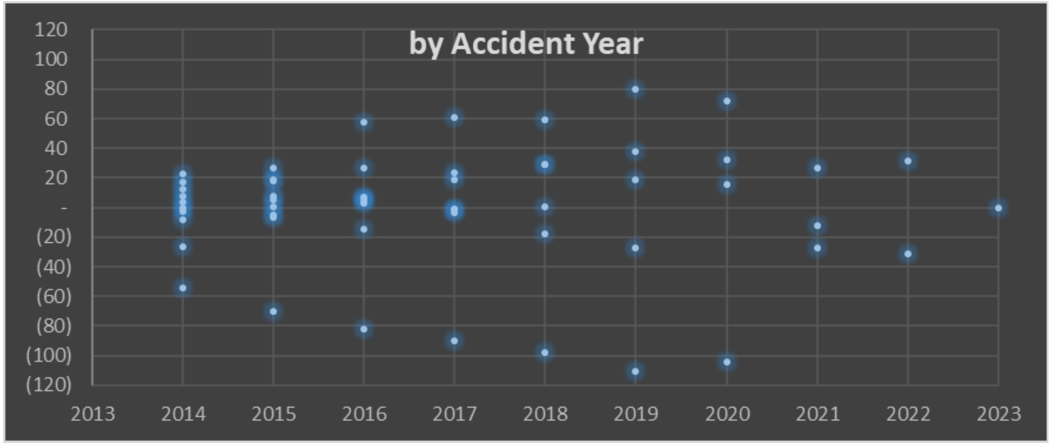
- Introduction to stochastic claims reserving
- Illustration of ODP Bootstrap including B-F and Cape Cod approaches
- ODP Model
- Sample diagnostics & Practical considerations
- Q&A

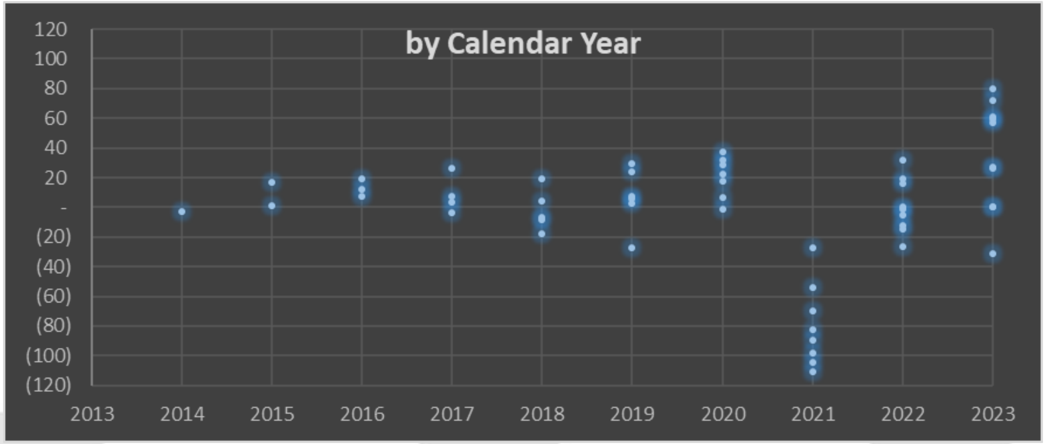
Residual Graphs

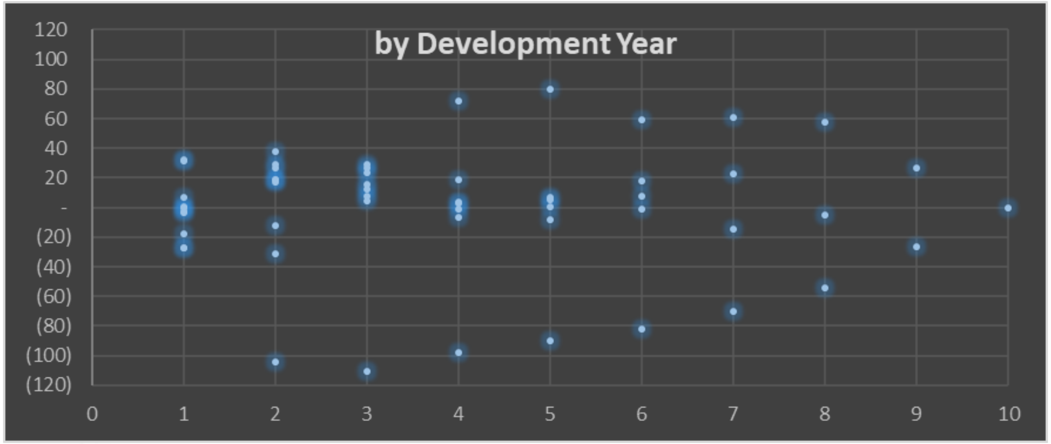
- Starting point is Unscaled Pearson Residuals
- These residuals are then "Standardized"
- Do they appear Independent & Identically distributed?
- Plot Standardized Pearson Residual by Origin period / Development period / Calendar period
- Check for unbiasedness and homoscedasticity
- Check for Outliers
- "Hetero-adjustment" to overcome heteroscedasticity

Standardized Pearson Residuals

This dataset is representative of Motor TP Paid Claims data. This is not specific to any particular insurer. Compiled from NL-38 public disclosure of a few insurers for FY2022-23, and a few assumptions on 1st development year. There is a clear calendar year effect in 2021.







ODP model with CY Effect

Model_name2 <- summary(glm(formula = Incremental_Claim ~ AY + DY + CY_EFFECT , family = tweedie(var.power = 1.0, link.power = 0), data = dataset_name))

Summary(Model_name2)

fitted(Model_name2)

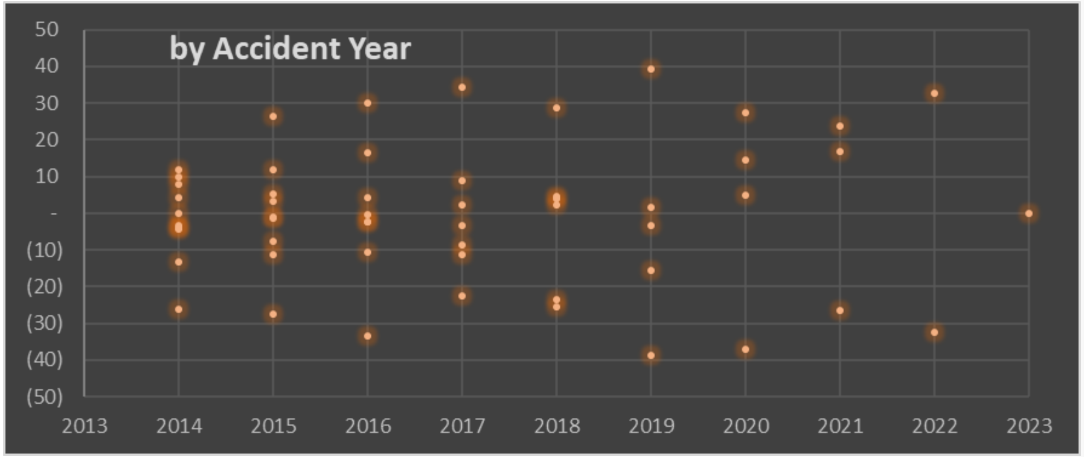
"CY_EFFECT" variables is YES for FY2020-21 and NO for others. Model is estimating Claims Paid in calendar year FY2020-21 to be lower by 44% relative to other years.

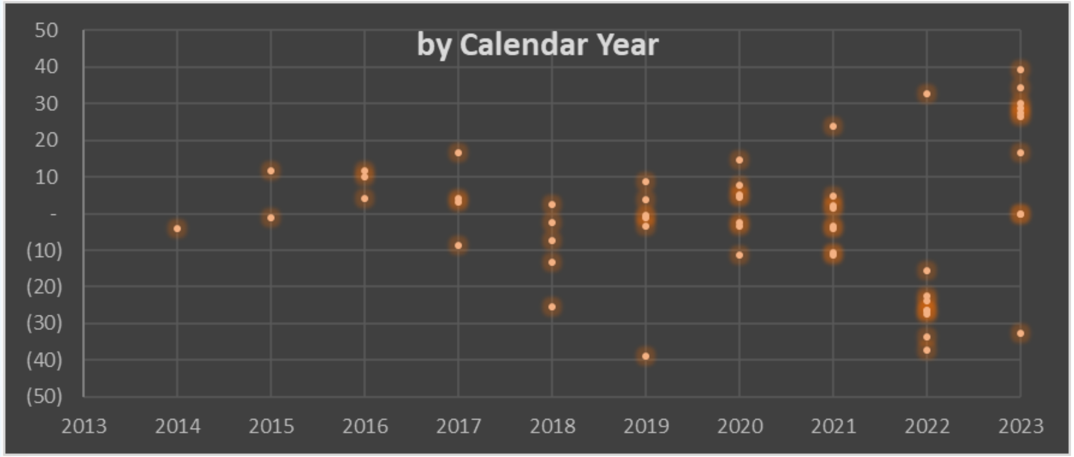
```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    8.52740    0.08047 105.964  < 2e-16 ***
AY2015         0.07713    0.05634   1.369 0.179661
AY2016         0.19870    0.05621   3.535 0.001169 **
AY2017         0.25355    0.05697   4.451 8.32e-05 ***
AY2018         0.31401    0.05856   5.362 5.36e-06 ***
AY2019         0.56356    0.05851   9.632 2.24e-11 ***
AY2020         0.77550    0.05968  12.995 5.92e-15 ***
AY2021         0.65969    0.06610   9.979 8.95e-12 ***
AY2022         0.89418    0.07814  11.443 2.22e-13 ***
AY2023         1.11368    0.16482   6.757 7.89e-08 ***
DY10           0.79233    0.18711   4.235 0.000158 ***
DY2            1.80315    0.07414  24.321  < 2e-16 ***
DY3            2.00171    0.07507  26.665  < 2e-16 ***
DY4            1.88292    0.07696  24.467  < 2e-16 ***
DY5            1.70696    0.08109  21.051  < 2e-16 ***
DY6            1.46855    0.08808  16.672  < 2e-16 ***
DY7            1.27814    0.09733  13.132 4.36e-15 ***
DY8            1.08164    0.11351   9.529 2.95e-11 ***
DY9            0.88728    0.13455   6.594 1.28e-07 ***
CY_EFFECTYES -0.57886    0.04833 -11.977 6.18e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Tweedie family taken to be 318.3081)

    Null deviance: 679958  on 54  degrees of freedom
Residual deviance:  11299  on 35  degrees of freedom
AIC: NA
```

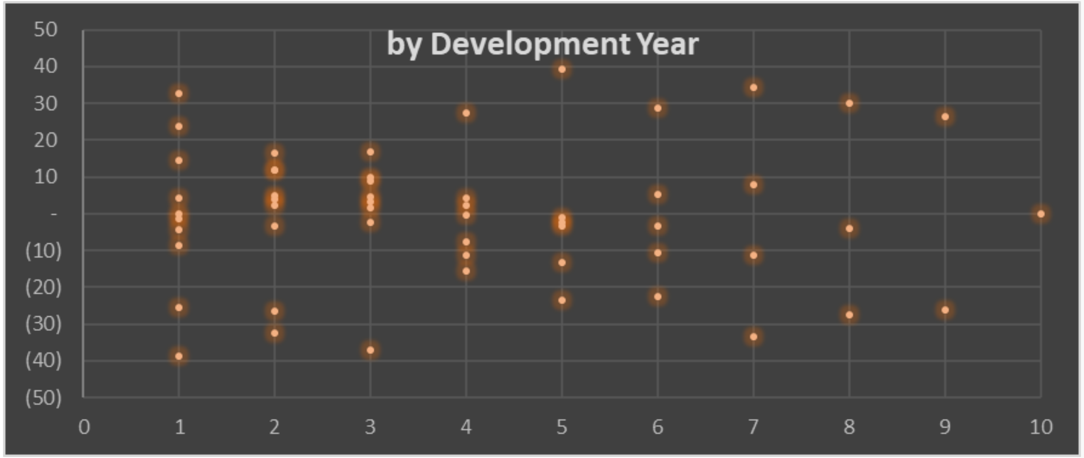
Institute of Actuaries of India

Standardized Pearson Residuals

The residuals are computed again, after including CY_EFFECT variable in the model. The values of residuals reduced sharply. Please note that the scale of the axis is different from previous graphs..







Other Aspects:

(a) Any CY effect in 2022 & 2023?

(b) Exclude CPA claims & check again

Poll Question 3

When there were extra ordinary circumstances (like pandemic), were there any adjustments made to incurred claims while projecting ultimate claims using chain ladder method for Motor TP LoB?

a. Incurred claims were loaded before projection

b. Prudent development factors were used

c. Other methodologies were used for arriving at ultimate claims – Frequency Severity etc and avoided using incurred claims data.

d. Any other

Practical Considerations

- Do Calendar Year effects impact results?
  - Calendar year such as one illustrated in previously slides in Motor TP Claims Paid could lead to lower Best Estimate and greater variability than actuals, if not addressed
- Chain Ladder for Recent Accident Years in Motor TP Claims Paid?
  - Estimate of Ultimate Claims depends under CL method greatly on the value of the paid or incurred till date which can be volatile and CDF can be high
  - Check if difference between Stochastic Best Estimate (SBE) vs. Actuarial Best Estimate (ABE) may be more than, say, 10%
  - Stochastic BF method or Stochastic Cape Cod method may be used.

Practical Considerations

- How to treat tail factor in ODP bootstrapping?

  - Assuming a distribution for tail factor parameter
  - Other considerations such as process variance, hetero adjustment factors etc., can be added to tail factors
  - Instead of simulating a single tail factor, it is appropriate to consider extrapolation of tail factors

Agenda

Institute of Actuaries of India

- Introduction to stochastic claims reserving
- Illustration of ODP Bootstrap including B-F and Cape Cod approaches
- ODP Model
- Sample diagnostics & Practical considerations
- Q&A