GUY CARPENTER

Workers Comp Excess Model

Credibility

Gary Venter

Simplified Version of Least Squares Credibility in General

Basic Setup

- Series of n observations for each of a group of m risks
- Want estimator for each risk's mean
- Each risk's sample mean is one estimator
- Mean of all m risks is another
- How to combine these?

Two Variances

- Within variance = process variance
  - Average over all risks of variance of each risk's own results from a single sample
  - Dividing individual risk variance by sample size gives variance of its sample mean
- Between variance = variance of hypothetical means
  - Variance across risk means
- These give measurements of errors for the two estimators
  - Small within variance makes sample mean a good estimator of risk mean
  - Small between variance makes group mean a good estimator of risk mean

Weighting Observations

- Suppose you have two independent samples X and Y from a population
  - one with variance σ^2/n
  - the other with variance σ^2/m
- Say you want to minimize the variance of $zX+(1-z)Y$,
  - $z^2\sigma^2/n + (1-z)^2\sigma^2/m$
- Set derivative to zero
  - $z\sigma^2/n = (1-z)\sigma^2/m = \sigma^2/m - z\sigma^2/m$
- So… $z = [\sigma^2/m]/[\sigma^2/n+\sigma^2/m] = n/[m+n] = [n/\sigma^2]/[m/\sigma^2+n/\sigma^2]$
- Weight on X is inverse of its variance, same for Y

Weight Sample Mean and Group Mean
by Inverse of Variances

- $$\frac{(\text{Sample mean})(n/\text{Within}) + (\text{Group Mean}) / \text{Between}}{n/\text{Within} + 1/\text{Between}}$$
- Let K = Within / Between, multiply top and bottom by Within
- Sample[n/(n+K)] + Group[K/(n+K)]
- Let Z = n/(n+K)
- (Sample)Z + (Group)(1 – Z)
- This turns out to be minimum variance linear estimator

Apply to Work Comp Classes

Workers Compensation
Excess Pricing Model

- Bureau excess prices traditionally based on hazard groups
- Excess potential - very different across hazard groups
  - but also within hazard groups
- Bureau methodology weights injury-type severity distributions by hazard group injury-type frequency splits
- We do that by class
  - Requires credibility procedure to get class distribution of losses by injury type

Severity by Injury Type, Massachusetts:
Large Loss Potential Is Driven by Fatal, PT

	Fatal	PT	Major	Minor	TT
Mean	$ 411,287	$ 896,725	$ 137,163	$ 15,826	$ 12,367
95th Percentile	$ 1,285,878	$ 2,566,482	$ 307,876	$ 42,187	$ 49,050
Ratio to TT					
Mean	33.3	72.5	11.1	1.3	1.0
95th Percentile	26.2	52.3	6.3	0.9	1.0

Differences in Injury-Type Frequencies Across and Within Hazard Groups:
Ratios to Temporary Total

Hazard group means are very different but significant variation exists within each hazard group

Means			
HG	Fatal:TT	PT:TT	Major:TT
1	0.21%	0.33%	6.10%
2	0.28%	0.44%	7.06%
3	0.69%	0.74%	11.61%
4	1.83%	1.44%	27.27%

95th Percentile Class*		
HG	Fatal:TT	PT:TT
1	0.86%	0.74%
2	0.97%	1.47%
3	2.82%	2.66%
4	4.79%	2.77%

**95th percentile of larger classes*

Correlation of Ratios to TT
Across Classes Hazard Group III

	PT	Major	Minor
Fatal	39%	45%	20%
PT		52%	31%
Major			28%

Use correlations to better estimate class frequencies.

Major predictive of fatal and PT.

Credibility Including Correlation

Credibility with Correlation

- Denote by **V, W, X, Y** - class ratios to **TT** for Fatal, PT, Major & Minor
- Credibility Formula for Fatal for Class **i**:
  - $\mathrm{E}\,v_i + b(V_i - \mathrm{E}\,V_i) + c(W_i - \mathrm{E}\,W_i) + d(X_i - \mathrm{E}\,X_i) + e(Y_i - \mathrm{E}\,Y_i)$
  - Here $\mathrm{E}\,v_i = \mathrm{E}\,V_i$ is the hazard group mean for **Fatal:TT; b** is usual Z
- Example credibilities for fatal for a class in HG III with 300 **TT** claims
  - **b = 32.6%, c = 5.0%, d = 1.3%, e = 0.2%**
- Major frequency - over 15 times fatal
  - so factor of 1.3% is in ballpark of being like 20% for fatal
- Minor frequency - over 50 times fatal
  - so factor of 0.2% has impact of a factor of 10% for fatal
    (assuming differences from mean are of same magnitude as the mean)
- How are these estimated?

Credibility with Correlation

Denote four injury types by *V, W, X,* and *Y*.

Assume that each class has parameters that determine its behavior up to random draws

but…
since the parameters are unobserved they too are considered random variables.

For the i^{th} class, denote the population mean ratios (i.e., the true conditional, or "hypothetical" mean) as v_i, w_i, x_i, and y_i.

Here these are mean ratios to TT.

Notation

$$\sum_{t=1}^{N} m_{it} W_{it} / \sum_{t=1}^{N} m_{it}$$

We observe each class **i** for each time period **t.**

Denote by W_i the class sample mean ratio for all time periods weighted by exposures m_{it} **(TT claims),** where there are **N** periods of observation.

Similarly for V, X, and Y.

Let m_i denote the sum over the time periods t of the m_{it}

m is the sum over classes i of the m_i.
Then within $\mathrm{Var}(W_{it} \mid w_i) = \sigma_{Wi}^2 / m_i$

Assume a linear model and minimize expected squared error, where expectation is taken across all classes in the hazard group.

For PT this can be expressed as minimizing:

$$\mathrm{E}[(\underline{a + bV_i + cW_i + dX_i + eY_i} - w_i)^2]$$

estimate of w_i

expected squared error is like sum of squared errors without the observations

The coefficients sought are a, b, c, d, and e.
Differentiating wrt a gives:

$$a = -\mathrm{E}(bV_i + cW_i + dX_i + eY_i - w_i)$$

Plugging in that for a makes the estimate of $w_i =$

$$\mathrm{E}w_i + b(V_i - \mathrm{E}V_i) + c(W_i - \mathrm{E}W_i) + d(X_i - \mathrm{E}X_i) + e(Y_i - \mathrm{E}Y_i)$$

We have $w_i =$

$$\mathrm{E}\,w_i + b(V_i - \mathrm{E}\,V_i) + c(W_i - \mathrm{E}\,W_i) + d(X_i - \mathrm{E}X_i) + e(Y_i - \mathrm{E}\,Y_i)$$

Since in taking the mean across classes $\mathrm{E}\,w_i = \mathrm{E}\,W_i$, c is the traditional credibility factor **Z.**

The derivative of $\mathrm{E}[(a + bV_i + cW_i + dX_i + eY_i - w_i)^2]$ wrt b gives:

$$a\mathrm{E}\,V_i + \mathrm{E}[V_i(bV_i + cW_i + dX_i + eY_i - w_i)] = 0$$

Plugging in for *a* then yields: $0 =$

$$\mathrm{E}(bV_i + cW_i + dX_i + eY_i - w_i)\mathrm{E}\,V_i + \mathrm{E}[bV_i^2 + cV_iW_i + dV_iX_i + eV_iY_i - V_iw_i]$$

Using $\mathrm{Cov(X,Y) = E[XY] - EXEY}$,this can be rearranged to give:

$$\mathrm{Cov}(V_i, w_i) = b\mathrm{Var}(V_i) + c\mathrm{Cov}(V_i, W_i) + d\mathrm{Cov}(V_i, X_i) + e\mathrm{Cov}(V_i, Y_i)$$

Doing the same for **c, d,** and **e** will yield three more equations that look like (3), but with the variance moving over one position each time.

Thus you will end up with four equations that can be written as a single matrix equation:

$$\begin{pmatrix} \mathrm{Cov}(V_i, w_i) \\ \mathrm{Cov}(W_i, w_i) \\ \mathrm{Cov}(X_i, w_i) \\ \mathrm{Cov}(Y_i, w_i) \end{pmatrix} = C \cdot \begin{pmatrix} b \\ c \\ d \\ e \end{pmatrix}$$

where C is the covariance matrix of the class by injury-type sample means ***Cov(*V_i*,*Y_i*)*** etc.

You need estimates of all covariances - like estimating the EPV and VHM

But…with these you can solve this equation for b, c, d, and e to be used for PT.
Repeat for the other injury types.